Solid Mechanics and its Applications

Dominique François
André Pineau
André Zaoui

Mechanical Behaviour of Materials

Volume 1: Micro- and Macroscopic Constitutive Behaviour

Springer

Mechanical Behaviour of Materials

SOLID MECHANICS AND ITS APPLICATIONS
Volume 180

Series Editors: G.M.L. GLADWELL
Department of Civil Engineering
University of Waterloo
Waterloo, Ontario, Canada N2L 3GI

Aims and Scope of the Series

The fundamental questions arising in mechanics are: *Why?*, *How?*, and *How much?* The aim of this series is to provide lucid accounts written by authoritative researchers giving vision and insight in answering these questions on the subject of mechanics as it relates to solids.

The scope of the series covers the entire spectrum of solid mechanics. Thus it includes the foundation of mechanics; variational formulations; computational mechanics; statics, kinematics and dynamics of rigid and elastic bodies: vibrations of solids and structures; dynamical systems and chaos; the theories of elasticity, plasticity and viscoelasticity; composite materials; rods, beams, shells and membranes; structural control and stability; soils, rocks and geomechanics; fracture; tribology; experimental mechanics; biomechanics and machine design.

The median level of presentation is the first year graduate student. Some texts are monographs defining the current state of the field; others are accessible to final year undergraduates; but essentially the emphasis is on readability and clarity.

For further volumes:
http://www.springer.com/series/6557

Dominique François • André Pineau • André Zaoui

Mechanical Behaviour of Materials

Volume 1: Micro- and Macroscopic Constitutive Behaviour

Springer

Dominique François
École Centrale de Paris
rue des Envierges 13
75020 Paris
France

André Pineau
École des Mines de Paris
ParisTech
Centre des Matériaux UMR CNRS
B. P. 87
91003 Evry Cedex
Academy of Engineering
Avenue F.D. Roosevelt
75008 Paris
France

André Zaoui
Académie des Sciences
23 Quai de Conti
75006 Paris
France
Academy of Engineering
Avenue F.D. Roosevelt
75008 Paris
France

This is the second updated edition of the earlier book entitled Mechanical Behaviour of Materials, Volume I, published by Kluwer Academic Publishers in 1998.

ISSN 0925-0042
ISBN 978-94-007-2545-4 e-ISBN 978-94-007-2546-1
DOI 10.1007/978-94-007-2546-1
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2011944979

© Springer Science+Business Media B.V. 2012
No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Man discovered a long time ago by a succession of trials and errors the way to produce steel and to increase its hardness by quenching. The empirical recipes were more or less kept secret within craft guilds. It is not until 1772 that René-Antoine Ferchault de Réaumur[1] shed some light on the difference in carbon content of cast iron and steel. He thus opened the way, and Alexandre-Théophile Vandermonde, Gaspard Monge, Claude Louis Berthollet[2] after him, to a true scientific understanding of metallurgy, enabling to master transformations and the effects of metal treatments. Scientific researches following, among which should be emphasised the pioneering works of Henry le Chatelier, of Floris Osmond, of Georges Charpy and of Léon Guillet, allowed understanding the chemistry of alloys, the links between microstructure and mechanical behaviour and improving industrial processes. The science of metals thus founded opened the way to the wider subject of materials science. In the same way the practical problems of construction led, earlier than metallurgy as shown by the interest of Leonardo da Vinci and Galileo Galilei for the strength of materials, scientists like Robert Hooke, Joseph Louis Lagrange, Leonhard Euler, Augustin Louis Cauchy to build solid mechanics as a branch of applied mathematics. For a long time the constitutive equations needed in structural mechanics remained crude idealisations of actual behaviour. The pioneers in this field could correspond equally well with their peers about metallurgy (or alchemy) as about mechanics (or astrology). Later, scientists have become more and more specialised, and there is yet not enough overlap between materials science and solid mechanics.

As technical equipment of ever-greater sophistication has become available, the risk of catastrophes, of a scale that can affect the environment and kill many people,

[1] René-Antoine Ferchault de Réaumur (1772) L'art de convertir le fer forgé en acier et l'art d'adoucir le fer fondu; ou De faire des Ouvrages de fer fondu aussi finis que de fer forgé, Michel Brunet Paris.

[2] Vandermonde A T, Monge G and Berthollet C (1790) Avis aux ouvriers en fer sur la fabrication de l'acier. Comité de salut public.

has increased; and safety has become a major concern. As we write this foreword, hour after hour we learn that the consequences of the earthquake and tsunami in Japan keep being more and more severe. Economic considerations press for longer lifetimes and smaller safety factors; these generate strong incentives to use more realistic constitutive equations and better failure criteria in the calculations, and the computer now makes this possible. Materials design has become much more of a practical possibility, and materials can be produced with better and more reliable properties.

All this shows that establishing relations, as quantitative as possible, between the microstructure of materials and their macroscopic properties is nowadays essential. Thanks to fruitful cooperation between materials scientists and solid-mechanics specialists, recent research has led to promising achievements in this direction; but the number of training programs covering both fields, which we considered to remain low when writing the first edition of this book, tends even to decrease. It was the awareness of the need for advanced courses here that led us, some 26 years ago, to create in France what was called a Diplôme d'Études Approfondies (DEA) – Advanced Studies Diploma – with the title "Mécanique et Matériaux" – "Mechanics and Materials". The notes provided for the courses were the root of two books written in French[3] concerning mechanical properties of materials. The need was probably greater in France than in English-speaking countries, where the famous book of McClintock and Argon, Mechanical Properties of Materials, was already much in use. This, however, was published in 1966 and so did not deal with recent developments. This gave us the incentive to embark on writing these books, even though we felt that it was hard to match McClintock and Argon.[4] The D.E.A. "Mécanique et matériaux" trained some 500 students. It is estimated that about 300 of those then prepared a PhD. They have pursued careers in University and Industry, contributing to continued technical progresses. Unfortunately, for some of those highly political reasons, the D.E.A. was discontinued. As a consequence, textbooks about mechanics and materials are probably more than ever needed.

In the mean time, Kluwer asked us the permission to translate the books in English. Dr. Jack Howlett was appointed for this thankless work. It gave birth to two new books: "Mechanical behaviour of materials" in the Solid mechanics and its applications series.[5] Apparently, they met a certain audience, so that Springer, continuing the series asked us to prepare a new edition. When we agreed to proceed, we underestimated the amount of revision required. Now at last, we are glad to finish Volume I.

---

[3] François D, Pineau A and Zaoui A (1991, 1992) Comportement mécanique des matériaux. I Élasticité et plasticité. II Viscoplasticité, endommagement, mécanique de la rupture, mécanique du contact. Hermes Paris.

[4] Argon AS and Mc Clintock FA (1966) Mechanical behaviour of materials. Addison Wesley, Reading.

[5] François D, Pineau A and Zaoui A (1998) Mechanical behaviour of materials. I Elasticity and plasticity. II Viscoelasticity, Damage, Fracture Mechanics, Contact Mechanics. Klüwer Academic Pub., Dordrecht.

The organisation of this volume follows the main classes of mechanical behaviour: elasticity, elastoplasticity, elastoviscoplasticity and viscoelasticity. Throughout we attempt to describe the physical processes that are responsible for the kinds of behaviour studied, the way in which the constitutive equations can represent the behaviour and how they relate to the microstructures. Revising the book, we improved much the existing material, in particular in modifying the organisation, and we added new up to date content. Understanding the subject matter requires a good knowledge of solid mechanics and materials science; we give the main elements of these fields in a set of annexes at the end of the first volumes. We thought interesting for the readers to give as footnotes some information about the many scientists whose names are attached to theories and formulae and whose memories must be celebrated. Wikipedia proved extremely helpful in doing so.

We are now to undertake the revision of the second volume, which will be devoted to fracture mechanics and damage as well as elements of contact mechanics, friction and wear. We have realised that exercises to illustrate the various chapters, and case studies also, would occupy too much space to be included in each book and thus this will need a third volume. Now, each volume will be self-sustained.

Whilst the present book, as well as the following ones, is addressed primarily to graduate students, part of it could possibly be used in undergraduate courses; and we hope that practising engineers and scientists will find the information it conveys useful. We hope also that English-speaking readers will be interested in the aspects of French culture, and more particularly of the French school of micromechanics of materials, which our treatment will undoubtedly display.

The authors are very grateful to all their colleagues, in particular those who participated in the DEA “Mécanique et Matériaux”, for their contributions and encouragements. We wish to thank all those people who have provided photographs to illustrate the book. We also thank Professor Gladwell of the University of Waterloo, Canada, for including it in the series of which he is responsible. In the course of the original translation the frequent questions and suggestions of Dr. Jack Howlett helped to improve many paragraphs significantly. We found cooperation with him very stimulating and we thank him for his excellent work. Nathalie Jacobs, of Springer, followed our work and kept helping and encouraging us in answering our many questions. Thanks to her are extended to all the people who took great care in editing a book of the best possible quality.

We are particularly indebted to Professor Georges Cailletaud (École nationale supérieure des mines de Paris) who contributed greatly to Chap. 4 about viscoplasticity, and to Professor Jacques Verdu (Arts et Métiers ParisTech) who not only wrote the part devoted to polymers in Chap. 5 about viscoelasticity but also contributed to Annex 1. Joëlle Pineau and Odile Adam have been extremely helpful and they deserve our warmest thanks.

Acknowledgements

Illustrations in this book are for the most part originals or adapted from various sources. Many figures were provided by courtesy of authors. Let them all be thanked.

Permissions for reproduction were solicited for the reproduction of original figures and photographs. Would publishers and authors who would not have been identified signal it to Springer so acknowledgments could be given in future editions.

The authors would like to acknowledge the following publishers for their permission to use a number of figures included in the text:

- MANEY Publishing

Pope SS, Ezz DP (1984) Critical resolved shear stress for Ni3Al. Int Mater Rev 29:136–167 – (Figure 13) for Fig. 3.58

Gourgues-Lorenzon AF (2007) Application of electron backscatter diffraction to the study of phase transformations. Int Mater Rev 52:65–128 – (Figures 1 & 13) for Fig. 1.21a, b

Flower HM, Gregson PJ (1987) Solid state phase transformations in aluminium alloys containing lithium. Mater Sci Technol 3:81–90 – (Figures 3 & 8) for Figs. 1.43 and 1.45

Gregson PJ, Dinsdale K, Harris SJ, Noble B (1987) Evolution of microstructure in Al-Li-Zn-Mg-Cu alloys. Mater Sci Technol 3:7–13 – (Figure 4a) for Fig. 1.44

- SPRINGER Netherlands

Cozar R, Pineau A (1973) Morphology of γ' and γ'' precipitates and thermal stability of Inconel 718 type alloys. Met Trans 4:47–59 – (Figure 4) for Fig. 1.27a

- TAYLOR & FRANCIS

Foreman AJE, Makin MJ (1966) Dislocation movement through random arrays of obstacles. Philos Mag 14:911–924 – (Figures 1, 3, & 5) for Figs. 3.95a, b

Campbell JD, Ferguson WG (1970) The temperature and strain rate dependence of the shear strength of mild steel. Philos Mag 21:63–82 – (Figures 5 & 11a) for Fig. 4.24
- ELSEVIER

Miller MK (2001) Contributions of atom probe tomography to the understanding of nickel based superalloys. Micron 32:757–764 – (Figure 11) for Fig. 1.27b
Clavel M, Pineau A (1982) Fatigue behaviour of two nickel base alloys II. Mater Sci Eng 55:173–180 – (Figures 4 & 5a) for Figs. 1.54 and 1.55
- IOP Publishing

Verdier M, Fivel M, Groma I (1998) Mesoscopic scale simulation of dislocation dynamics in fcc metals. Model Simul Mater Sci Eng 6:755–770 – (Figure 13) for Fig. 3.63
- TRANS TECH Publications

Polak J, Man J, Obrtlik K (2005) Atomic force microscopy study of the early fatigue damage. Mater Sci Forum 482:45–50. doi: 10.4028/www.scientific.net/MSF.482.45 – (Figures 1 & 3) for Figs. 1.25a, b
- WILEY & Sons

Willems B, Nistor LC et al (2005) Strain mapping around dislocations in diamond and cBN. Phys Status Solidi A 202:2224–2228. doi: 10.1002/pssa.200561923 – (Figure 3) for Fig. 3.41
Motoyashiki Y, Bruckner-Foit A, Sugeta A (2007) Investigation of small crack behaviour under cyclic loading. Fatigue Fract Eng Mater Struct 30:556–564 – (Figure 8) for Fig. 1.17

Contents

Chapter 1
Introduction

Abstract General considerations specify the nature of materials, and their importance, in particular regarding their mechanical properties. The main classes of materials derive from the atomic bonds types, which determine their overall mechanical properties. The various mechanical behaviours and modes of failure are described. This helps in the choice of a material for a particular application. A description is given of the main instruments for the observation of microstructures at various scales. A number of typical microstructures are displayed. Mechanical testing machines are described. The way to perform tensile tests in particular, as well as other usual tests, is briefly detailed in keeping with international standards. Indications are given on the measurement of elastic properties. The various mechanical behaviours determine the main classes of constitutive equations. They must fulfil several conditions, which are detailed. The stages followed in the treatment of heterogeneous materials are explained and illustrated by examples. Indications are given about anisotropy and its consequence on the constitutive equations.

1.1 The Main Classes of Materials from the Point of View of Mechanical Properties

1.1.1 What Is a Material?

The physical chemist will study the properties of *matter*, of pure metals, of alloys, carbides, nitrides, glasses, ceramics or polymers, but nevertheless seldom those of the corresponding *materials*. These last are matter that has been worked on in some way by man in order to fabricate objects. Thus flint (matter) became a material when our early ancestors worked it into the form of tools. Thus aluminium, a precious metal a century ago, became a material when used for constructing aeroplane fuselages. Thus alumina is not the same *material* in a polishing paste, a synthetic hip joint or an oven wall.

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1_1, © Springer Science+Business Media B.V. 2012

A material represents a combination of fabrication processes, microstructures and physical properties (Table 1.1). The causal chain is that the fabrication processes – casting, sintering, rolling, extruding, forging, machining, etc. – give it a certain microstructure, which in turn determines its chemical, physical and mechanical properties. But in keeping to this triangle an essential dimension has been forgotten: the *performances* of the material, meaning its efficiency, its manufacturability, its reliability, its durability, its absence of toxicity, its potential for recycling, its cost, its appearance, etc. – all in the light of the application under consideration. What we should be considering is not a triangle but a tetrahedron, whose edges represent the mutual interactions that materials engineers and scientists must learn to master, as shown in Fig. 1.1. It is good practice always to refer to these four aspects at the beginning of a paper or a report dealing with materials**.**

1.1.2 Industrial Importance

Materials are the base on which industries, economies and civilisations are built; their names distinguish the great periods of the history of mankind – the Stone Age, the Bronze Age and the Iron Age and in the twentieth century the Silicon Age. Here are a few examples that illustrate their importance.

1. It was the development of steel for rails that, through the spread of the railway system (the "Iron Road"), largely determined the growth of the Industrial Age of the nineteenth century- in America, the conquest of the Far West. Today, the TGV, transporting 2,000 passengers at 300 km/h, surpasses even the jet airliner, with 250 passengers at 1,000 km/h, in terms of passenger-kilometres per hour. But the maximum speed of the TGV – the world record is 520 km/h – is limited by the speed of propagation of flexural waves on the power catenaries, and therefore by the tension applied to these: the elastic limit of copper, chosen for its high electrical conductivity, must not be exceeded. Could we do better? with a copper-nanotubes composite, perhaps?
2. Construction of the "Grande Arche" at La Défense in Paris was made possible by the development of fused silica concrete: the very small particle size of the SiO_2 gave a much higher density than was possible with traditional Portland cement, resulting in much better mechanical properties. This kind of concrete allowed also the foundation of the Millau viaduct piles, the highest in the world, because, in particular, it generates less heat than ordinary concrete in setting.
3. Carbon-fibre composites made possible the manufacture of structures sufficiently light and strong to form the casings of the Ariane space rockets. They are widely used now in aircrafts and helicopters.
4. The increase in the power delivered by gas turbines (jet engines) is a consequence mainly of the progress in development of nickel-based, which have made it possible to run these at higher temperatures (Fig. 1.2). In particular, the absence of grain boundaries has given a higher creep resistance. The production of single-crystal turbine blades (Fig. 1.2) has contributed largely to the development of jet engines.

Table 1.1 Elements entering the definition of a material

Fabrication processes		Casting, sintering, electrolytic deposition, powder metallurgy, new processes
		Forming by deformation: rolling, extruding, drawing, forging, stamping
		Forming by removal of material: machining, grinding; removal by torch, laser beam, electron beam, water jet, electrolysis
		Heat treatment: annealing, quenching, ageing, tempering
		Assembling: welding, adhesion, riveting, bolting, screwing, friction welding
		Surface treatment: shot-peening, grinding, surface quenching, electrolytic or vapour-deposition, spray coating
Micro-structural elements		Bond type
		Molecules type and size
		Crystal structure
		Order, short-range
		Order, long-range
		Polymers: level of cross-linking
		Polymers: level of crystallinity
		Chemical heterogeneities: GP zones, segregation at boundaries
		Spinodal decomposition
		Precipitates (coherent/incoherent) – dimensions and shape – volume fraction
		Needles, laths, packets (martensite)
		Lamellae (pearlite) – size
		Twins, grown or by deformation – type – dimension
		Grains – size and shape
		Nodules (polymers) – size
		Texture
		Inclusions – size and shape – volume fraction
		Fibres: short, long, whiskers, nanotubes
		Layers (composites)
		Aggregates: nature – form – size – volume fraction
		Porosities
		Epitaxy
Properties	*Mechanical*	Elastic moduli
		Elastic limits: (proportionality, at 0.2%, upper, lower yield strengths) . . .
		Work-hardening, flow stress . . .
		Ultimate strength
		Ductility: extension, at rupture, reduction of area
		Creep rate
		Viscosity
		Damping capacity
		Hardness
		Fatigue resistance
		Wear resistance

(continued)

Table 1.1 (continued)

	Chemical	Corrosion resistance
		Oxidation resistance
		Stability, reactivity
		Equilibrium phase diagrams
	Physical	Specific mass
		Electrical conductivity
		Thermal conductivity
		Specific heat
		Temperatures and latent heats of transformations
		Coefficient of thermal expansion
		Reflectivity
		Emissivity
		Transparency
		Birefringence
		Refractive index
		Surface energy
		Binding energy
Performances		Availability
		Reproducibility
		Efficiency
		Machinability, ease of forming
		Weldability, ease of assembly
		Reliability
		Durability
		Maintainability
		Absence of toxicity
		Capacity for recycling
		Destructibility
		Appearance
		Cost
		Familiarity, experience gained

Many processes have been developed over the past two decades, such as: Hot Isostatic Pressing (HIP), Air (or Vacuum) Plasma Spraying (A/V PS); Direct Metal Deposition (DMD) using a laser; Selective Laser Melting (SLM) and Selective Laser Sintering (SLS). SLM and SLS involve the use of powders (metal or polymers)

5. The introduction on the market of swimming suits made of a new polyurethane material increased suddenly the swimming competition records.
6. Few inventions have changed the world as profoundly, and rapidly, as the microprocessor. The first brain chip was born in 1971. Today the world chip production includes more than several tens billion microprocessors. Cars typically have at least ten microprocessors and the new models have between 50 and 100 of them. The "transistor" effect was discovered in germanium by Bardeen, Brattain and Shockley[1] in 1947. The first transistors were made from

[1] John Bardeen, Walter Houser Brattain and William Bradford Shockley obtained the Nobel price in physics in 1956 for their researches on semi-conductors and their discovery of the transistor effect.

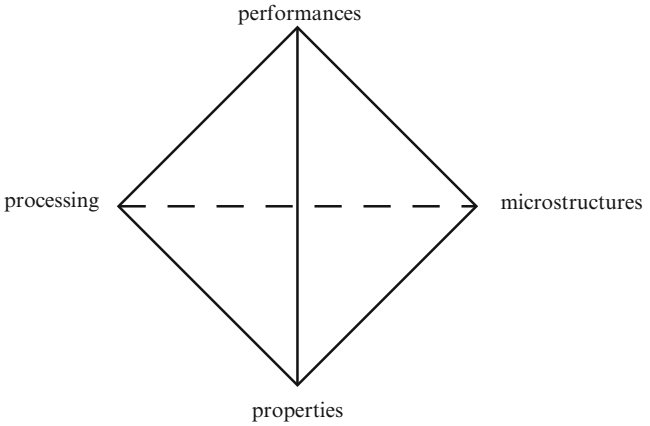


Fig. 1.1 A *material* is the outcome of a synthesis of processing, microstructure, properties and performances

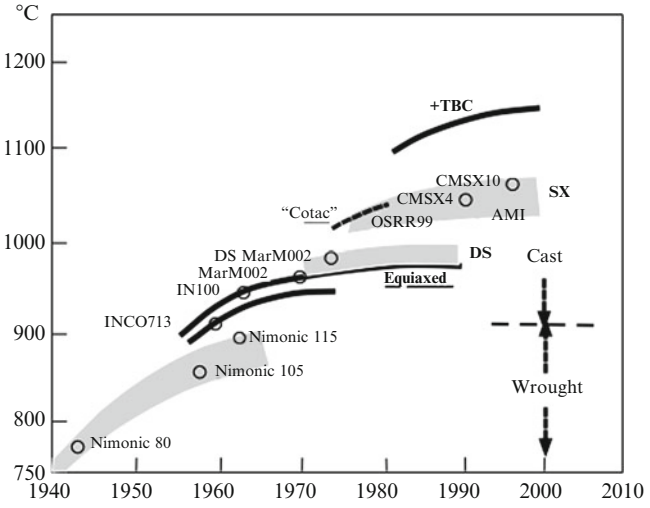


Fig. 1.2 The steady increase over time in maximum temperature in superalloys submitted to an applied stress of 150 MPa for 1,000 h. The major changes in technologies: wrought → cast (equiaxed & unidirectional solidification (DS)) → single crystal (SX) → thermal barrier coating (TBC) are indicated. The introduction of cast materials has allowed to increase the volume fraction of γ' particles. In the recent single crystalline materials the introduction of elements such as Re, Ru has led to an increase of the maximum temperature. The names of the alloys are indicated

germanium. Silicon, which has the same crystal structure (diamond cubic) as germanium and is in the same column of the periodic table (Group IV), has now replaced the first germanium-based transistors. The use of silicon in semiconductor devices demands a much greater purity than metallurgical grade. The

Table 1.2 Production of various materials (unit 1,000 t)

Material	Region	1985	2005	2008
Steel	–	–	–	–
–	European Union	–	–	198, 000
	World	710, 000	–	1, 279, 980
Aluminium	–	–	–	–
	Africa	473	1, 753	1, 715
	America	5, 944	7, 773	8, 443
	Asia	1, 096	3, 139	3, 923
	Europe	3, 695	8, 546	9, 276
	World	15, 576	23, 453	25, 654
Copper	–	–	–	–
	America	2, 289	–	5, 800
	Asia	1, 320	–	7, 800
	Europe	1, 513	–	3, 800
	World	9, 728	16, 573	18, 232
Zinc	–	–	–	–
	World	4, 826	9, 930	11, 655
Lead	–	–	–	–
	World	–	7, 005	8, 673
Plastics	World	68, 000	–	190, 000

Rare metals are not listed in this table, in spite of the high strategic importance of some of them, such as indium widely used in materials for electronic components (Behrenat et al. 2007)

main purification technique is that of "zone melting" also called "zone refining" which is used for the fabrication of wafers grown into mono-crystalline, ingots up to 300 mm in diameter using the Czochralski[2] process.

It is important to realise that choosing the most advanced materials does not always give an economic advantage. A manufacturer who produces 25,000 million aluminium cans per year is not going to consider changing his material when he can gain 25 millions of euros a year by finding a way to save one-tenth of a cent per can. This is why the traditional materials – steels, aluminium alloys, concrete – are far from obsolescence (Table 1.2 and Fig. 1.3). For one thing, the very fact of having been in use for so long has given them qualities of availability and reproducibility, and especially has established standards and construction rules; and for another, their properties and performances are being continually improved, and even if the improvements are small they can give large financial gains in view of the quantities used.

The importance of *recycling*, when the availability of many materials of high industrial usage is becoming crucial, should be stressed.

[2]Jan Czochralski (1885–1953) was a German-Polish scientist. One of his biggest achievements (in 1916) was upgrading the process for acquiring single crystals of metals.

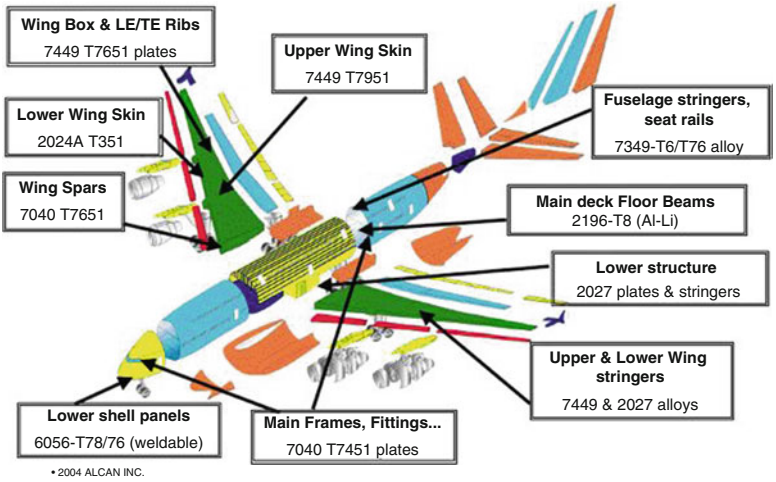


Fig. 1.3 The main aluminium alloys for Airbus A380

1.1.3 Importance of Mechanical Properties

Consideration of mechanical properties is essential for any material that is to be used in any structure – machinery, plant, vehicle, civil engineering project, prosthesis, etc. Structural materials are among those produced in the largest quantities: aluminium alloys, steels, glasses, polymers and elastomers, timbers, concretes. They are equally important to the big organisations that control their production and to the very large number of users; the level of need for a better understanding of their mechanical properties and of the relation of these to their fabrication and their microstructure is measured by the daily occurrence of accidents and catastrophes. Consider the following:

1. In an engineering workshop a steel reduction gear, about 1 m diameter, shattered during a routine over-speed test. One half went through the roof and wrecked several cars parked outside the adjacent factory; the other careered around, seriously damaging several machines but, by a miracle, did not kill anyone. Subsequent examination revealed cracks of about 10 cm diameter; this represented a problem of flakes in the fabrication of the material.
2. On December 12, 2000 a crack opened in a weld in the 3 m in diameter shaft of the Grande Dixence dam, discharging about 27,000 cubic metres of water down the mountainside. This dragged down rocks, trees and mud sweeping cottages, killing three people. The slide temporally formed a dam on the Rhone River. This was the result of a poor control of welding operation.
3. In Japan, a Boeing crashed into a mountainside after half an hour flight: control had been destroyed by the failure of an aluminium panel that closed the rear of

the passenger compartment. It was found to have had fatigue cracks that had been badly repaired: insufficient mastery of resistance to fatigue.

4. The Alexander Keelan drilling platform sank as a result of a number of errors: one of the welds in the tubular structure was found to have fatigue cracks several centimetres long. Even if this was not the main cause of the accident, it showed a lack of understanding of fatigue in such a structure.
5. On 4 November 2010, an Airbus A380 aircraft operated as Quantas flight 32 departed from Singapore for Sidney, New South Wales in Australia. A total of 469 passengers were on board. One of the four engines exploded during the climb at 7,000 ft above mean see level. A subsequent examination of the aircraft indicated that this engine had sustained an uncontained failure of the intermediate pressure (IP) turbine disc. Although the examinations of components removed from the failed engine are still ongoing, the investigators have identified the presence of fatigue cracking within a pipe that fits oil into the high pressure (HP)/IP bearing structure.
6. Alloy 600 (Ni – 16Cr – 8Fe – 0.10C) was initially selected for the fabrication of the tubes of heat exchangers in pressurised water reactors (PWR). This alloy gave rise to the formation in service of stress-corrosion cracks. This very damaging and expensive phenomenon has necessitated to replace many tubes in many exchangers all around the world and to use an alloy with higher chromium.

Such examples could be multiplied, all underlining the inadequacy of our understanding. On the other hand, there are cases where the accumulation of knowledge has made it possible, if not to avoid a catastrophe entirely, but at least to reduce the severity of its consequences.

(a) The use of ultrasonics in non-destructive testing made it possible to detect cracks of about 7 mm depth in the carbon steel under the stainless steel coating in the vessel of a nuclear reactor. The knowledge then built up by CEA, EDF and Framatome enabled them to convince the safety authorities that none of these cracks will extend under fatigue sufficiently to pierce the inner coating within 40 years. What was at stake here was half of France's electricity generation.
(b) Similarly, investigation by non-destructive testing of the welds in the Alaska oil pipeline enabled the National Bureau of Standards to demonstrate that cracks found in these were not detrimental enough to require the whole structure to be rebuilt.

These two examples, with their enormous economic consequences, show very clearly the amount of testing and calculation that must be undertaken in order to get a full understanding of fatigue failure when one is dealing with a variety of materials, all with their characteristic properties and microstructural imperfections, subjected to variable and complex cycles of mechanical and thermal stress at a wide range of temperatures.

Table 1.3 Materials classified according to bond type

1- Covalent[a]	Boron (fibres)
	Diamond carbon (abrasive powders, fibres)[b]
	Silicon, germanium, semiconductors (electronic components)
	Carbides, nitrides (grinding wheels, bearings, refractory fibres)
	Thermosetting polymers
2- Ionic compounds	*Alumina*, *silica, zirconia* (abrasives, refractories)
	Glasses (window glass, containers, fibres)
	Uranium oxide and carbide (nuclear industry)
3- Metallic	*Aluminium* (light alloys)
	Magnesium
	Zinc (building materials, castings)
	Titanium (aeronautics)
	Zirconium (chemical industry)
	Copper (brasses, bronzes)
	Gold, silver (coinage, plating)
	Tin, lead (plumbing, plating)
	Tantalum (refractories)
	Iron (steels, castings), cobalt, *nickel* (superalloys)

[a]The refractory oxides have bonds which are partially covalent
[b]Graphite carbon has covalent bonds and free electrons, which give it metallic properties

1.1.4 Bond Types for the Main Classes of Materials

The properties of a material are highly correlated with the type of bond between its atoms, for it is this that determines its energy of cohesion (the force needed to separate the atoms either by a small amount or by enough to result in rupture), the coefficient of thermal expansion, the conductivities of heat and electricity, the stability of the various phases, the reactivity, etc. There are five types of bond: *covalent, ionic, metallic, van der Waals and hydrogen*. Table 1.3 lists the main materials, grouped according to bond type, which defines also the main classes of material. These classes are characterised by properties that are familiar to the man in the street, who has no difficulty in identifying ceramic, glasses, metals, plastics, rubbers, woods, cements and concretes. Covalent materials have particularly high bond energies; they are usually refractories, insulators or semiconductors, hard and brittle. Refractory oxides have characteristics in common with covalent materials; they are light, become electrically conducting at high temperature; in glass form they are transparent. Metals are good conductors of heat and electricity, are malleable and reflect light. Polymers are poorly resistant to heat but offer great possibilities for moulding. Elastomers can be subjected to large elastic deformations. Woods are very anisotropic and are sensitive to changes in humidity. Cements and concretes resist compression well but are brittle. All the materials of this last class are of high porosity and are poor conductors of heat and electricity.

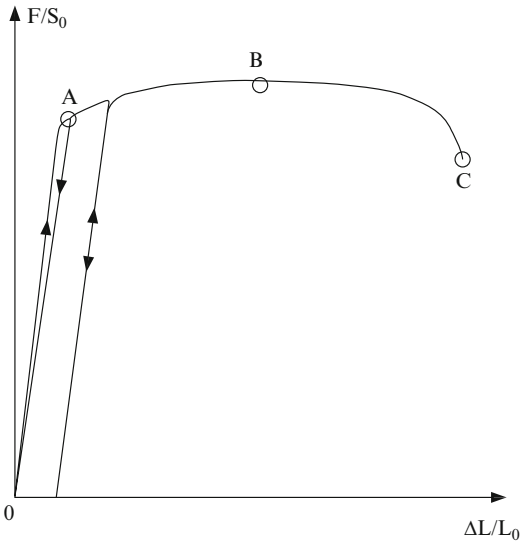


Fig. 1.4 Schematic stress–strain recording

To this list should be added *composite materials* and multi-materials, which are combinations of various materials.

The main properties of materials are given in standard manuals such as *Handbook of Chemistry and Physics*, *Metals Handbook*, etc.

1.1.5 Main Types of Mechanical Behaviour

Mechanical properties are characterised by measurements made in simple tests, easily performed in any laboratory: hardness, tension, compression, torsion, impact tests. These tests have been standardised (cf. ISO standards) and when comparing results it is important to ensure that the standard conditions have been applied. However, much more information is obtained in the course of a test than is contained in the standardised results, and it is important to preserve this, as it could prove very valuable for research. Further, the constitutive laws that express the behaviour are tensorial in form, and to determine them fully requires much more than the current standard tests.

In many cases, a tensile or a compression test shows three stages in the response of the material (Fig. 1.4). The first is the *elastic region*, OA in the figure, in which no permanent deformation remains after the load has been removed. The second, AB, is the *plastic region*, characterised by a permanent residual deformation. Pure plastic behaviour is completely independent of time, in particular of the rate at which the load is applied. The final stage, BC, ends in rupture; we shall describe this in more detail later. Before doing so we should note that in most cases pure time-independent plastic behaviour is not observed and the form of the curve can be affected by the

rate at which the load is applied: holding the load constant will result in progressive deformation by *creep*, holding the deformation constant will give a *relaxation* of the load, and cyclic loading will give phenomena of *hysteresis*. These last are a manifestation of the viscosity of the material, the corresponding phenomena being either *visco-elastic* or *visco-elasto-plastic*. Not all materials show both of the two first stages; the point C of final rupture can occur early or late, in particular it can precede point B.

The response to compression, and in general to any other type of applied stress, can be very different from that to tension.

1.1.6 Modes of Failure

The failure modes of a structure depend on applied loadings – mechanical, thermal, chemical –, on the internal defects it may contain and on the properties of the material(s) of which it is built.

They include the phenomena of *buckling*, instabilities resulting from a reduction in energy under small perturbation of the boundary conditions.

Chemical attack manifests itself in *oxidation* and *corrosion*, localised to a greater or lesser extent. These are very important effects, but we shall not deal with them except when they are accompanied by mechanical stresses.

The phenomena of fracture derive from *damage* of the material, that is, the development of new surfaces. On the atomic scale there are three basic types of damage (Fig. 1.5): cleavage, slip with formation of surface steps, and appearance of cavities resulting from diffusion of vacancies. The last can occur only when the temperature is high enough; it is the dominant effect in *creep* at high temperatures, and occurs also when *neutron or other particles irradiation* creates a large excess of vacancies over the equilibrium concentration.

Slip is responsible for plastic deformation. This can lead to the structural instability of *necking* or *plastic collapse* when the reduction in cross-section that plastic deformation produces is no longer compensated by the strain-hardening of the material. The deformation is then localised and premature for low strain-hardening material. Plastic deformation can also lead to the formation of small internal *cavities* in the material, which can grow, coalesce and finally cause fracture. Under cyclic loading slip is not perfectly reversible: after a time it can cause the surface to deteriorate, ending with *fatigue* failure.

The term *cleavage* is used only in connection with crystalline materials, but an analogous mechanism, the breaking of bonds normal to the plane of crack (called Mode I in fracture mechanics), is responsible for the *crazing* of polymers, and for the fracture of concrete and glass. This is also the case for *intergranular fractures*. These various fracture phenomena interact with chemical mechanisms, such as migration of gases into the cavities, blocking of slips by solute atoms or by precipitates, changes in boundary or surface energies due to migration of impurities, chemical reactions on the new surfaces created. The various types of damage are summarised in Table 1.4.

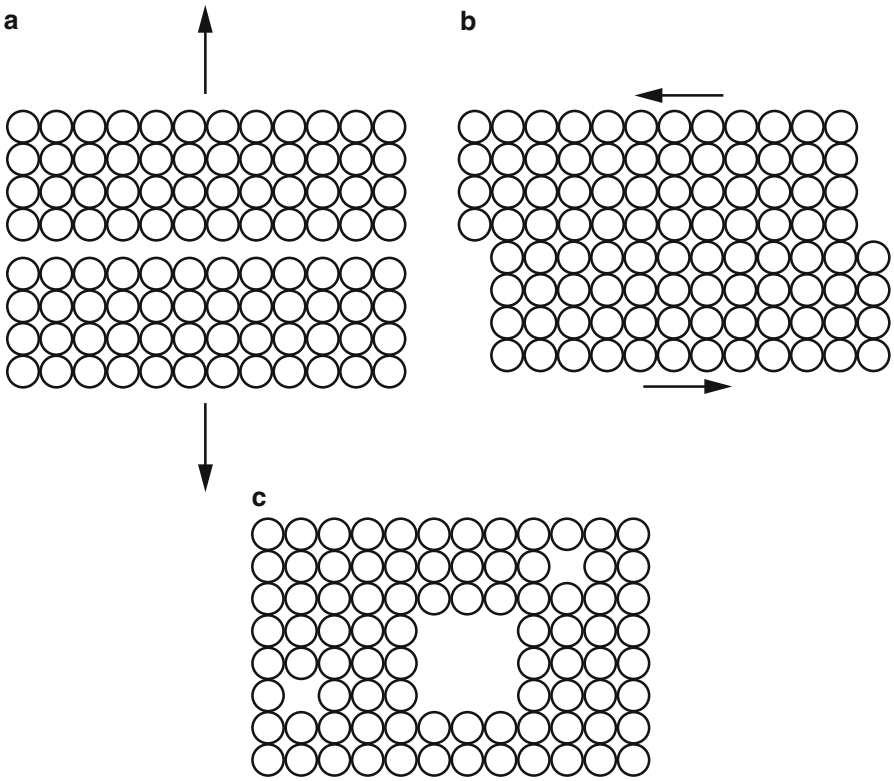


Fig. 1.5 Atomic scale damage: (**a**) cleavage; (**b**) slip; (**c**) appearance of cavities

Table 1.4 Various modes of damage

Fracture type	Volume damage		Mixed damage	Surface damage
Sudden fracture	Development of cavities (Trans or inter-granular)	Cleavage (Trans or inter-granular)	–	Liquid metal embrittlement (Hg, Ga . . .)
Delayed Fracture	–	Impurities embrittlement	←−−−−→	Oxidation Corrosion
		Hydrogen embrittlement		Stress-corrosion
	–		–	Corrosion-fatigue
	High temperature creep		Creep-fatigue	Fatigue
	Irradiation embrittlement		–	–

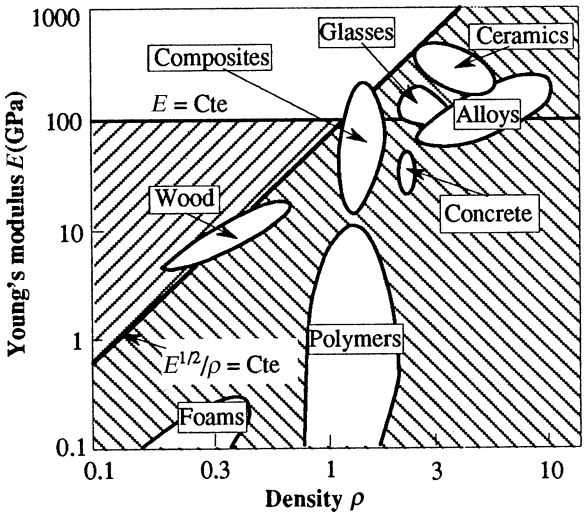


Fig. 1.6 Maximising resistance to buckling (for a column)

1.1.7 Mechanical Properties as Criteria for Choice of Materials. From Properties to Performances

A good choice of material for a particular application constitutes an illustration of how a sensible combination of materials properties is able to meet performances requirements. Because they have the same or similar bond types and structures, the materials in any one class have similar physical and mechanical properties. We can regard each class as occupying some region in the space of general properties. Particular projections of this space can be used to guide the choice of material(s) for a particular application. This is illustrated by Figs. 1.6, 1.7, 1.8, 1.9, which we took from (Ashby 1989, 1999).

The parameters to be considered depend on the application. Obviously, when rigidity is the most important design criterion, the elastic moduli need to be taken into account; and the endurance limit when it is fatigue resistance. Now, there are also constraint limitations such as weight or cost. The best choice of a material must then combine the properties in an appropriate way.

Two simple examples illustrate the procedure. We want a slender bar of length l to be as strong and as light as possible. It is loaded either in tension or in bending. The cross section S_0 can be chosen at will. In tension the stress F/S_0 must be less than R_m the tensile strength of the material while the mass M is $S_0 l\rho$, ρ being the specific mass. The way to maximise the ratio F/M is found by eliminating S_0: the ratio $(1/l)(R_m/\rho)$ must be as large as possible. Thus the material index to be maximised in this particular application is R_m/ρ. In bending, a different index is to be considered as $F < (bh^2/6\,l)R_m$, where b is the width and h the height of the beam; eliminating S_0 for a beam of square cross section shows that the index to be maximised this time

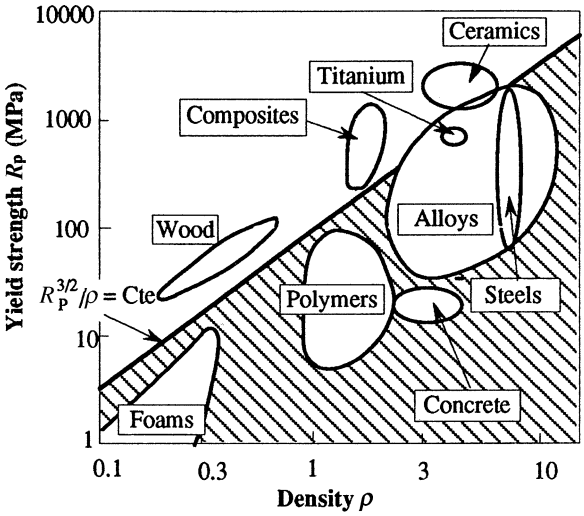


Fig. 1.7 Maximising specific resistance to torsion

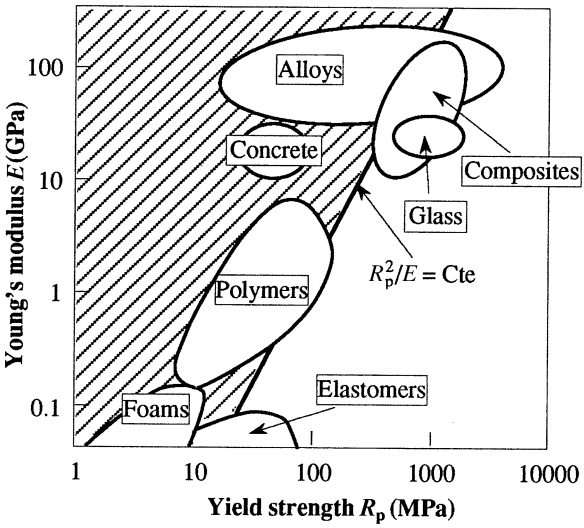


Fig. 1.8 Maximising stored elastic energy (in a spring)

is $R_m{}^{2/3}/\rho$. If it is the height which can be chosen at will, then the index is $R_m{}^{1/2}/\rho$, and it is R_m/ρ when it is the width.

Table 1.5 gives the criteria for optimising various structures with respect to weight, and Figs. 1.6, 1.7, 1.8, and 1.9 are projections of the space of properties relevant to the specific applications given in the captions. These diagrams are logarithmic, so that material indexes of given value lie on straight lines of various slopes. It is then easy to visualise the materials, which yield the maximum performance index for a particular application.

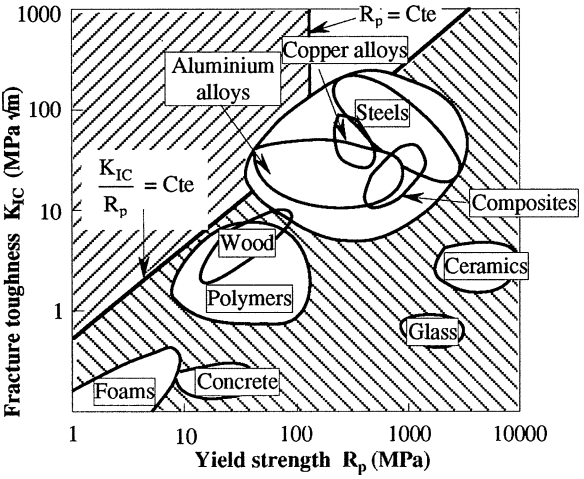


Fig. 1.9 Maximising fracture toughness (pressure vessel)

Table 1.5 Performance indexes for the choice of materials to maximise specific stiffness and specific strength

Structural element	Constraint	Stiffness	Strength
Bar in tension	Free cross section	E/ρ	R_m/ρ
Beam in bending	Fixed length	$E^{1/2}/\rho$	$R_m{}^{2/3}/\rho$
	Free cross section		
	Fixed length	$E^{1/3}/\rho$	$R_m{}^{1/2}/\rho$
	Free height		
	Fixed length	E/ρ	R_m/ρ
	Free width		
Shaft in torsion	Free cross section	$\mu^{1/2}/\rho$	$R_m{}^{2/3}/\rho$
Column in buckling	Free cross section	$E^{1/2}/\rho$	R_m/ρ
Cylindrical pressure vessel	Free wall thickness	E/ρ	R_m/ρ
Spherical pressure vessel	Free wall thickness	$E/(1\text{-}\nu)\rho$	R_m/ρ

1.1.8 As a Conclusion: The Structure of the Books

What we wish to achieve is to provide students, professors, engineers with up to date tools to have a good command of mechanical properties of materials, so as to be able to design and maintain industrial components and structures. Furthermore, we would hope to give basis for sound research programs to be established in this field. This ambition was inspired by the book of Argon and Mc Clintock (1966), which gave emphasis on the importance to treat together mechanical properties over different scales, including microstructures descriptions.

The logic we will follow is to adhere to the various classes of mechanical properties. Thus, we will treat in succession elastic, elasto-plastic, elasto-visco-plastic and visco-elastic behaviours. This will be followed, in the second book, by chapters devoted to different modes of failure of materials, according to their preceding description (Sect. 1.1.6). Lastly, we will address contact mechanics associated with friction and wear. In all cases, microstructural aspects will be linked with macroscopic approaches.

As the level at which the various treatments will be addressed is that of graduate studies, it was felt that more elementary summary of materials science and solid mechanics would be useful and they are included as annexes. They have the other advantage to put within easy reach fundamental formulae.

A third book will include exercises related to the various chapters of the first two books for which solutions will be outlined.

Emphasising once more the importance of microscopic observations and data acquisition of mechanical properties, we will now give some indications about instruments and test machines useful for such purposes. They will be accompanied by descriptions of typical microstructures and by the way mechanical tests, in particular the tensile test, ought to be performed. This will be followed by general considerations about constitutive equations, essential for the computation of deformations of structures and of the stresses they must bear to be possible.

1.2 Microstructures of Materials

1.2.1 The Importance of Microstructural Observations

The mechanical behaviour of a material and the damage it can sustain are very closely dependent on its microstructure, that is, not simply its overall chemical composition but also the arrangement of the various phases that it contains. This is so not only for artificial composites (for example, materials into which reinforcing elements or matrices were incorporated at the time of manufacture) but also to materials that have been in use for a long time and which are mostly natural composites. Understanding the in-service behaviour of materials increasingly requires the investigation of the microstructure of the materials. Modern techniques of observation, of local chemical analysis and of image analysis have contributed to the development of a “new” mechanics of solids, sometimes referred to as meso-mechanics to show how it fits between continuum mechanics on the one hand, in which matter is regarded as a “black box”, and micro-mechanics (or even nano-mechanics) on the other, in which the accent is on such basic phenomena as the displacement of a dislocation line or the interaction between a point defect and a dislocation.

Table 1.6 Characteristics of four observation techniques

Technique	Macroscope	Optical microscope	Scanning Electron Microscope (SEM)	Transmission Electron Microscope (TEM)
Resolution	0.1 mm	1 μm	5 nm	0.2–2 nm
Magnification	×20	×1,000	×50,000	×500,000
Wave length	0.5 μm	0.5 μm	0.01 nm	0.0001 nm
Mode of observation	Reflection – transmission	Reflection – transmission	Reflection	Transmission
Size of objects	A few cm^3	A few cm^3	A few cm^3	<1 μm^3
Thickness of objects	–	–	–	0.1–1 μm

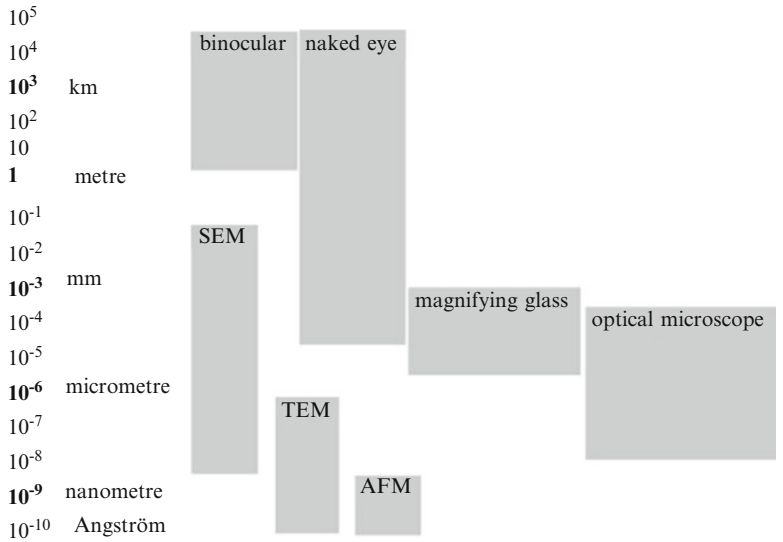


Fig. 1.10 Observation scales and main instruments used

1.2.2 Scales of Observation. Available Means of Observation

Many types of instruments are now available for studying microstructures; each has its own field of application, according to the magnification required, the method of observation (reflection or transmission) and the size of the object to be studied.

Table 1.6 summarises the characteristics of four methods currently in use. Figure 1.10 gives the ranges of size at which observations can be made with the various instruments. Figs. 1.11 and 1.12 relate to microstructures of materials, giving respectively grain sizes and reinforcements.

We now describe commonly used instruments: the optical microscope and two types of electron microscope and more newly developed techniques; and follow this

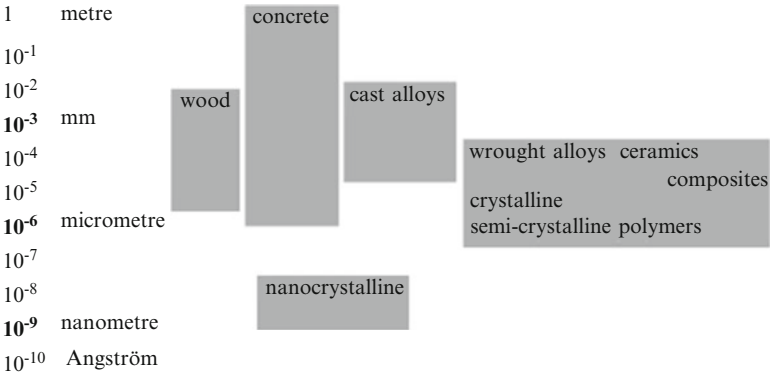


Fig. 1.11 Scales of granular microstructures for various materials

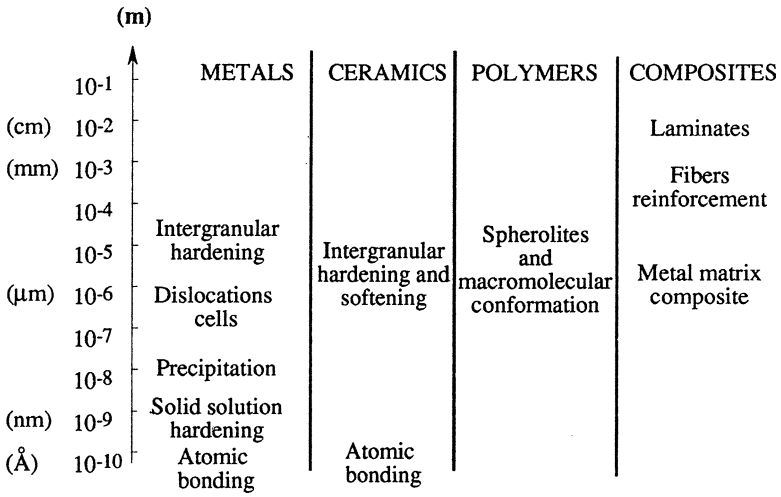


Fig. 1.12 Scales of reinforcements for various materials

with examples that illustrate the main fields covered in the book. We are limiting ourselves to instruments that enable us to "see" a microstructure, so that we shall not consider, for example, X-ray crystallography.

1.2.2.1 Optical Microscope

Metallic and ceramic materials, being opaque, are viewed by reflection; certain glasses and polymers can be viewed by transmission. Careful preparation of the surface is necessary in most cases. Figure 1.13 shows, schematically, the structure and principle of the optical microscope, set up for viewing by reflection.

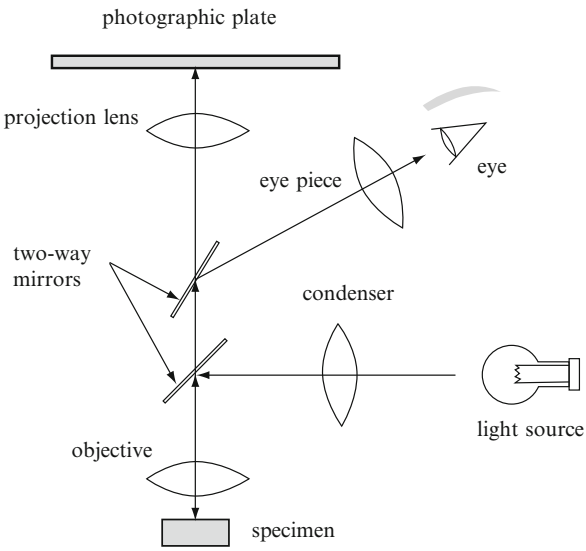


Fig. 1.13 Sketch showing the principle of the optical microscope

The resolution of the instrument is limited by the wavelength of the light used to illuminate the specimen, and the magnification possible does not exceed about 1,500 times. The resolution limit is of the order of 1 μm. Various refinements such as differential contrast and illumination by polarised light are often used.

The specimen must be specially prepared: mechanical polishing, first with abrasive papers of increasing fineness, from 100 down to 8 μm, followed by finer abrasive powders (alumina or diamond, 15 μm down to 0.05 μm) as suspensions in water or oil. The microstructure of the material is made evident by etching the surface, usually chemically or electrolytically in the case of metals and by thermal etching in the case of ceramics. The chemical etchants to be used here depend on the alloys to be studied; they are listed in metallography treatises, and the names of some recall the great figures who pioneered metallography in the twentieth century (Jacquet, Bechet-Beaujard, Murakami, etc.).

1.2.2.2 Transmission Electron Microscope (TEM)

Transmission electron microscopy (TEM) is a technique whereby a beam of electrons is transmitted through an ultra thin specimen (0.1–1 μm). This beam interacts with the specimen as it passes through. An image is formed from the interaction of the electrons through the specimen; the image is magnified and focussed onto an imaging device, such as a fluorescent screen, or is detected by a sensor such as a CCD camera. The first TEMs were built in the 1930s but the commercial "friendly" microscopes arrived in the materials science laboratories only at the end of the 1960s (Egerton 1996, 2005, Reimer and Kohl 2008; Haque and Saif 2001).

TEMs are capable of imaging at a significantly higher resolution than light microscopes, owing to the small de Broglie[3] wavelength of electrons (Table 1.6). TEM image contrast is due to the absorption of electrons in the material. At high magnification complex wave interactions modulate the intensity of the image, requiring expert analysis of observed images. Alternate modes of use allow for the TEM to evidence modulations in local chemical composition, crystal orientation and electronic structure.

The principle of a modern TEM is schematically shown in Fig. 1.14. From the top down, the TEM apparatus consists of an emission source (gun), which may be a tungsten filament or a LaB_6 source. By connecting this gun to a high voltage source (typically ∼100–300 kV) the source will begin to emit electrons either by thermionic or field electron emission. The upper lenses of the TEM allow for the formation of the electron beam to the desired size and location for later interaction with the specimen.

The specimen holders are adapted to hold a standard size of grid (3 mm diameter) upon which the sample is placed. The specimen holder device includes mechanisms for the translation of the sample in the XY plane, for Z height adjustment and usually for at least one rotation degree of freedom for the sample. Most modern TEMs provide the ability for two orthogonal rotations.

Contrast formation in the TEM depends greatly on the mode of operation. The most common mode of operation is the bright field (BF) imaging mode. Samples can exhibit diffraction contrast, whereby the electron beam undergoes Bragg[4] scattering. The desired Bragg reflections can be selected to obtain dark-field (DF) images. An example of an image showing dislocations is given in Fig. 1.15.

Many TEM are equipped with the advanced technique of electron energy loss spectroscopy (EELS). This technique allows for the selection of particular energy values and can be used to generate an image, which provides information on elemental composition. EELS spectrometers can often be operated in both spectroscopic and imaging modes.

The diffraction pattern in a single crystal is dependent upon the orientation of the specimen. An example of crystalline diffraction pattern from a twinned grain of FCC austenitic steel is shown in Fig. 1.16.

Sample preparation for TEM observation can be a complex procedure. This preparation is specific to the material under analysis and the desired information to obtain from the specimen. Materials that have small enough dimensions to be electron transparent, such as powders or nanotubes, can be quickly prepared for direct observation. In material science and metallurgy the specimens must be prepared as a thin foil by mechanical polishing, electrolytic thinning and ion etching.

[3]Louis de Broglie (1892–1987) was a French mathematician and physicist who won the Nobel price in physics in 1929 for his discovery of the wave nature of electrons.

[4]William Lawrence Bragg (1890–1971) was an Australian physicist who won the Nobel price in physics in 1915 with his father for their service in the analysis of crystal structure by means of X-rays.

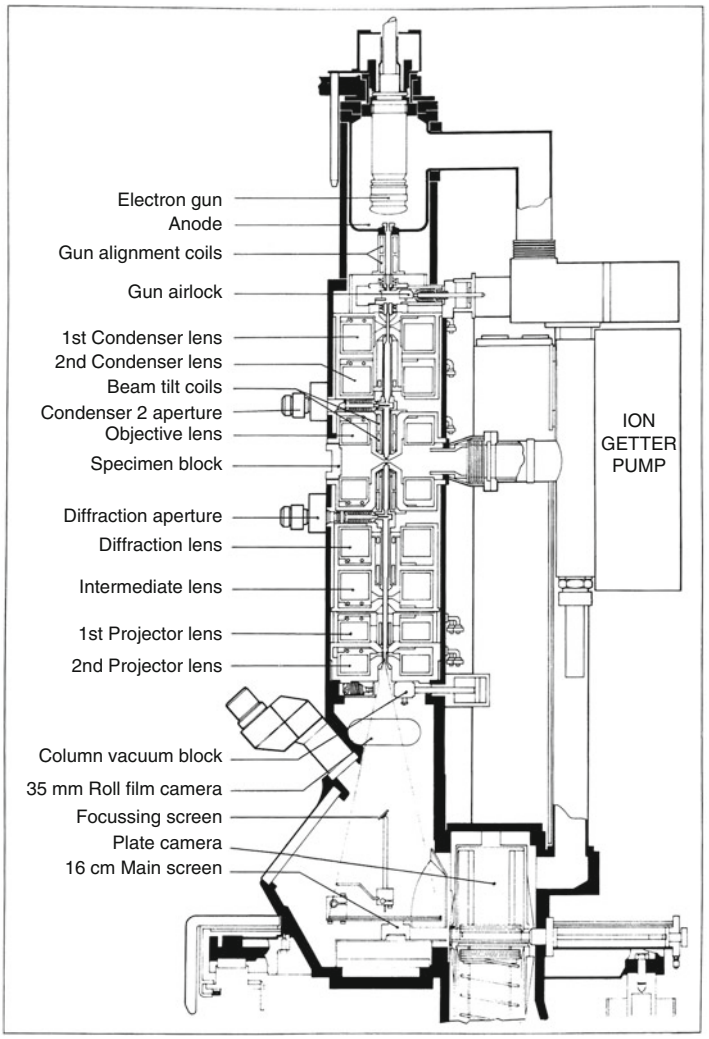


Fig. 1.14 Principle of the Transmission Electron Microscope

More recently focussed ion beam (FIB) methods have been used to prepare samples. Because FIB can be used to micro-machine samples very precisely, it is possible to mill very thin membranes from a specific area of interest in a sample, such as a fracture surface (Fig. 1.17). Unlike inert gas ion sputtering, FIB uses significantly more energetic gallium ions and thus may alter the composition or structure of the material through gallium implantation. The FIB tomography technique with the help of EBSD measurement allows for three-dimensional investigation of small objects (second-phase particles, cavities, microcracks, etc.).

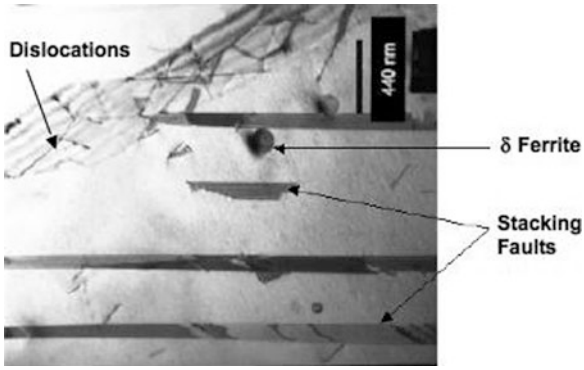


Fig. 1.15 Austenitic stainless steel. TEM observations showing dislocations, stacking faults and a particle of δ (BCC) ferrite (by courtesy of Nanga 2008)

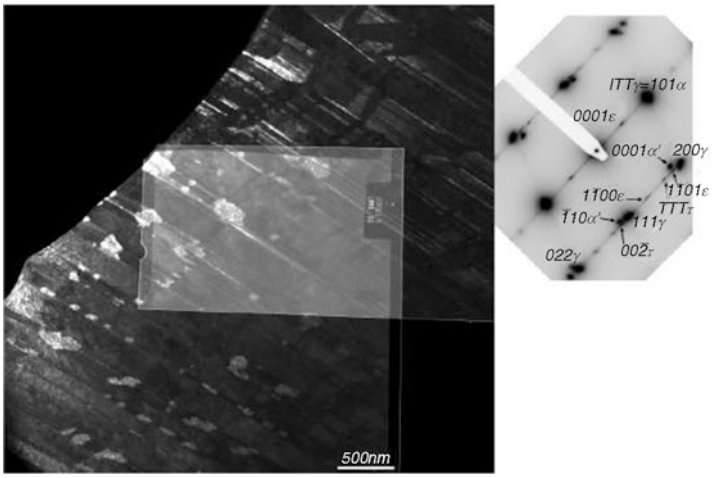


Fig. 1.16 Austenitic stainless steel, prestrained by 20% at 23°C. Dark field TEM observation showing mechanical twins, ε (HCP) martensite and α′ (BCC) revealed by the diffraction pattern (by courtesy of Nanga 2008)

1.2.2.3 Scanning Electron Microscope (SEM)

The principle of the scanning electron microscope (SEM) is very different from that of the TEM. The scanning microscope makes direct use of the various effects produced when an object is bombarded with a beam of electrons.

As Fig. 1.18a–c show, a number of signals are emitted, each originating at a characteristic depth in the material, the main ones being the following.

Back-scattered electrons, resulting from elastic scattering of incident electrons by sub-surface atoms; their energy is close to that of the incident beam, E_s, which is in the range 5–50 keV. The energy shift depends on the atomic number of the elements in the irradiated zone and, in some cases, on the orientation of the

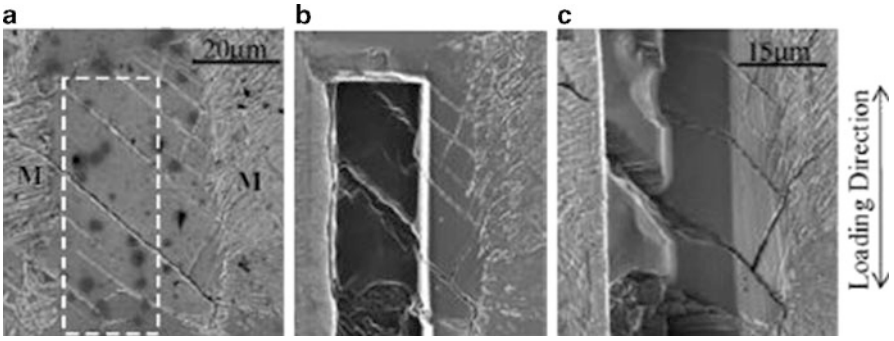


Fig. 1.17 SEM observation of fatigue microcrack in a dual phase steel with an FIB tomographic technique; (**a**) *Top view* before FIB processing; (**b**) *Top view* after FIP processing; (**c**) *Side view* with a tilt angle of 45°. This technique allows for the three-dimensional orientation of fatigue microcracks (Motoyashiki et al. 2007)

surface; it is not less than 50 eV. These back-scattered electrons can be used to give contrast effects that show local variations in the composition of the material (Fig. 1.19).

Secondary electrons, resulting from inelastic scattering in a layer closer to the sample surface than the preceding (Fig. 1.18b); the energy of these is below 50 eV and the intensity of the signal depends mainly on the local orientation of the surface to the detector used, giving a relief contrast.

Auger electrons, resulting from interaction between the incident electrons and the electrons in the inner shells of the atoms in a layer in the immediate vicinity (within about 1 nm) of the surface of the sample. The spectrum depends on the elements and is used for surface chemical analysis – Auger spectroscopy, which in recent instruments can give a lateral resolution of less than 1 μm. Auger electrons are not used in most scanning microscopes.

X-rays: emission of these depends on the atoms within a depth of about 1 μm below the surface of the sample, the beam voltage and the density of the target (Fig. 1.18c). They are used for local chemical analysis (Castaing[5] microprobe).

Figure 1.20 shows the principle of the SEM. The incident beam, of diameter less than 100 nm, is focussed by electromagnetic lenses on the surface to be examined. The emitted (back-scattered or secondary) electrons are captured by one or more detectors and the resulting signal is used to modulate the beam of a cathode ray tube (CRT). Scanning coils enable the sample surface to be traversed, line-by-line or point-by-point, and the CRT beam follows synchronously. Thus to each position of the microscope beam on the sample surface there corresponds a spot on the CRT screen, making it possible to reconstruct the surface being scanned.

[5]Raimond Castaing (1921–1998) was a French physicist.

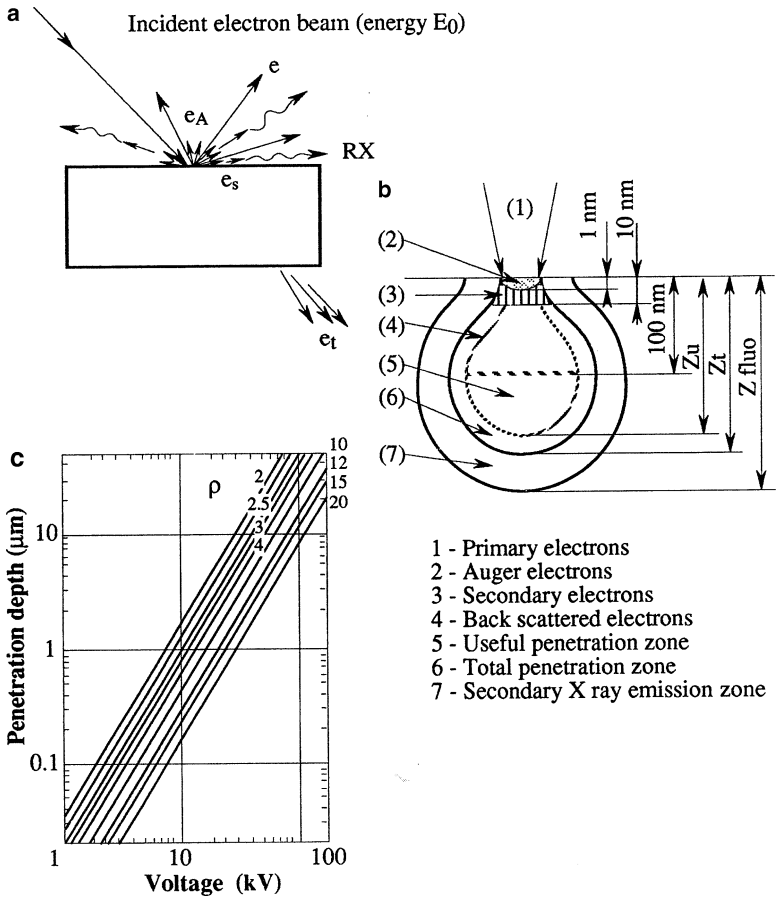


Fig. 1.18 Interaction between an electron beam and the sample. (**a**) Back-scattered electrons e; secondary electrons e_S; Auger electrons e_A; X-rays (**b**) depths of origin of the irradiation products (**c**) penetration depth as a function of beam energy and density ρ of the material

The resolution depends on the dimensions of the zone of origin of the signal being used: about 1 μm for back-scattered electrons and X-rays, about 5 μm for secondary electrons. The magnification can be up to 50,000. Equally important is the depth of field, about 100 times greater than that of the optical microscope: 500 μm at 100× and 30 μm at 2,000×.

No special preparation is needed for the samples, provided they are electrically conducting; thus metallographic sections and fracture surfaces can be examined directly. Electrical insulators are given a thin (about 10 nm) conducting coating (C, Ag, Au, Pt), by vacuum deposition from vapour or by cathodic sputtering.

The sample-holder compartment is much larger than in the case of the transmission instrument, making it possible, with certain stages, to handle samples of

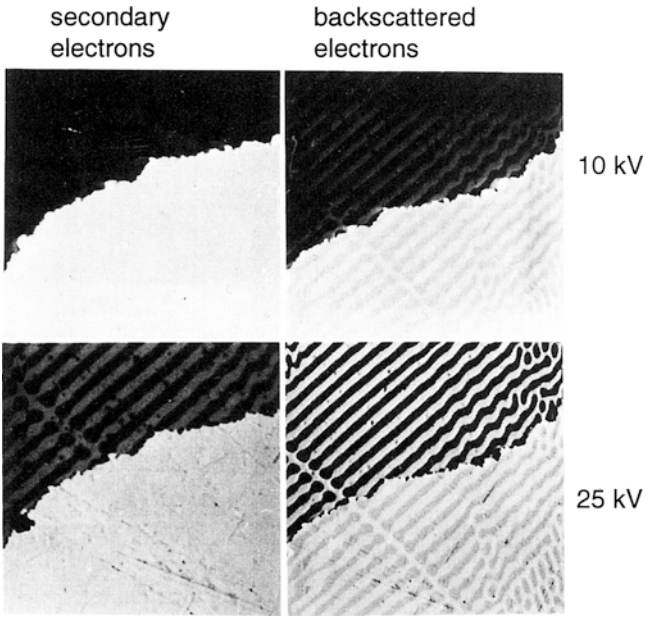


Fig. 1.19 Scanning Electron Microscope pictures of an eutectic Al-Cu alloy at beam voltages of 10 and 25 kV. Note the better contrast given by back-scattered electrons

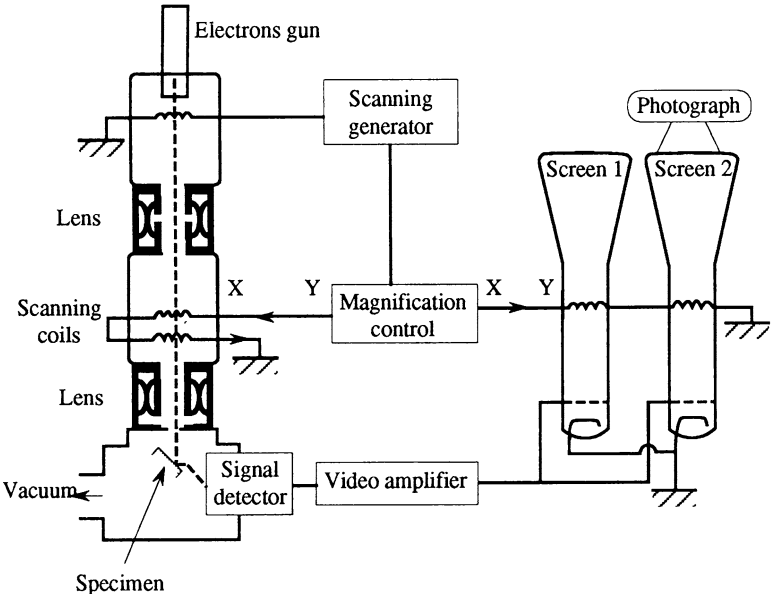


Fig. 1.20 Principle of the Scanning Electron Microscope

10–100 cm^3. Some stages have means for applying mechanical loads (tension, compression, torsion) to the sample, so that the mechanical properties can be observed *in situ*. Finally, when coupled to an X-ray spectrometer the SEM enables chemical analyses to be performed on samples of the order of 1 μm^3 and X-ray images to be produced which show the spatial distribution of individual atoms.

1.2.2.4 Electron Backscatter Diffraction (EBSD)

Electron backscatter diffraction (EBSD), also known as *backscatter Kikuchi diffraction* (BKD) is a technique used to examine the crystallographic orientation of many materials. The knowledge of these orientations is important in the study of phase transformation and plasticity and fracture in polycrystalline materials. EBSD can be used to index and identify crystal systems, and as such it is applied to crystal orientation mapping, defect studies and regional heterogeneity investigations. Traditionally these types of studies have been made using X-ray diffraction (XRD), neutron diffraction and/or electron diffraction in a TEM.

Experimentally EBSD is carried out using a SEM equipped with a phosphorus screen, and low light CCD camera. A flat/polished crystalline specimen is placed into the normal position in the SEM chamber, but is highly tilted ($\sim$70° from horizontal) towards the diffraction camera in order to increase the contrast in the resultant EBSD pattern.

An EBSD pattern is obtained when many different crystallographic planes diffract to form Kikuchi bands. It is possible to relate these bands to the underlying crystal phase and orientation of the material within the electron interaction volume. Each band can be indexed individually by the Miller[6] indices of the diffracting plane. Most commercial systems use tables with crystal data bases to perform indexing. The principle of EBSD analysis is shown in Fig. 1.21.

Scanning the electron beam results in many maps, which give local orientation of each grain in a polycrystal (orientation, image quality). An application of this technique to cast duplex (austenite + ferrite) stainless steels is illustrated in Fig. 1.22 where a detailed analysis shows that austenite colonies across parent ferrite grain boundaries keep near Kurdjumov-Sachs (KS) orientation relationship (RO) with both ferrite grains. These observations and many others on variant selection at parent grain boundaries have largely contributed to a better understanding of intergranular nucleation of a new phase during a phase transformation.

Another application of EBSD mapping technique is illustrated in Fig. 1.23 where the orientations of each individual grain in a polycrystal are shown. These observations combined with other information such as measurements of local strains using either microgrid technique or image correlation are now largely used for the detailed study of plasticity in polycrystalline materials.

[6] William Hallowes Miller (1801–1880) was a British mineralogist.

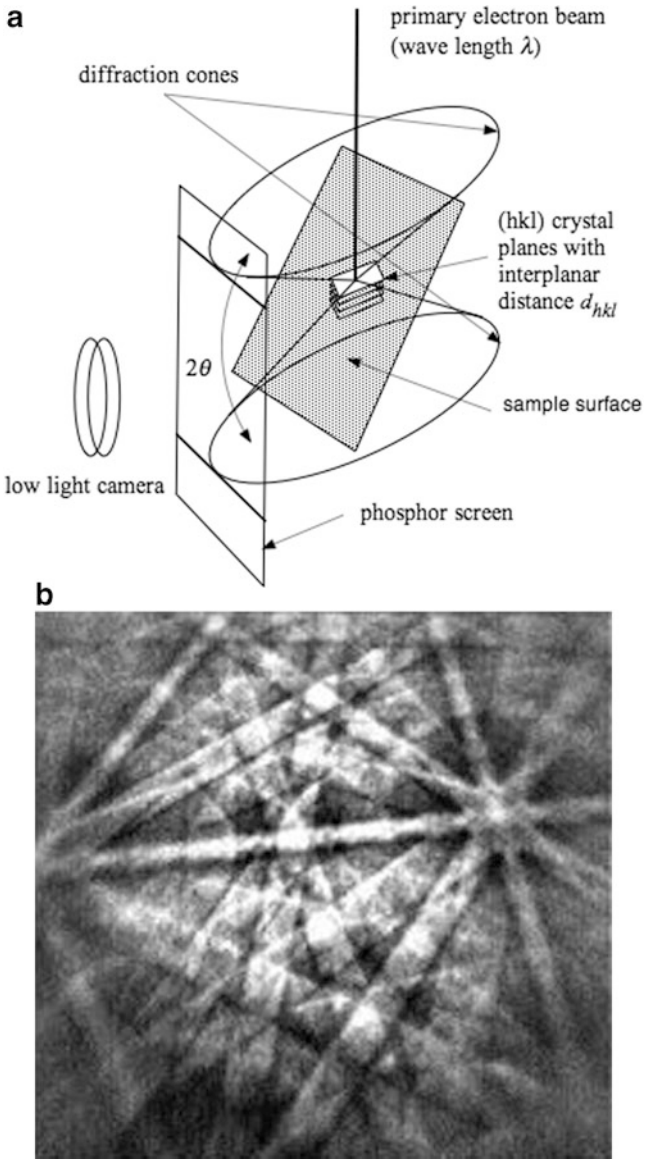


Fig. 1.21 (**a**) Electron diffraction according to Bragg's law ($2d_{hkl}\sin\theta = n\lambda$); (**b**) EBSD pattern from a bainitic steel with a BCC structure (Gourgues-Lorenzon 2007)

1.2.2.5 Scanning Tunnelling and Atomic Force Microscopy

Atomic force microscopy (AFM) or scanning force microscopy (SFM) is a very high-resolution type of scanning probe microscopy, with resolution of the order

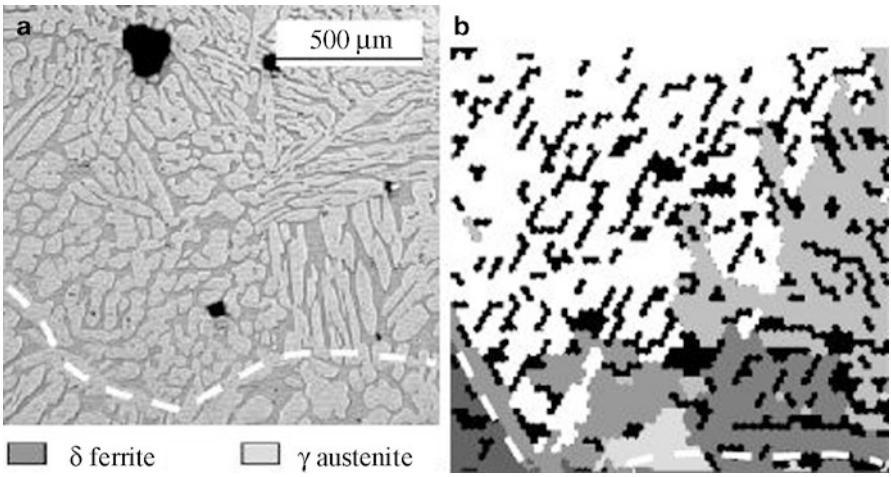


Fig. 1.22 Duplex (austenite + ferrite) stainless steel observed by SEM and EBSD technique (Gourgues-Lorenzon 2007)

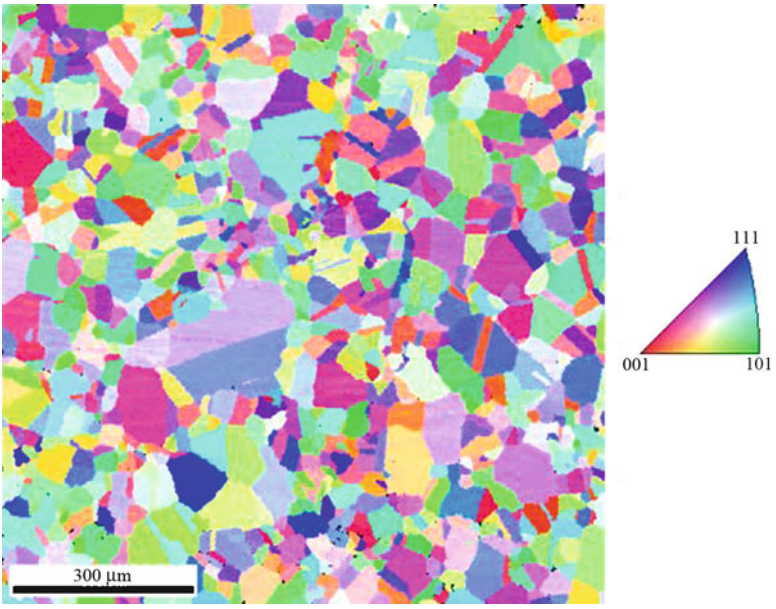


Fig. 1.23 EBSD map of 304 (18Cr-10Ni) stainless steel. The colours indicate the orientation of the grains (private communication by courtesy of Le Millier)

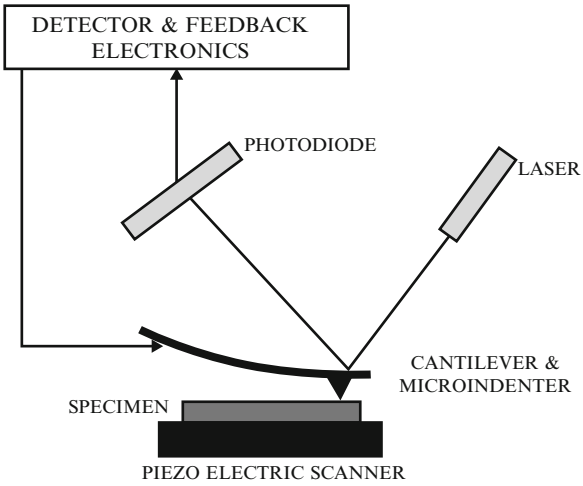


Fig. 1.24 Block diagram showing the principle of atomic force microscopy (AFM)

or fraction of a nanometre. The precursor to the AFM, the *scanning tunnelling microscope* (STM), was developed by Gerd Binning and Heinrich Rohrer in the early 1980s at IBM Research-Zurich, a development that earned them the Nobel Prize in Physics in 1986. The first commercially available AFM was introduced in 1989. The AFM is one of the foremost tools for imaging and measuring at the nanoscale.

A block diagram of an atomic force microscope is shown in Fig. 1.24.

The AFM consists of a cantilever with a sharp tip (probe) at its end that is used to scan the specimen surface. The cantilever is typically silicon or silicon nitride with a tip radius of curvature of the order of nanometres. When the tip is brought close to a sample surface, forces between the tip and the sample generate an elastic deflection of the cantilever. Depending on the situation, these forces include mechanical contact, van der Waals forces, capillary forces, chemical bonding, electrostatic and magnetic forces. The deflection can be measured using a laser spot reflected from the top surface of the cantilever into an array of photodiodes (Fig. 1.24). Other methods can also be used such as optical interferometry, capacitive sensing or piezo-resistivity elements.

In most cases a feedback mechanism is used to adjust the tip-to-sample distance in order to maintain a constant force between the tip and the specimen, which is scanned. The resulting map represents the topography of the specimen. An example of AFM images of fatigue persistent slip bands (PSB) observed in an austenitic stainless steel is shown in Fig. 1.25. The height of these PSBs, which represent the early stage of fatigue damage, can be measured as a function of the number of applied fatigue cycles. AFM is thus an excellent technique to investigate the early stages of deformation and damage.

It is also extremely useful for nano-indentation measurements (see Sect. 1.3.5.3).

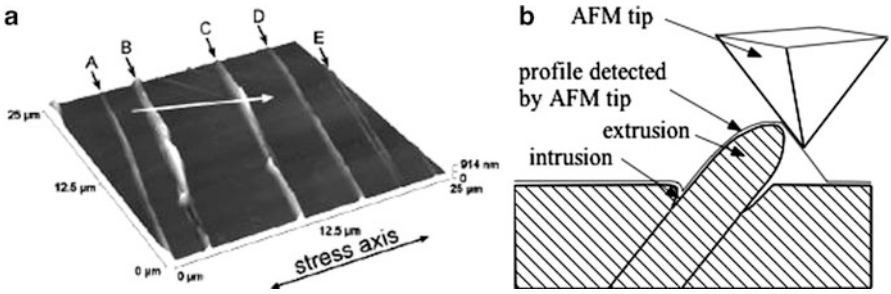


Fig. 1.25 (**a**) AFM image of the free surface of 316L austenitic stainless steel fatigued at $\varepsilon = 2 \times 10^{-3}$ for 20,000 cycles; (**b**) True profile of a persistent slip band and profiles detected by AFM tip Polak et al. 2005)

1.2.2.6 Electron Microprobe (EMP), Secondary Ion Mass Spectroscopy (SIMS) and Rutherford Backscattering Spectrometry (RBS)

Electron microprobe (EMP) was invented by R. Castaing as a non-destructive method for chemical analysis of surfaces of solid samples. A focussed electron beam is applied to excite characteristic X-rays (see Fig. 1.18) revealing the identity and concentrations of chemical elements present in the sample. The resolution is of the order of one micron and all elements heavier than boron can be detected to a level as low as ∼10–100 ppm.

The basic components of an electron microprobe are: electron gun, two or more sets of lenses, sample holder and an X-ray analyser, which can be either an *energy-dispersive spectrometer* (EDS) or a *wavelength-dispersive spectrometer* (WDS). The wavelength recorded in WDS or photon energy recorded in EDS are characteristic of each chemical element, while the intensities in the X-ray spectrum are used for quantitative analysis by reference to standards. Often, electron microprobe spectrometers (mainly EDS) are coupled with a high quality imaging microscope such as TEM or SEM to allow for a visually correlated analysis of a specimen. Such a powerful combination of instruments has applications in a wide range of areas due to its non-destructive nature.

Secondary ion mass spectrometer (SIMS) is also an instrument for local chemical analysis. The SIMS microprobe uses an internally generated beam of either positive (*e.g.*, Cs) or negative (*e.g.*, O) ions which constitutes the primary beam focussed on a sample surface to generate ions that are then transferred and analysed in a mass spectrometer across a high electrostatic potential, and are referred to as secondary ions (Fig. 1.26).

One of the main SIMS advantages is its sensitivity which means that samples with low concentration levels (down to ppb) can be analysed. As a result, the SIMS is used to determine trace elements in geochemistry and in the semi-conductor industry.

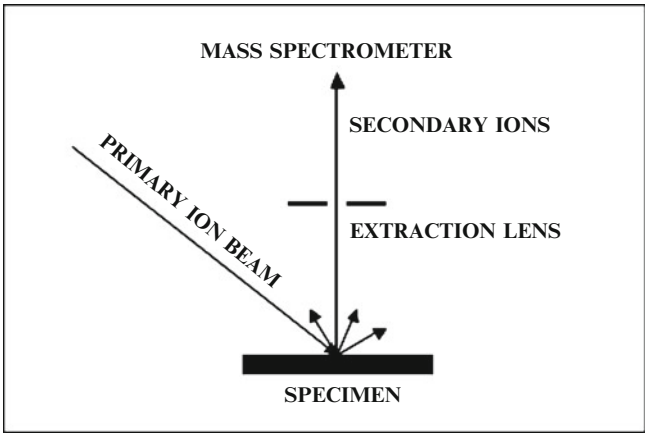


Fig. 1.26 Principle of secondary ion microscope mass spectrometer (SIMS)

Rutherford[7] *backscattering spectrometry* (RBS) is another analytical technique used to determine the composition of materials by measuring the backscattering of a beam of high energy ions impinging on a sample. Rutherford backscattering can be described as an elastic (hard-sphere model) collision between a high kinetic particle from the incident beam (the projectile) and a stationary particle located in the sample (the target). In practice a compositional depth profile can be determined from an RBS measurement. The elements contained in a sample can be determined from the position of peaks in the energy spectrum. Depth can be determined from the width and shifted position of these peaks, and relative concentration from the peak heights. This technique is very useful for studying nano-layers deposited on metallic surfaces for improving their hardness, their corrosion resistance or even their bactericidal properties as shown later in Sect. 1.2.3. As compared to SIMS technique RBS is not destructive.

1.2.2.7 Atomic Probe Tomography

The technique of *atomic probe tomography* (APT) is able to determine the spatial coordinates and the identities of the individual atoms in a metallic specimen close to atomic resolution (∼0.3 nm). The surface atoms of a cryogenically-cooled needle-shaped specimen are individually ionised by the application of a short duration (10 ns) high voltage pulse superimposed on a standing voltage on the specimen. These ions are then projected from the specimen in a time-of-flight

[7]Sir Ernest Rutherford (1871–1937) was a physicist sometimes referred as the father of nuclear physics. He won the Nobel price in chemistry in 1908 for his investigations in the disintegration of elements and the chemistry of radioactive substances.

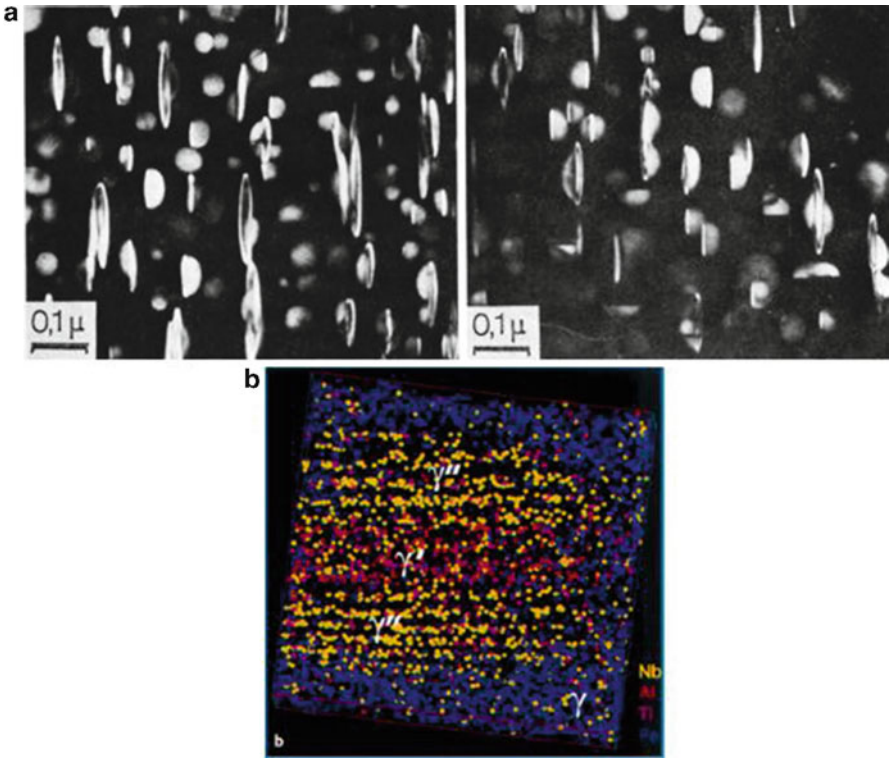


Fig. 1.27 (**a**) Dark field TEM observations in alloy 718 showing the co-precipitation of γ' ($L1_2$) and γ'' (DO_{22}) particles (Cozar and Pineau 1973); (**b**) Atom maps of 2 γ'' precipitates surrounding a γ' particle. The colours indicate the elements (Nb, Al, Ti, Fe) (Miller 2001)

mass spectrometer and detected with a single atom position-sensitive detector. In these three-dimensional atom probe instruments, the typical magnification of the specimen surface is $\sim 5 \cdot 10^6$ times. A full description of the technique may be found in a monograph by Miller (2000).

An example of the use of APT is given for the determination of the precipitates composition in Alloy 718 (Ni – 20Cr – 20Fe – 5Nb – 3Mo – 25Ti – 2Al). This material is a widely used niobium-containing nickel base superalloy that obtains its high temperature strength from a dispersion of lenticular DO_{22} – ordered γ'' precipitates and roughly spherical secondary γ' precipitates of $L1_2$ structure in an FCC matrix. The composition of the DO_{22} – ordered can be described as an $L1_2$ – ordered structure with a [1/2, 1/2, 0] displacement every other (001) plane. In addition the compositions of the DO_{22} – ordered and $L1_2$ – ordered structures can both be described as Ni_3(Al, Ti, Nb) with different levels of aluminium, titanium and niobium. In many cases γ' and γ'' are assembled, as illustrated in Fig. 1.27a which is dark field TEM image. Analysis of the compositions of the phases reveal that iron, chromium, molybdenum and cobalt preferentially partition to the γ matrix while

niobium and titanium partition to the $\gamma'$and γ'' precipitates with an enrichment in niobium in γ'' particles, as shown in Fig. 1.27b. The information derived from these observations allowed to develop a compact morphology (cube-shaped γ' particles covered by γ'' lenticular particles) with an improved thermal stability (Cozar and Pineau 1973).

1.2.2.8 X-Ray Tomography

Tomography is imaging by successive sections through the use of any kind of penetrating wave (X-ray, gamma ray, electron-positron annihilation, electrons, ions). The device used in tomography is called a *tomograph*, while the image produced is a *tomogram*. X-ray tomography is largely used for medical applications. Medical tomography is based on sectional images through a body by moving an X-ray source and the film in opposite directions during the exposure. The image contrast is due to the difference in absorption between the various components of the body. Recently the development of high energy X-rays sources produced by synchrotron has led to the use of a new technique called *X-ray tomographic microscopy* (*SRXTM*). This technique allows for detailed three-dimensional scanning of metals. The X-ray source is fixed and the specimen is progressively rotated. The energy of synchrotron equipment allows for the use of specimens of metallic materials with a section of the order of a few millimetres, depending on the nature of the materials. The lateral resolution is continuously increasing and is now of the order of 0.6 μm.

The SRXTM is a very powerful technique for *in-situ* observations when combined with mechanical tests, such as tensile or even fatigue tests. This technique is under a considerable development and constitutes one of the dreams of the researchers in materials science, which is the *in-situ* observation of solid materials.

1.2.2.9 Digital Image Correlation and Full-Field Measurement Techniques

Digital image correlation (DIC) is initially an optical method that employs tracking and image registration techniques for accurate two- and three-dimensional measurements of information including deformations and displacements from the digital images. DIC technique has been developed in the early 1980s and has been optimised and improved in recent years (see a recent book on this subject edited by Grédiac and Hild 2011). During the last two decades, the improvements in image processing with microcomputers and the advances in camera technology have largely contributed to the development of non contact measurement techniques such as DIC, but also *moiré interferometry, speckle, microgrid* method or *infrared thermography*. All these techniques are becoming more and more popular in the experimental solid mechanics community. This is clearly illustrated by the increasing number of research papers in which these techniques are used to improve the knowledge, at various scales, of the mechanical behaviour of materials and structures.

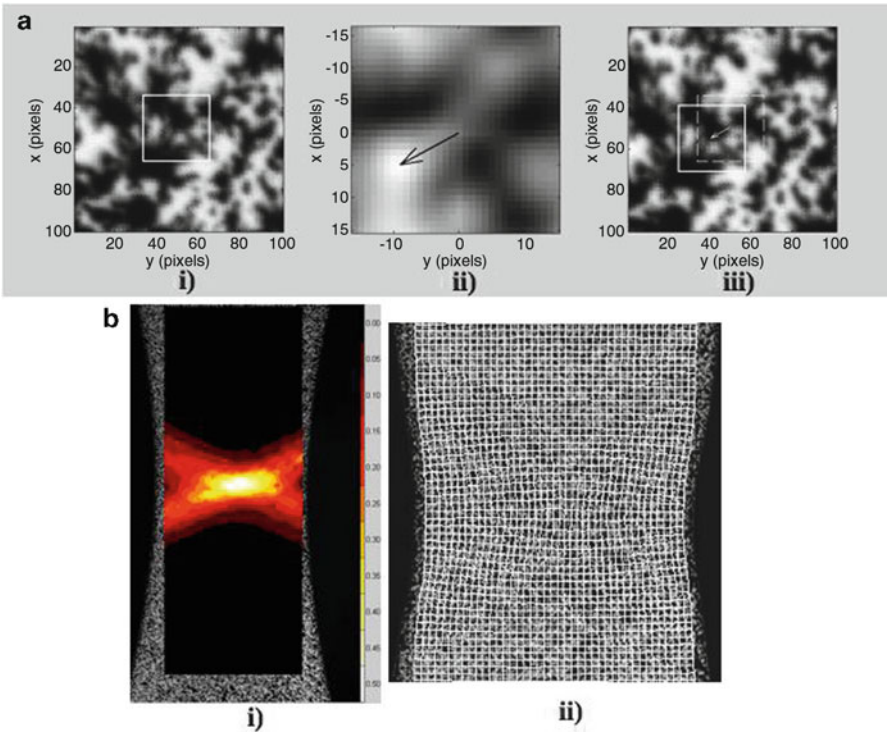


Fig. 1.28 (**a**) Basics of digital image correlation (DIC). (i) Reference configuration with a zone of interest (ZOI) (box). Random texture created by spraying black and white paint; (iii) Deformed configuration. The dashed box shows the location of the ZOI when no motion is assumed. In the present case, the average motion of the considered ZOI is determined by maximising the correlation product of both ZOIs (ii). Its maximum allows us to estimate the mean translation (*black arrow*). The solid box (ii) shows the location of the ZOI in the deformed configuration for which the measured translation was applied (*white arrow*); (**b**) (i) Longitudinal strain measured in a tensile test on 18MND5 steel. Strain localisation occurs in the central part of the sample where necking is located. The mean strain over the whole analysed region is equal to 5%, whereas the local strain exceeds 50%. (ii) Corresponding deformed mesh superimposed on the picture in the deformed configuration. The initial mesh was made of identical Q4 squares (by courtesy of Hild)

DIC technique is based on the maximisation of a correlation coefficient that is determined by examining pixel intensity array subsets on two or more corresponding images and extracting the deformation mapping function that relates the images (Fig. 1.28). An iterative approach is used to maximise the correlation coefficient by using nonlinear optimisation techniques.

Several techniques are used to prepare the specimen surface for image contrast analysis. The *embedded microgrid* (EMG) technique is worth being mentioned with this respect. EMG, combined with digital image processing of high definition SEM images, is now largely used to investigate, at the micrometre scale, the deformation mechanisms of materials. Microgrids are deposited on the surface with a pitch of

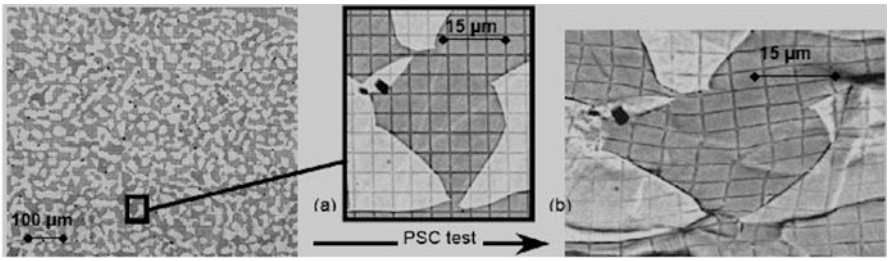


Fig. 1.29 Microgrids used to determine the local strains in a duplex (austenite + ferrite (darker phase)) deformed at 850°C under plain strain compression (30%, 1 s^{-1}); (**a**) Initial; (**b**) Final microstructures (Hernandez-Castillo et al. 2006)

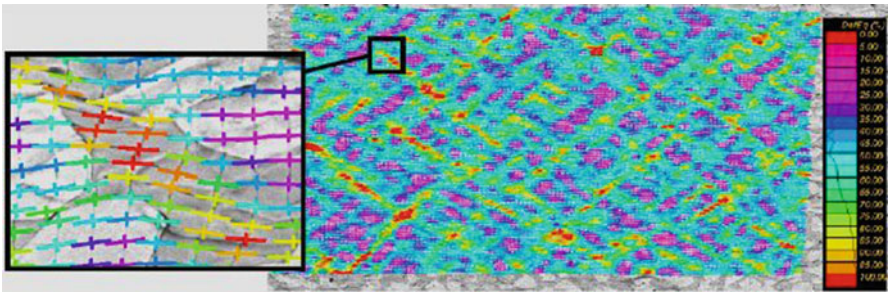


Fig. 1.30 Microgrids used to study the local deformations in a duplex (austenite + ferrite) stainless steel. The colours indicate the level of von Mises equivalent plastic strain superimposed on the deformed microstructure (Hernandez-Castillo et al. 2006, Pinna et al. 2000)

a few micrometres, as shown in Fig. 1.29 where the EMG technique is applied to the study of the deformation at high temperature, in plane strain compression, of a duplex (ferrite + austenite) stainless steel. These pictures clearly show many details in particular the occurrence of sliding at a number of phase boundaries. Figure 1.30 provides an example of strain map obtained on a larger area with 111×122 grid intersections (*i.e.* 500×550 μm^2). Deformation bands are clearly observed as in other materials. This information is very important for modelling the mechanical behaviour of materials.

1.2.3 Examples of Microstructures

The mechanical behaviour of a material is closely determined by the microstructure and this in turn is affected by the method(s) used in its manufacture: for metal alloys, casting, powder metallurgy, forming in the solid state, assembly by welding, etc.; for polymers, transformations in the liquid state, possibly followed by crystallisation;

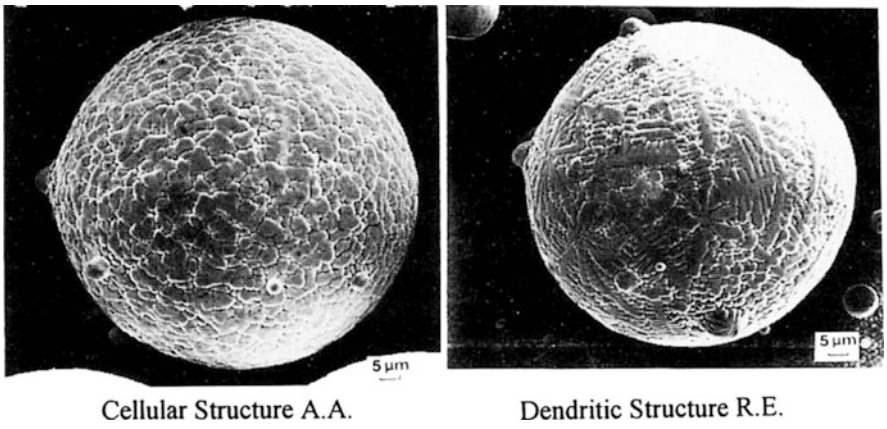


Fig. 1.31 SEM pictures of two particles of a nickel-based powder, produced by two atomisation processes: rotating electrode (RE) and argon atomisation (AA)

for ceramics, sintering, changes during their lifetime (ageing). We now give a few examples to illustrate the influence of these factors and follow these with a discussion of some microstructural features related to deformation, damage and fracture. A number of illustrations showing the microstructure of materials using various observation techniques have already been given in the preceding subsection.

1.2.3.1 Relations Between Microstructures, Fabrication and Ageing of Materials

Casting is still the most commonly used method for forming objects in metallic alloys, at least in the early stages. Since solidification necessarily involves the appearance of major and minor segregations – see Annex 2 – the parameters of the casting process must be controlled very carefully.

When a liquid is cooled very quickly, for example by atomisation, the size of these segregations can be limited and then, using the techniques of powder metallurgy, such as sintering, forging, extrusion or hot pressing, new alloys can be produced that could not be made by the classical method of casting followed by solidification.

The powder metallurgy of alloys of aluminium and nickel in particular has seen considerable developments. Figure 1.31 shows scanning electron microscope pictures of two particles of a *nickel-based alloy* (Astroloy: 55 Ni, 15Cr, 5Mo, 17Co, 3.5Ti, 4Al) obtained from different atomisation processes: rotating electrode, in which one electrode is heated and the liquid formed on it is thrown off by the very high speed of rotation and forms small droplets which cool very quickly; and argon atomisation, in which a jet of argon is blown through a bath of the liquid alloy. These pictures show different surface microstructures, the first cellular and the second

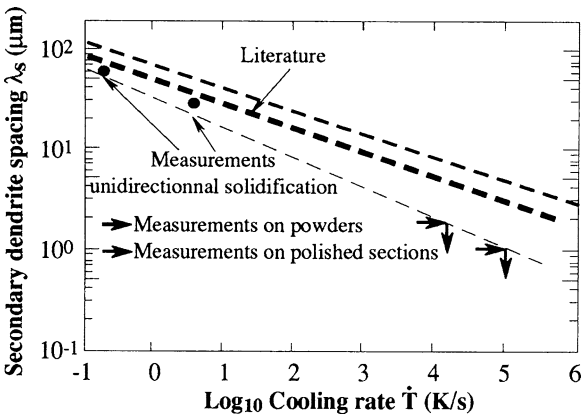


Fig. 1.32 Variation of the distance λ_S between the arms of the dendrites with the solidification rate

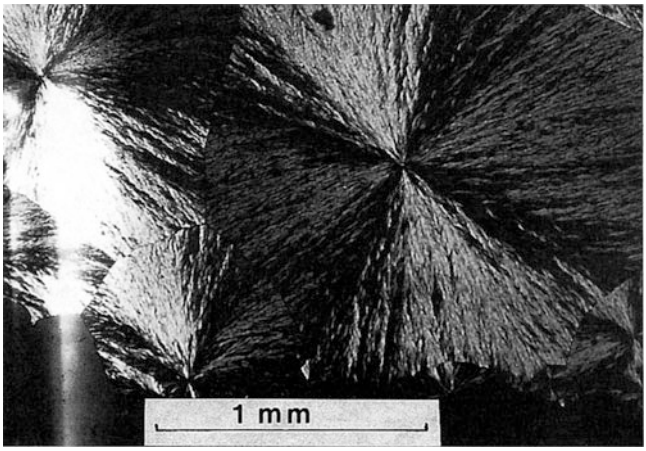


Fig. 1.33 A spherulite of polypropylene heated to 200°C viewed in the optical microscope, in transmission and in polarised light

dendritic. Figure 1.32 gives, for this alloy, the relation between the spacing of the arms of the dendrites (λ_s) and the cooling rate ($\dot{T}$); it shows that these atomisation processes give very high cooling rates, 10^4–10^5 °C/s, and consequently very small distances between the segregation zones, a few microns. A natural conclusion is that these processes can give a very homogeneous material.

Our second example concerns *crystalline or semi-crystalline polymers* such as polypropylene. This material is largely used in particular in the automobile industry for components such as shock absorbers, side-protection bars, etc. Typically, polypropylene consists of about 50% of a crystalline phase distributed through the material, under certain production conditions, as spherulites visible with an optical microscope when viewed in transmission with polarised light. Figure 1.33 shows an example.

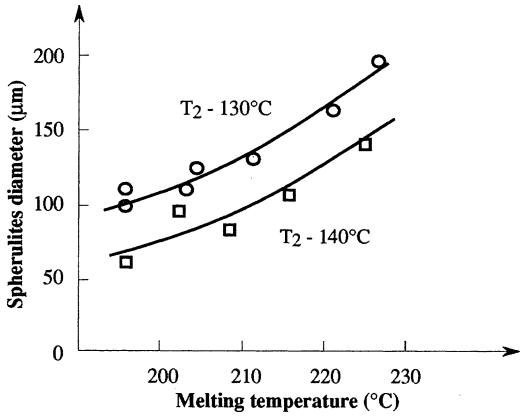


Fig. 1.34 Diameter of polypropylene spherulites as a function of the temperature of the melt, for two crystallisation temperatures, T_2, 130°C and 140°C

The crystallisation of *polypropylene*, which has a melting point close to 160°C, is a function of a number of parameters including the molecular mass, the temperature to which the molten polymer has been raised, especially the temperature at which the crystallisation takes place. Figure 1.34 shows that the size of the spherulites increases with the temperature at which the material was melted and decreases with that of crystallisation.

Heterogeneity is an aspect of materials with which this book is very much concerned. It is often something that is suffered rather than desired. (However, nanostructured materials largely used in particular in the electronic industry are deliberately heterogeneous). The structure of *wood* is an example, or of *cork*, shown in Fig. 1.35. This is a natural cellular material with a small cell size of about 50 μm, and lends itself well to the methods for homogenisation of periodic-structure materials. Other *cellular materials*, artificial in this case, have a coarser structure, as for example the honeycomb structures that can be created in aluminium (Fig. 1.36a), in ceramics (Fig. 1.36c) or in polymers (Fig. 1.36b).

Assembled material by one or other of the processes available such as welding, brazing, sticking, etc., have properties, which clearly will depend on the microstructures formed at the component interfaces or with the material used for the assembly. In order to take advantage of the best features of this or that material, such as the resistance to wear of ceramics or the toughness of metal alloys, methods of assembly have been developed over the years for "marrying" materials whose physical and chemical characteristics are not easily reconcilable, for example a metal and a ceramic. Figure 1.37 shows how an intermediate thin layer of aluminium can be used to attach the ceramic zirconia (ZrO_2) to a metal, steel or cast iron; this requires the development of a reaction zone whose extent, about 50 μm, must be precisely controlled if defects are to be avoided.

Nano-structured materials are intensively investigated since the 2000s. Two examples illustrating the benefit of high resolution analytical and microstructural observations are given below.

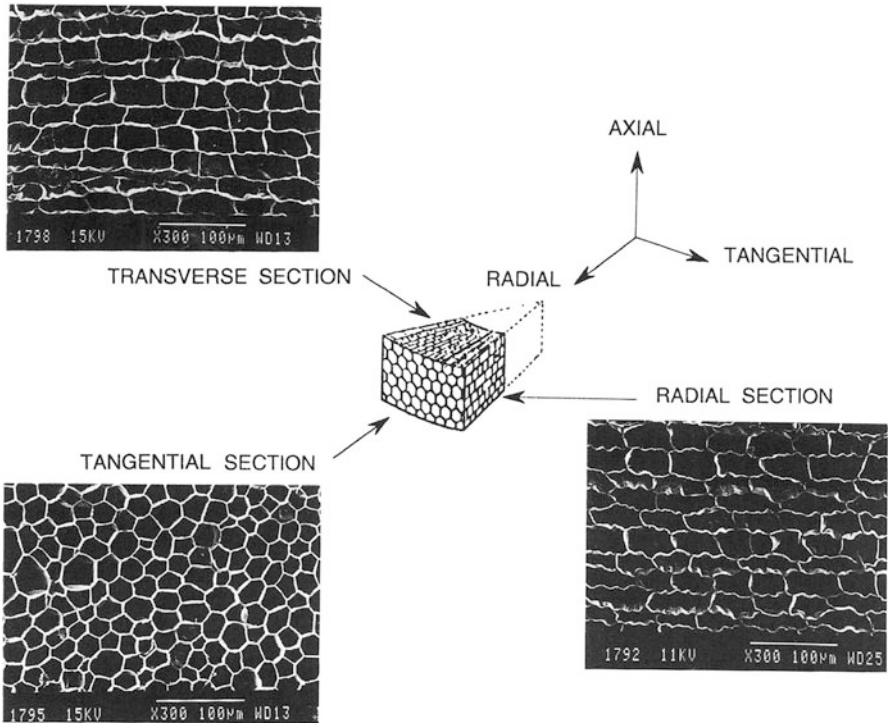


Fig. 1.35 Structure of cork: three mutually orthogonal sections

The first example deals with nano-structured semiconductor ceramics TiO_2 – SnO_2, which are promising candidates for new thermo-electrics with improved properties. Superlattices alternating very thin layers of different compositions have been shown to drastically reduce thermal conductivity by phonon scattering while enhancing electrical conductivity by charge confinement. It is well known that in metallic materials these two transport properties are always working hand in hand. Encouraging results to "marry" these two antinomic physical properties have been obtained recently in a SnO_2 – TiO_2 system (de la Pena et al. 2011). These authors have developed a novel route to produce by aqueous co-precipitation of $Ti_{0.5}Sn_{0.5}O_2$ nano-powders as precursors for the elaboration of dense ceramics. The nano-structures developed by spinodal decomposition (see Annex 2) are shown in Fig. 1.38. In this figure alternate lamellars of SnO_2 and TiO_2 oxides are observed. High resolution TEM (HRTEM) observations in [010] zone axis shows that the interface between the two ceramics is perfectly coherent.

The second example deals with physical vapour deposition (PVD) of TiO_2 films on carbon membranes fixed on a flat sheet of stainless steel. These films have photo-catalytic properties, which give them anti-bactericidal properties. The surface "functionalisation" of steel substrates by photo-catalytic TiO_2 coatings allows for

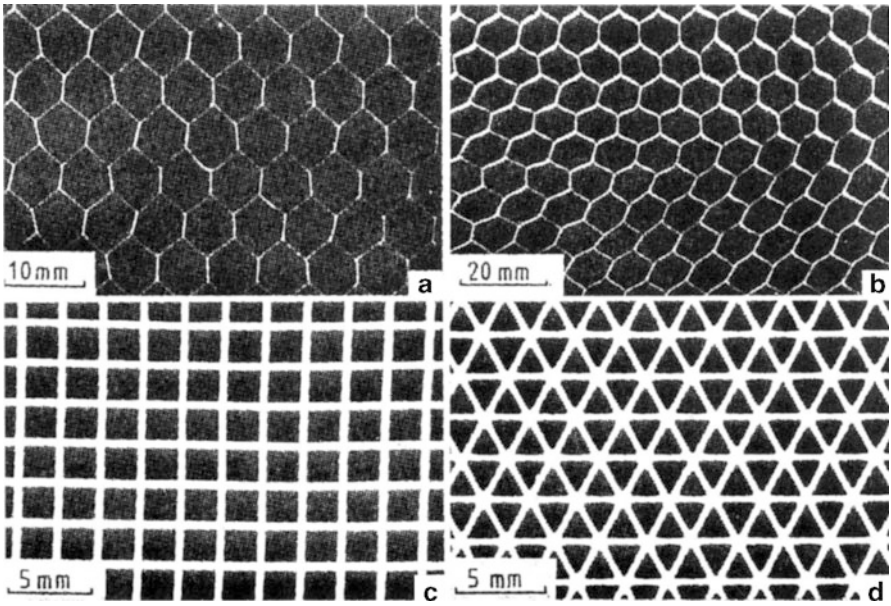


Fig. 1.36 Examples of honeycomb cellular structure: (**a**) aluminium (**b**) phenolic resin (**c&d**) ceramic, with square and triangular cells (Gibson and Ashby 1988)

the development of new steel products with added value, new properties and new applications such as air or water purification, deodorising, bactericidal effect and self-cleaning. TiO_2 coatings can be prepared on steel substrates by a plasma vapour deposition process. Optimum properties are obtained when the thin film is formed of titanium oxide grains within an amorphous matrix phase. Figure 1.39 shows a 15 nm crystallised grain of anatase phase.

The characteristics of a material are affected by the way its microstructure changes under heat treatment, or with ageing in use. Often these developments, such as changes in the hardening phases or growth in precipitate size, can only be detected by TEM at high magnification.

It is usually very difficult to predict the behaviour of a material that ages in use. However, much progress has been made only as a result of detailed study of the mechanisms of phase change and of the causes of deformation and damage. This is illustrated by the three examples that follow.

The first example concerns the microstructures of *nickel alloys* used for the fabrication of turbine blades for jet engines. Figure 1.40 shows the effect of the chromium content on the carbide-forming elements (Cr, Ti, Mo, W, Te), and of the elements (Ti, Al, Nb or Ti) which confer hardness by forming the γ' phase [Ni_3 (Ti, Al)]. This figure gives also some examples of alloys with the largest amounts of this phase, such as MAR M 246 which can contain up to 50%. This figure illustrates the complexity of these alloys, which are still the best choice for the blades when

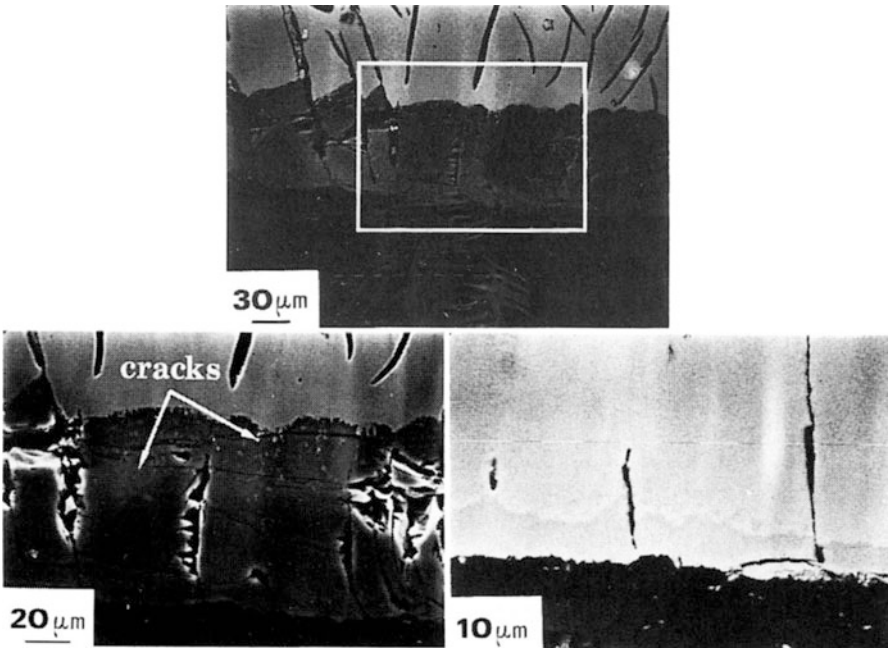


Fig. 1.37 Assembly of ceramic (ZrO_2) and steel or cast iron through the intermediary of a layer of aluminium; it shows the reaction zone, and cracks (by courtesy of C. Colin)

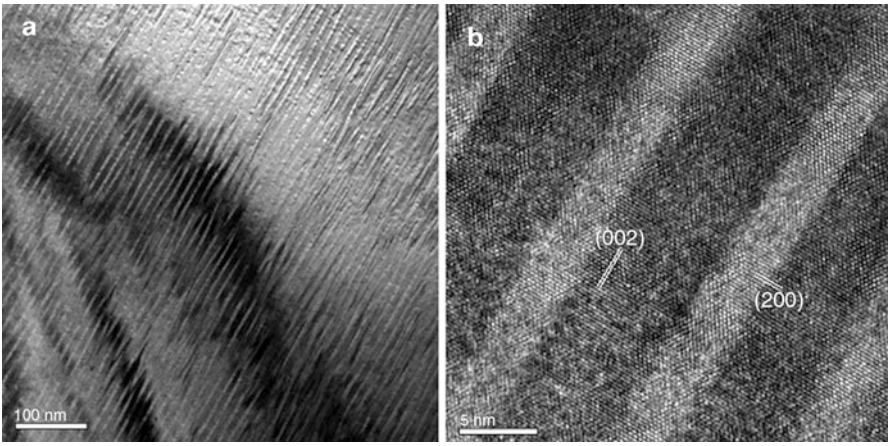


Fig. 1.38 TEM image of a SiO_2-TiO_2 ceramic sintered from co-precipitated powders. Microstructure obtained from spinodal decomposition. The HRTEM in [010] zone axis reveals coherent interfaces (de la Pena et al. 2011)

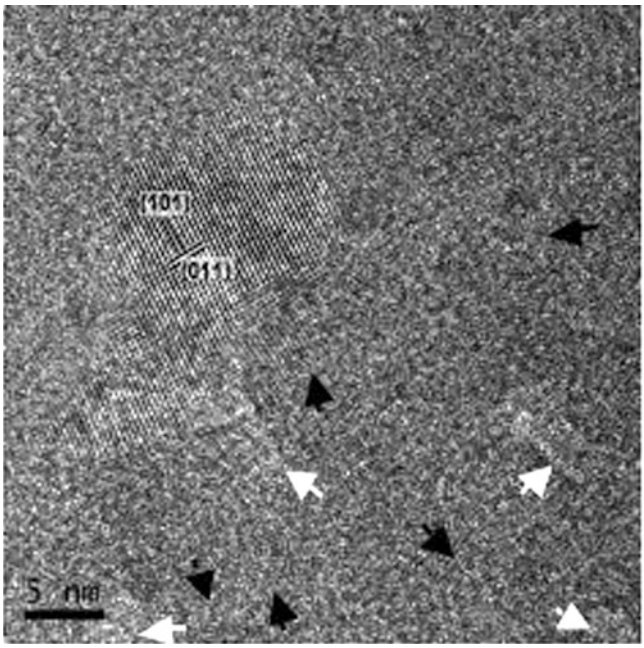


Fig. 1.39 Focussed High Resolution TEM picture showing an anatase grain. Black arrows point seeds of crystallisation, white arrows lighter second-phase (Berger 2011)

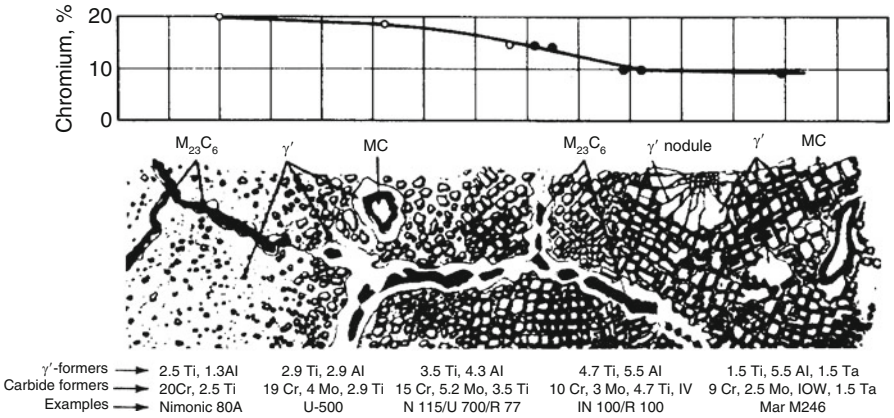


Fig. 1.40 Changes in the microstructure and chromium content of nickel-based superalloys (Brooks 1982)

these have to withstand high temperatures (>1,000°C) for long periods. Typically, the large particles of the γ′ phase are of the order of 0.5 μm, as can be seen from Fig. 1.41. All the phases (γ′, carbides) must maintain the highest possible stability under operational conditions.

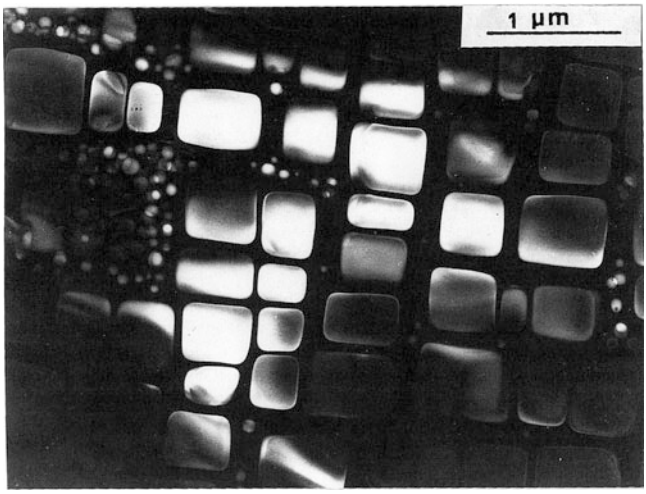


Fig. 1.41 TEM picture showing the microstructure of a monocrystalline turbine blade of the alloy AM1 (6.5 Co, 7.5 Cr, 2Mo, 5.5 W, 5.3Al, 1;2Ti, 8Ta, Bal Ni). The γ' precipitates are cubic in shape. The position of this alloy in the performance diagram of Ni-based Alloys is shown in Fig. 1.2.

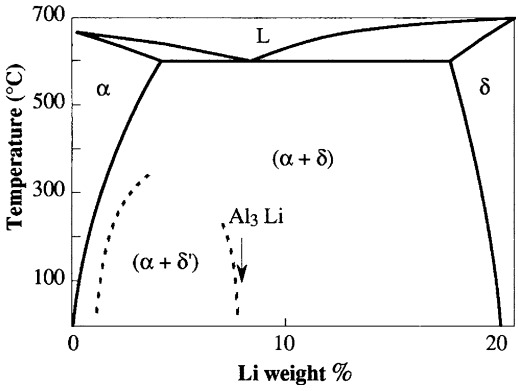


Fig. 1.42 Al-Li phase diagram: full lines correspond to equilibrium, dotted lines to the metastable δ' phase (Al_3Li)

Our second example concerns *light alloys*, based on Al-Li, on which the aeronautics industry relies, because of the compromise between strength and weight that they make possible. They can give reductions of 10–15% in the weight of an airframe compared with more traditional high-strength light alloys. There is now a strong competition between the new Al-Cu-Li alloys and the composite materials. The *Al-Li alloys* contain other elements such as copper, zinc and magnesium. The phase diagram of Fig. 1.42 (Al_3Li, stable for phase δ, metastable for δ') shows that these alloys are hardened during heat treatment by the precipitation of very fine particles (a few tens of nanometres) of the δ' phase. Other phases can

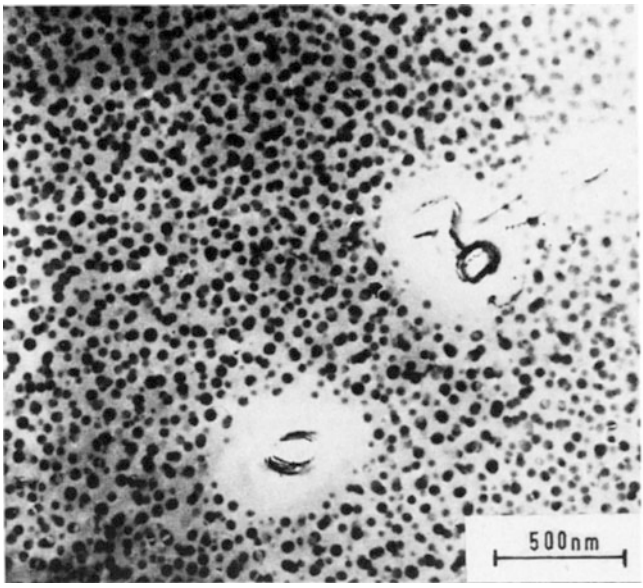


Fig. 1.43 Transmission Electron Microscope picture of an Al-Li-Mg alloy (Flower and Gregson 1987)

occur, however: Fig. 1.43 shows the formation of zones inside the grains free from precipitates of δ', in the neighbourhood of larger particles, of type Al_2 Mg Li. The presence of large numbers of such particles, at separations of 0.5–1 μm, reduces the strength of the alloy.

Light alloys involving lithium are very sensitive to intergranular fracture, which can make them brittle. One reason for this may lie in the microstructure of the grain boundaries, which must be carefully controlled. In alloys containing copper, zinc and magnesium in addition to lithium, particles rich in these elements and about 0.1 μm in size are found after certain heat treatments (Fig. 1.44). The same figure illustrates the value of an analysis that gives the enrichment of the particles in Zn, Mg and Cu (full lines) relative to the matrix (hachured zones). The local composition of the matrix is affected over a distance which, whilst small ($\approx$0.1 μm), is great enough to give a zone free of the δ' phase responsible for hardening in the interior of the grains. This "soft" zone can then be fractured preferentially, resulting in intergranular fracture, ductile in nature. Changes to the composition and to the heat treatment can initiate a *cellular reaction* of the δ phase, which appears at the grain boundaries as plates (Fig. 1.45). Precipitation of this phase, in this form, should be avoided, since it leads to a microstructure, which favours intergranular fracture.

Our last example concerns *composites with a metallic matrix*, materials now in the course of development. Composites with an aluminium or aluminium alloy

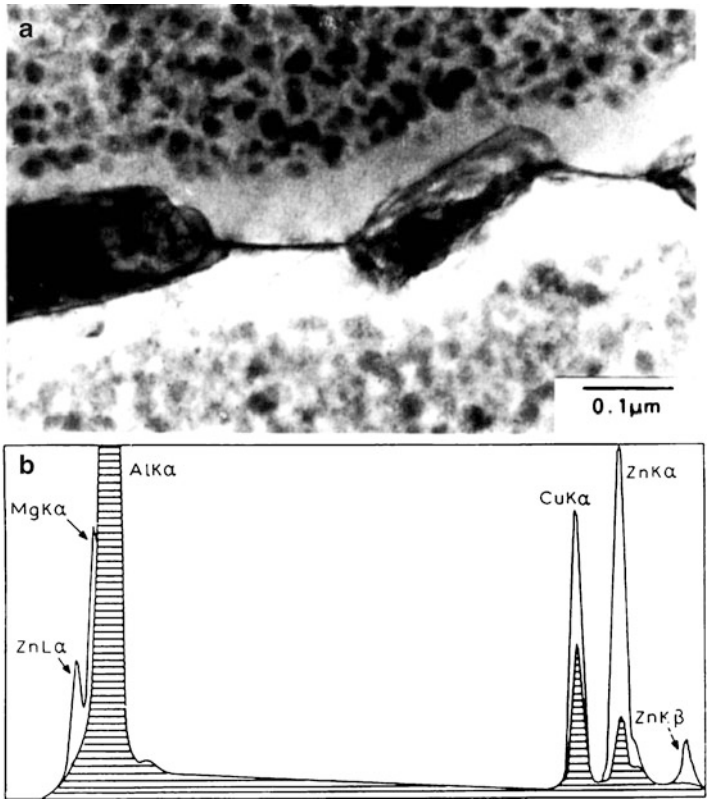


Fig. 1.44 Development of the microstructure near a boundary in an Al-Li-Mg-Cu alloy (**a**) formation of intergranular precipitates and a zone free of particles of the δ' phase (**b**) X-ray spectra of particles (*full lines*) and of matrix (*hachured*) (Gregson et al. 1987)

matrix, reinforced with silicon carbide (SiC) particles or whiskers, have been much studied; these can have up to 25% of the ceramic phase.

For certain uses the interface between the matrix and the particles in a material of this type must be as strong as possible, and this demands control of all the parameters of the fabrication. Figure 1.46 is a transmission electron micrograph of such an interface. The fine particles, in the form of platelets, correspond to the hardening phases of the binary Al-Cu system (see Annex 2). Following certain heat treatments, in particular keeping at high temperature for long periods (100 h at 500°C), the SiC phase, which is impure, may be oxidised locally; this is shown in Fig. 1.47, where the presence of particles of oxides such as MgO, Al_2O_3 and $MgAl_2O_4$ is revealed by electron micro-diffraction. The properties of the composite are greatly modified by such treatments.

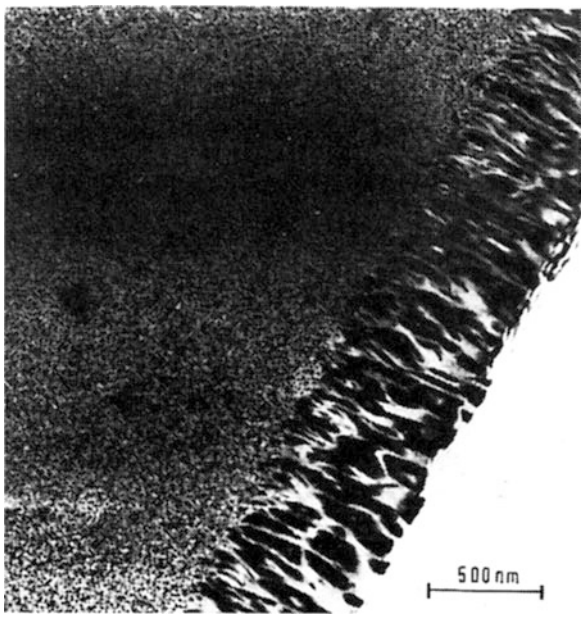


Fig. 1.45 Transmission electron micrograph of Al-3Li (% by weight) alloy, showing precipitates of δ phase located at grain boundaries (Flower and Gregson 1987)

1.2.3.2 Relations Between Microstructure and Behaviour, Damage and Fracture

The knowledge of the elementary micromechanisms of deformation is required for the development of macroscopic constitutive equations. This broad philosophy, which is the main concern of the whole books, is discussed and illustrated in the next few chapters; the methods of observation we have described are the favoured tools for the study of the elementary mechanisms. In this introductory chapter we shall limit ourselves to showing, with the help of two examples, with what a variety of size scales we constantly have to deal when we concern ourselves with the behaviour of polycrystalline aggregates.

The first level is the grain and the grain boundaries; or more generally, the interfaces between the different phases in polyphase alloys. Figure 1.48 is an optical micrograph of the polished surface of a two-phase alloy that has been deformed by tension. The material is a *cast stainless steel* (18 Cr – 12 Ni – 2.5 Mo), consisting of a relatively soft austenitic matrix (γ, FCC) and about 20% of much harder ferritic phase δ, BCC). The traces seen in the two phases correspond either to slip lines or to mechanical twins produced by plastic deformation; and at least two slip planes can be seen in the two phases. The discontinuities in the traces at the δ/γ interface is due to the fact that the two phases have different crystal structures and consequently different slip systems. It is easy to appreciate that maintaining compatibility of deformation in the two phases poses a problem: this is particularly evident in the δ phase, where the relaxation of the compatibility conditions has been achieved by

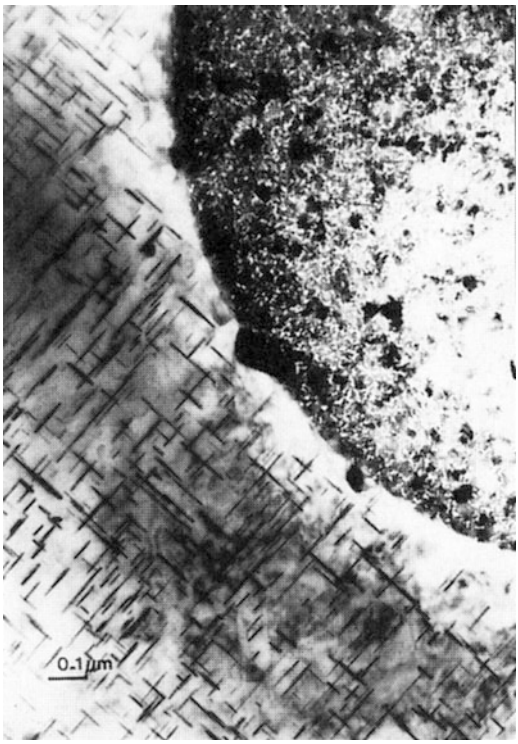


Fig. 1.46 Transmission electron micrographs of the interface between SiC fibres (*upper right part*) and the matrix of an aluminium-based composite material

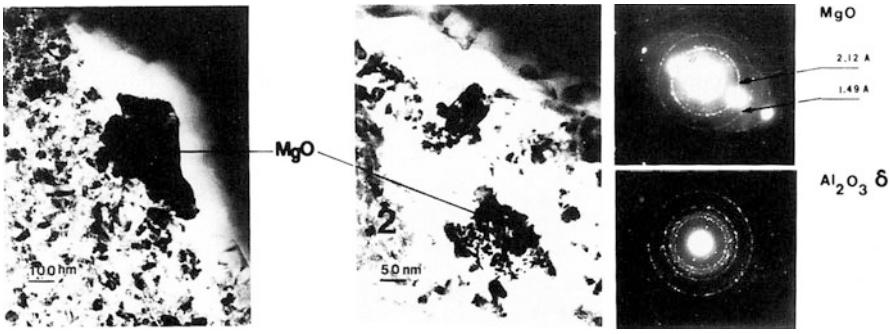


Fig. 1.47 Transmission electron micrograph showing, by electron micro-diffraction, oxide particles at the interface when the composite is aged at high temperature (100 h at 500°C)

the formation of a micro-crack in the δ phase. This micrograph illustrates also the first level in models of polycrystalline deformation, that of intergranular hardening.

At a lower level we must also concern ourselves with the distribution, more or less homogeneous, of the slip within the grains and with that of the dislocations. Figure 1.49 provides an example. This concerns a *nickel-based alloy*, Waspaloy

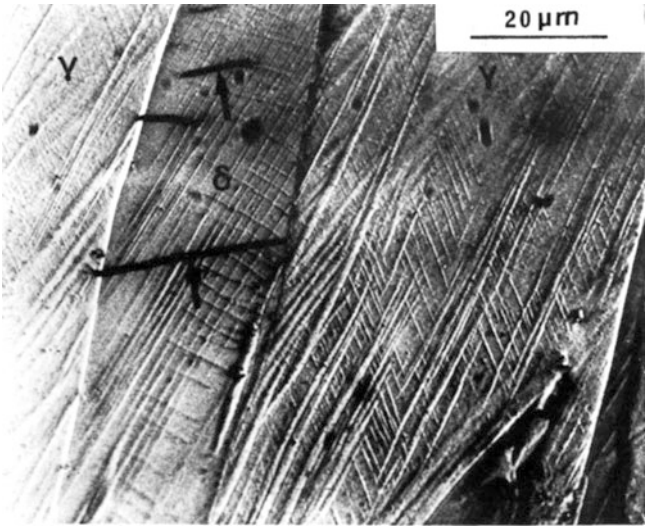


Fig. 1.48 Cast stainless steel, with about 20% ferrite (δ) in an austenite (γ) matrix , embrittled by accelerated ageing. Traces of slip planes and twinning in the two phases; formation of cleavage micro-cracks (shown by *arrows*) in the ferrite (by courtesy of P. Joly)

(18.7 Cr, 14.2 Co, 3.90 Mo, 3 Ti, 1;44 Al, Bal. Ni), deformed at ambient temperature by low-cycle fatigue (Clavel and Pineau 1982a, b). This alloy had been heat-treated in such a way that it contained two populations of the hardening phase γ'[Ni_3 (Ti Al)] with mean particle sizes 200 and 10 nm respectively: this can be seen in Fig. 1.49b, for which the conditions (dark field) were such that only the γ'phase was visible. Figure 1.49a shows that the deformation is localised in the clearly seen system of fine bands a few microns apart. At the still lower level of the region between a pair of bands, Fig. 1.49b shows that in this case the precipitates, initially spherical, have narrow slices cut through them. This micrograph illustrates another aspect of cyclic deformation, its more-or-less irreversible character.

Scientists and engineers specialising in Mechanics and Materials Science who have the task of developing new lines face the challenge of producing materials that have good resistance both to deformation and to damage and fracture. The study of the latter – damage and fracture – poses many problems related to scale effects; we discuss and illustrate the mechanisms associated with the various kinds of damage later; here we mention only two aspects:

– the strength of the grain boundaries in crystalline materials, especially at high temperatures, subjected to slowly-varying load (creep);
– the resistance to propagation of transgranular cracks in metal alloys subjected to cyclic load (fatigue).

The question of strength of *grain boundaries* arises in connexion with both metals and ceramic*s*. In the latter some boundaries include a very thin layer (a few

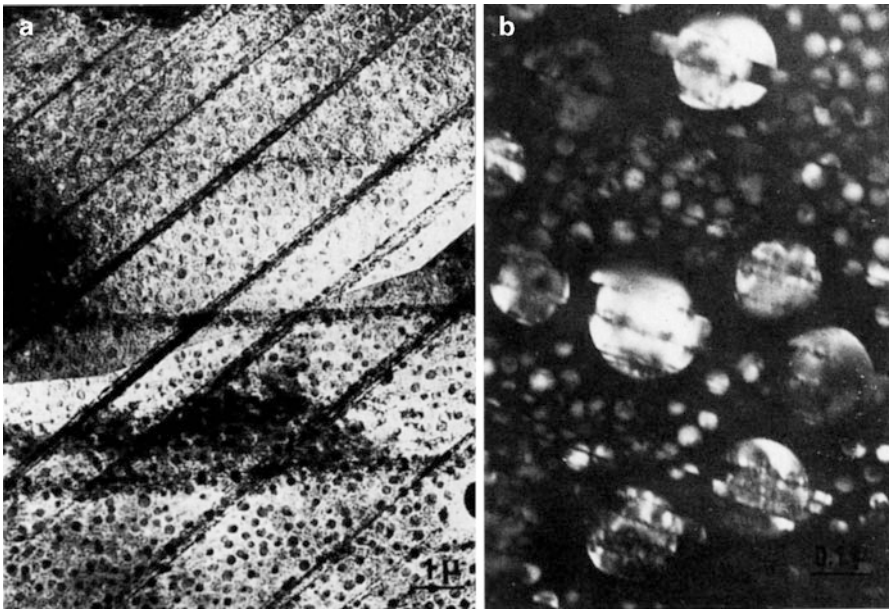


Fig. 1.49 Dark-field transmission electron micrograph showing the structure produced by cyclic deformation of Waspaloy (**a**) heterogeneous deformation distributed in bands along $\{111\}_\gamma$ planes (**b**) interior of a band showing slicing of the γ' precipitates (Clavel and Pineau 1982a)

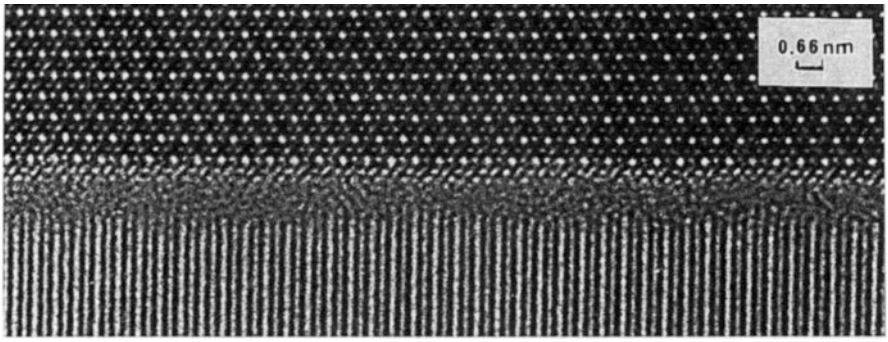


Fig. 1.50 Structure in the neighbourhood of a grain boundary in Si_3N_4; high resolution (0.2 nm) transmission electron micrograph showing rows of atoms

nanometres) of a phase that is not perfectly crystalline. Figure 1.50 illustrates this with the *silicon nitride* Si_3N_4: the "disordered" amorphous phase along the length of the boundary between the two crystalline phases is easily seen. At high temperature the existence of this phase would lead not only to viscous behaviour of the ceramic but also to *intergranular fracture*.

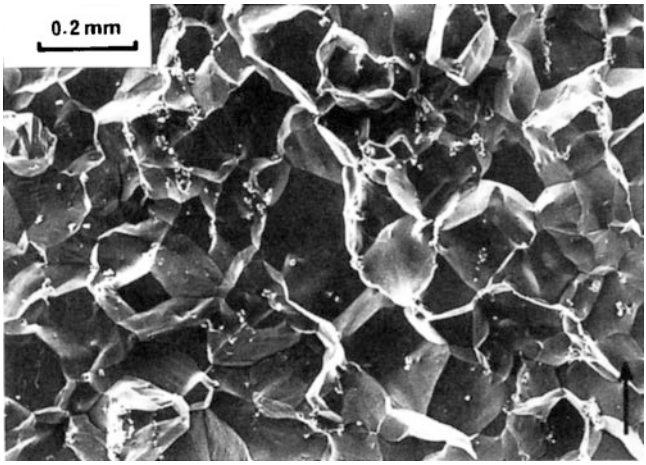


Fig. 1.51 Scanning electron micrograph of creep-induced intergranular fracture in alloy 718

The scanning electron microscope is the instrument now most used for studying fracture. Figure 1.51 shows *creep-induced intergranular fracture* in the nickel-based alloy 718 (18 Cr, 18 Fe, 3 Mo, 1 Ti, 5.2 Nb, Bal. Ni) used mainly for turbines. These fractures are often the result of coalescence formed at the grain boundaries, themselves initiated by particles lying along these. This is shown in Fig. 1.52 for a Cr – Mo – V steel used for steam turbines. This micrograph shows again the very heterogeneous distribution of the damage, and makes evident the underlying problems of scale. The number of such cavities could be used as a metallographic measure of the damage, but to find this would clearly need a great deal of care. However, we must have such measurements if we are to set up soundly based models to enable us to predict the lifetime of a component in service.

Examining the features of *fatigue-induced fracture* can not only be valuable as a help in identifying the underlying mechanisms but can also serve practical ends. The scanning electron micrographs of Figs. 1.53 and 1.54 provide examples. In Fig. 1.53 the vertical line probably corresponds to the emergence of a twin boundary at the fracture surface; on either side of this we see *striations* that are characteristic of a fairly high speed of propagation, about 0.1 μm per cycle. Figure 1.54 shows the completely different features sometimes called *pseudo-cleavage,* characteristic of the lowest fatigue crack growth rates, less than 10 nm per cycle. Careful examination of the central region of Fig. 1.54 will show that this feature consists of a number of faces, all corresponding to {111} slip planes in the same grain of this FCC material.

This is shown clearly by Fig. 1.55, an optical micrograph of a polished and etched section normal to the fracture surface, in which traces of the slip planes can be seen.

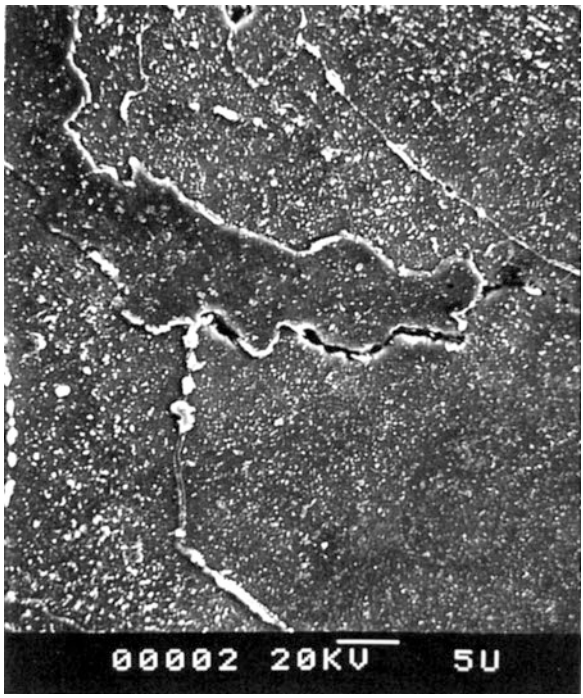


Fig. 1.52 Creep-induced intergranular cavity formation in steel (1 Cr, 1 Mo, 1.25 V). Some of the cavities are very small

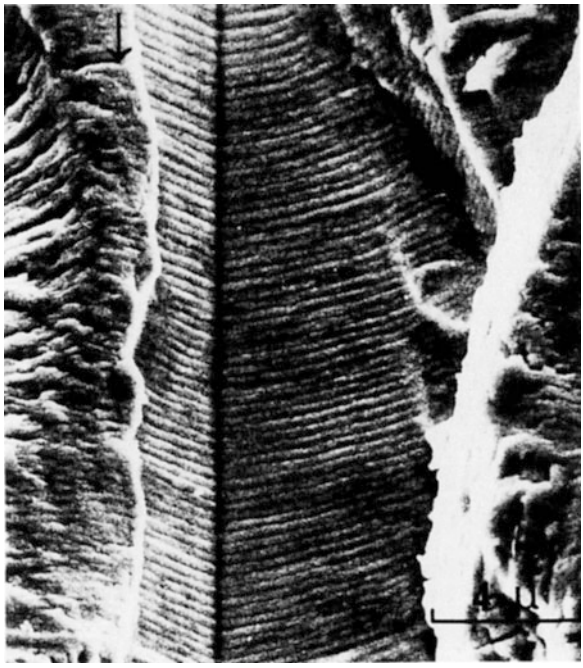


Fig. 1.53 Fatigue striations on either side of a twin boundary in alloy 718

Fig. 1.54 Scanning electron micrograph of fatigue fracture surface in Waspaloy (Clavel and Pineau 1982a, b)

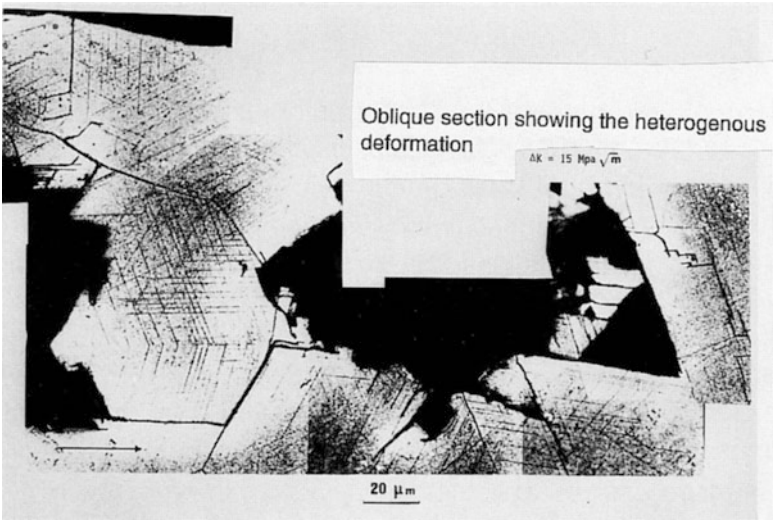


Fig. 1.55 Optical micrograph showing the section of a fatigue fracture surface in Waspaloy. Notice the very ragged surface, which in many places is parallel to the traces of the slip planes (Clavel and Pineau 1982a, b)

1.3 Characterisation of Mechanical Properties

1.3.1 Aim of Mechanical Tests

The engineer has to know the mechanical properties of the materials he intends using: in particular, for the prediction of the load that the structure he is designing will bear before starting to deform irreversibly or to be damaged and to fail. He must therefore have at his disposal results of mechanical tests that he can use in computing strengths of structures and designing components. In general, the relevant tests – tensile tests, for example – are simple in concept and interpretation. The stress states are simple and well known as in a tensile test. It is essential that laboratory tests are reproducible; therefore standards have been laid down by national and international bodies such as ISO (International Standards Organisation, www.iso.org), AFNOR (Association Française de Normalisation) and ASTM (American Society for Testing and Materials).

The research worker studying the mechanical behaviour of materials needs fuller and often more complex test data. Thus to find the yield conditions for a material under multiaxial loading requires something more complex than a simple tensile test, for example tension-torsion tests, or tests combining tension and internal or external pressure and performed on thin-walled tubes. He will need the same resources if his aim is to find laws that will describe the development of fatigue damage under a triaxial stress system, in particular the directions in which fatigue cracks initiate and grow; and similarly for damage and fracture resulting from creep. Yet are these tests isothermal. A researcher working on thermo-elasto-viscoplasticity is likely to have to develop his own methods.

It is perhaps rather strange that the resources available for such testing are still so limited, in view of what is at stake on the one hand and on the other the widening gap between the computing power now available for design of components and the amount of test data available to the designers. Thus there are only a few machines capable of performing in-plane tension-compression tests in two orthogonal directions; the number of tension-torsion machines is a little higher, but still small.

Servo-hydraulic and servo-controlled electromechanical machines are available in most laboratories. We therefore explain the principles on which they work and then describe the tensile test in more detail, followed by short accounts of other tests.

1.3.2 Servo-Controlled Testing Machines

The complete system is shown diagrammatically in Fig. 1.56. The test piece is held between pairs of jaws, one attached to the cylinder, which applies the load, the other to the moveable cross-head in which the load-measuring equipment is held. The jack can be either mechanical or hydraulic and can apply the load either linearly

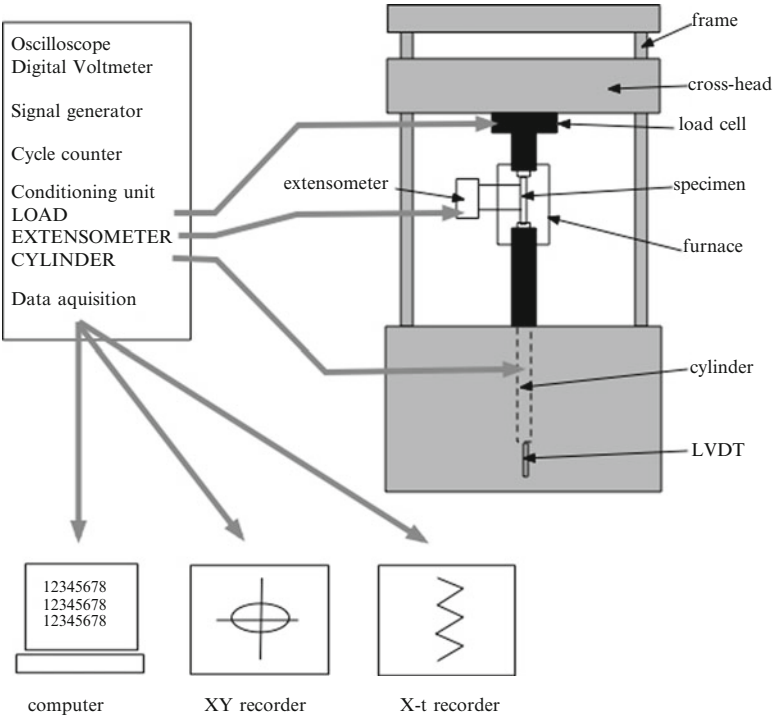


Fig. 1.56 Operation principle of a servo-controlled testing machine

(tension or compression) or rotationally (torsion); its displacement is picked out by a sensor (LVDT). The deformation of the test piece is recorded by an extensometer, usually mechanical but sometimes optical. It can be designed in such a way that the test piece could be heated in a furnace.

The servo-control can be based on the applied load, the displacement of the cylinder, the deformation of the test piece or any other measuring device; a signal generator (analogue or digital) enables, in principle, any desired variation of load with time, including cyclic, to be realised. Recording and processing of the results are becoming increasingly computer-based, but many installations still have X-Y and X-t recorders and oscilloscopes as output devices.

1.3.3 Tensile Test

1.3.3.1 Test Principle

The tensile test is the commonest mechanical test (François 2008; ISO 2009a); it consists, simply, of applying a tensile load F to a sample of the material and

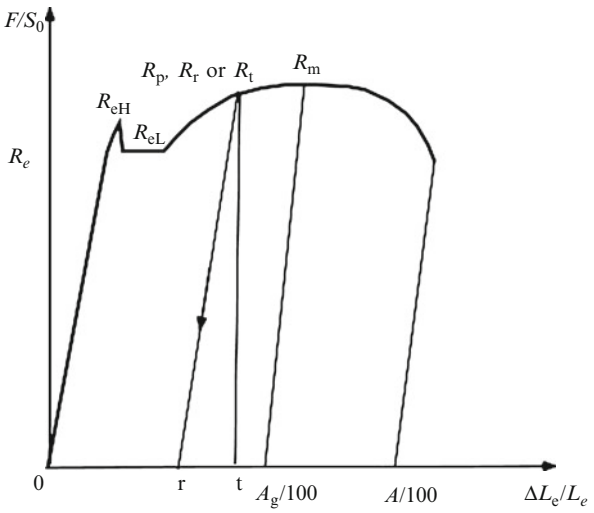


Fig. 1.57 Standard tensile test data

measuring the resulting increase in length ΔL. According to the type of testing machine used, the measurement is performed in load or in elongation rate control.

The *specimen* (ISO 1997) can be either cylindrical or prismatic; measured between reference marks it will have initial length and cross-section L_0 and S_0 respectively. As a general rule $L_0 = k\sqrt{S_0}$, with $k = 5.65$ (that is $5(4S_0/\pi)^{1/2}$). The heads can be either threaded or flat, according to the kind of jaws in which they will be held for the test; they must be smoothly joined to the stem, to reduce the possibility of stress concentration. The extension is measured between marks inscribed on the specimen or by an extensometer.

It is most important that the upper and lower jaws are accurately aligned so as to ensure as nearly as possible pure traction, with no unwanted bending; but however great the care taken, the alignment will never be perfect. To counter this, articulated links are often introduced, such as universal and ball-and-socket joints. This in turn may not suffice for very anisotropic materials, such as single crystals, composites and strongly textured sheets, or when the direction of the tension does not coincide with an axis of symmetry of the material. *Off-axis testing* requires special equipment.

Testing can be performed at *elevated temperature* (ISO 2009b) by surrounding the test piece with a furnace, or at *low temperature* (ISO 2000a; ISO 2004) by immersing in a cooling bath.

1.3.3.2 Standard Notations

Figure 1.57 is a schematic record, which could be obtained in a tensile test featuring the standard notations. We can derive from the curve the values of a number of quantities, characteristics of the material and of importance for the engineer.

Records completed during the test give the force measured by the load cell in relation to the initial cross-sectional area S_0 of the parallel length of the test piece, according to the elongation related to the initial length between marks L_0 or to the gauge length of the extensometer L_e. These are engineering values of the stress and strain. The standard specifies the following standard data (Fig. 1.57):

Upper yield strength: R_{eH}. This is the engineering stress value when a load drop is observed for the first time.

Lower yield strength: R_{eL}. This is the minimum engineering stress value during deformation.

Proof strength, non-proportional extension: R_p. This is the stress for a given conventional value of plastic percentage extension (extension is the increase of the extensometer gauge length L_e at a given moment of the test). In general, 0.2% is adopted and, in this case, the proof strength is designated by $R_{p0.2}$. For another value of plastic deformation, the index will be modified accordingly.

Tensile strength: R_m. This is the value of the engineering stress at the maximum force during the test.

The proof strength for a plastic percentage extension x is designated by R_{px}. In limit load computations, an average plastic flow stress value R_f is often used, which is equal to the average between the proof strength $R_{p0.2}$ and the tensile strength R_m. It is also possible to define the permanent set strength R_r, which is the engineering stress for a given residual percentage extension after removal of the load, as well as the proof strength total extension R_t, which is related to a total given extension (elastic plus plastic).

True tensile strength R_u is equal to the load quotient at the moment of fracture in relation to the fracture area.

Percentage elongation after fracture A is equal to 100 times the engineering strain at the moment of fracture.

Percentage elongation non-proportional at maximum force A_g is equal to 100 times the corresponding engineering plastic strain. Since plastic deformation takes place at constant volume, A_g is also almost equal in absolute value to the percentage relative variation of the cross-sectional area of the test piece. It can be measured after fracture along the parallel length outside of the necking zone.

The percentage reduction of area Z is equal to the maximum relative variation of the cross-sectional area of the test piece measured after fracture.

Can also be defined the *limit of proportionality* $R_e = F_e/S_0$, where F_e is the load at which the proportionality of extension to tension (Hooke's law[8]) ceases to hold.

[8]Robert Hooke (1635–1703) was an English experimental scientist who discovered the law ("ut tensio sic vis") which bears his name in 1660.

1.3.3.3 Analysis of the Tensile Curve

Analysis of a tension – elongation curve is simple in the case of a homogeneous, isotropic material. We define the *true* stress σ and strain ε at all points before the maximum of the curve – that is, before the onset of necking – as follows, taking care to use the extension after removal of the load so as to keep the plastic deformation only:

$$\sigma = \frac{F}{S} = \frac{F}{S_0}\exp(-\varepsilon) \tag{1.1}$$

$$\varepsilon = \log(L/L_0) = \log(1 + \Delta L/L_0) \tag{1.2}$$

Note that this definition of the strain is not exactly the same as the Lagrange[9] strain $\Delta = \Delta L/L_0 + (1/2)(\Delta L/L_0)^2$. The advantage of the definition (1.2) is that the strains are additive.

Various empirical stress–strain relations are used to represent mathematically the experimental curve.

The most widely used is the Hollomon[10] formula:

$$\sigma = \sigma_0 \varepsilon^n \tag{1.3}$$

A similar formula is the Ramberg[11]-Osgood law:

$$\frac{\varepsilon}{\varepsilon_0} = \alpha\left(\frac{\sigma}{\sigma_0}\right)^N \tag{1.4}$$

Here σ_0 is the yield strength R_p and ε_0 the corresponding strain, namely σ_0/E. The Ludwick's law is also often used:

$$\sigma = \sigma_0 + \sigma_1 \varepsilon^n \tag{1.5}$$

σ_0 and σ_1 having here a different meaning.

The strain rate sensitivity can be introduced in the representation of the tensile behaviour, writing then:

$$\sigma = \sigma_0 \varepsilon^n \dot{\varepsilon}^m \tag{1.6}$$

$\dot{\varepsilon}$ being the strain rate.

---

[9] Joseph Louis, count of Lagrange (in Italian Giuseppe Lodovico Lagrangia), was born in Turin in 1736 and died in Paris in1813. He was a mathematician and astronomer. Born in Italy, he spent 30 years in Piémont then 21 years in Berlin and the rest of his life in Paris.

[10] John Herbert Hollomon Jr. (1919–1985) was an American engineer.

[11] Walter Ramberg (1904–1985) was an American scientist.

The deformation remains homogeneous so long as the relation $d\sigma/d\varepsilon > \sigma$ holds. Necking occurs when $d\sigma/d\varepsilon = \sigma$, the *Considère's criterion.*[12] For materials for which the law of work-hardening can be put in the form $\sigma = \sigma_0 \varepsilon^n$, this occurs when the homogeneous deformation is equal to:

$$\varepsilon_g = \log\left(1 + A_g/100\right) = n \tag{1.7}$$

For a strain-rate sensitive material, whose behaviour can be represented by Eq. 1.6, the homogeneous deformation depends on the exponent m. The relation can be determined by studying the stability of a slight local perturbation of the cross-section of the specimen (Hutchinson and Neale 1977). Intuitively, it can be surmised that high rate sensitivity increases the homogeneous deformation because a local decrease of the cross-section increases the strain rate and thus the stress. The relation is:

$$\varepsilon_{hg} = \frac{n}{1-m} \tag{1.8}$$

Considère's criterion: necking of a tensile specimen begins when the homogeneous deformation is equal to the strain-hardening exponent n (or to $n/(1\text{-}m)$ for a strain-rate sensitive material).

When $m = 1$ (corresponding to Newtonian flow in viscosity) the homogeneous deformation becomes infinite. The material is *super-plastic*.

When necking occurs the specimen is no longer being subjected to a simple tensile stress state. Several approximate relations have been devised to describe the stress–strain state then holding in the minimal section of a circular cylindrical test piece, the one most commonly used being that proposed by Bridgman,[13] which assumes that the deformation is uniform over the minimal section (detailed calculation is given in Volume II).

Let R be the radius of curvature of the neck, and a the radius of the minimal cross-section. Then at any point distant r from the axis of symmetry, in polar coordinates:

$$\sigma_{rr} = \sigma_{\theta\theta} = \sigma_{zz} - \bar{\sigma} \tag{1.9}$$

where $\bar{\sigma}$ is the equivalent stress. And:

$$\sigma_{zz} = \bar{\sigma}\left[1 + \log\left(1 + \frac{a^2 - r^2}{2aR}\right)\right] \tag{1.10}$$

[12] Armand Considère (1841–1914) was a French engineer, a pioneer of reinforced concrete. He published his criterion in 1885.

[13] Percy Williams Bridgman (1882–1961) was an American physicist winner of the Nobel Prize in 1946 for his work on high pressures.

Fig. 1.58 Longitudinal section through a tensile tested specimen, test of which was interrupted before fracture, showing a crack in the central region of the neck

This corresponds to a uniform deformation at the section. It follows that the mean stress on the minimal section $\bar{\sigma}_{zz}$, measured beyond the maximum load, can be expressed as

$$\bar{\sigma}_{zz} = \bar{\sigma}\left(1 + \frac{2R}{a}\right)\log\left(1 + \frac{a}{2R}\right) \tag{1.11}$$

This relation enables, in principle, to determine the form of the work-hardening curve $\sigma = \mathrm{f}(\varepsilon)$, well beyond instability corresponding to the onset of necking, provided that we can measure the two radii a and R. This can be useful in the study of large deformations since for many materials the homogeneous deformation is small, *e.g.* $n \approx 0.1$ for ferritic steels.

The greatest principal stress σ_{zz} is maximum at the centre, $r = 0$ in Eq. 1.10.

Similarly for the stress triaxiality ratio:

$$(\sigma_m / \bar{\sigma})_{r=0} = (\sigma_{zz} / \bar{\sigma})_{r=0} - 2/3 = 1/3 + \log(1 + a/2R) \tag{1.12}$$

This result shows immediately why ductile fracture initiates preferentially at the centre of the neck, as shown in Fig. 1.58; this is the place where the hydrostatic stress σ_{m} is maximum.

In anisotropic materials, in particular in thin sheets of rolled aluminium alloys, or low carbon steels, tensile tests are also carried out to determine the *Lankford coefficient* (Lankford et al. 1950), also called Lankford R value. In a thin sheet rolled along the direction L, with T the transverse direction and S the thickness direction

(axis 3), a flat tensile specimen is cut along a direction, which makes an angle α with the L direction. The definition of the R value is:

$$R = \frac{\mathrm{d}\varepsilon^{\mathrm{p}}\,(\alpha + \pi/2)}{\mathrm{d}\varepsilon^{\mathrm{p}}_{33}} \tag{1.13}$$

where $\mathrm{d}\varepsilon^{\mathrm{p}}$ is the increment of plastic strain. In a purely isotropic material, $R = 1$. In usual metals, $R \neq 1$ even initially (*initial anisotropy*) and the R value changes with plastic strain (*induced anisotropy*). For sheet metals the R values are usually determined for three different orientations (0°, 45°, 90°) to the rolling direction and the mean value of R is taken as:

$$\bar{R} = \frac{1}{4}\,(R_0 + 2R_{45} + R_{90}) \tag{1.14}$$

The value of $\bar{R}$ is used extensively as an indicator of the formability of thin sheet metals. The values of R are determined experimentally using either contact extensometers or digital image correlation (see Sect. 1.2.2.9).

1.3.4 Compression Test

The compression test (ISO 1979) is similar in principle to the tensile test: the specimen, usually a circular cylinder, is submitted to two axially opposed forces by putting it between the tables of a press. This seems simple, but in practice there are at least two difficulties.

The first concerns the risk of *buckling*, an instability either elastic or plastic; to avoid this the aspect ratio (length/diameter) of the specimen must be kept below a critical value. The second concerns the friction between the press plates and the heads of the specimen. Under compression, the diameter of the cylinder increases as the length decreases, and this is opposed by the friction; the consequence is that the cylinder is distorted into a barrel shape. Further, it is difficult to use the test to find the ultimate strength under compression if the material is ductile. For this reason the compression test is most often used with brittle materials such as concrete and ceramics.

1.3.5 Other Tests

Many tests besides tension and compression are performed on materials; we shall just indicate the principles without going into details.

1.3.5.1 Bend Tests

The 3-point bend test (or 4-point, to develop a constant bending moment over a section of the specimen length) has the advantage of needing but simple shaped test pieces (ISO 2005a). Like the compression test, it does not in general enable fracture to be attained in ductile materials. These tests are currently used in quality control and to find the ultimate strengths of brittle materials.

1.3.5.2 Torsion Tests

Torsion tests are widely used to investigate the deformation of alloys at high temperatures in order to determine their constitutive laws and to study the development of their microstructures such as polygonisation and static and dynamic recrystallisation. For this purpose the tests are performed on cylindrical test pieces, most often at high strain rates ($\mathrm{d}\varepsilon/\mathrm{d}t \approx 1\text{–}10\ \mathrm{s}^{-1}$) so as to match the conditions which exist in forming operations. This deformation mode avoids necking, which makes it particularly interesting for the study of large deformations ($\varepsilon > 100\%$).

The main difficulty in this test is that the deformation is not uniform over the cross-section of the specimen, being maximum at the periphery and zero at the centre. In the regime of perfect plasticity, if R is the radius of the cylinder, C the applied torque and φ the angular strain, the maximum shear stress τ_{R} at the periphery is given by:

$$\tau_{\mathrm{R}} = \frac{1}{2\pi R^3}\left(3C + \frac{\mathrm{d}C}{\mathrm{d}\varphi}\right) \tag{1.15}$$

As we emphasised in Sect. 1.3.1, torsion tests are used in conjunction with tensile and compression tests to investigate the constitutive behaviour laws for materials, most often with small cyclic deformations. Here tubular specimens are used, with walls thin enough for the shear stress to be taken as constant over the section. The difficulty here concerns the instrumentation. Measurements can be made by strain gauges stuck to surface of the specimen if the test is to be made at or near room temperature and if, in a fatigue test, the interest is in the cyclic behaviour rather than in the damage. Indeed, gluing strain gauges does not allow perfect polishing of the surface of the test piece. Otherwise special measuring devices have to be used, of which there are two main types: mechanical, acting by direct contact, and optical, possibly using lasers. It is worth mentioning that digital image correlation (DIC) is a technique for strain measurements, which is now becoming more and more popular (see Sect. 1.2.2.9).

1.3.5.3 Hardness Tests

The hardness test, used mainly for control, measures the resistance of the material to penetration by a probe of standard shape and size (Pommier 2008); examples are

- hard steel ball, diameter 10 mm (Brinell test)[14] (ISO 2005b) or 1.59 mm (Rockwell B) (ISO 2005c)
- diamond cone, vertex angle 120° (Rockwell C) (ISO 2005c)
- diamond pyramid, square base, angle 136° between opposite faces (Vickers) (ISO 2005d)

The tests are simple and quick to perform and the results are reproducible. There are empirical correlations between hardness measures and tensile strengths (ISO 1989): thus

> the Vickers hardness is approximately 3 times the tensile strength.

A theoretical derivation of this factor 3 is given at the end of Chap. 3 (*cf.* Fig. 3.141), by considering the flat indentation of a semi-infinite body.

Micro-hardness tests provide results concerning local properties, such as the depth of penetration of a surface treatment or the extent of a heat affected zone in welding.

Equipments have been developed for *nano-hardness measurements*, which require instrumentation of the indenter (ISO 2002a). In this way properties of individual grains can be investigated (see Sect. 1.2.2.5). In nano-indentation small loads and tip sizes are used. The resulting indentation area is of the order of a few micrometres or nanometres. An indenter with a geometry known to high precision (usually a Berkovich tip which has three-sided pyramid geometry) is used. During the course of the instrumented nano-indentation test, the depth of penetration is recorded, and then the area of the indent is determined using the known geometry of the indentation tip, as shown in Fig. 1.59. These curves can be used to determine the mechanical properties of the material at the nano-scale, in particular the "local" Young modulus can be measured from the slope of the curve, $\mathrm{d}P/\mathrm{d}h$, upon unloading.

An example of the use of nano-indentation technique applied to a dual-phase (ferrite + martensite) steel is given in Fig. 1.60. In this material with an extremely fine microstructure, the sizes of the nano-indents were measured using SEM. The results shown in Fig. 1.60b clearly indicate the difference in hardness between the ferrite phase and the martensite, as expected. Rather surprisingly it is observed that the ferrite / martensite interface is softer. This result can be explained from a detailed analysis of the phase transformation taking place in this type of material.

[14] Johan August Brinell (1849–1925) was a Swedish engineer who proposed the hardness test which bears his name in1900.

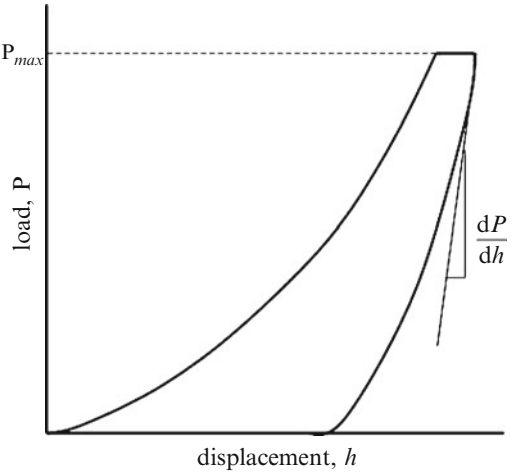


Fig. 1.59 Principle of the nano-indenter. Load–displacement curve

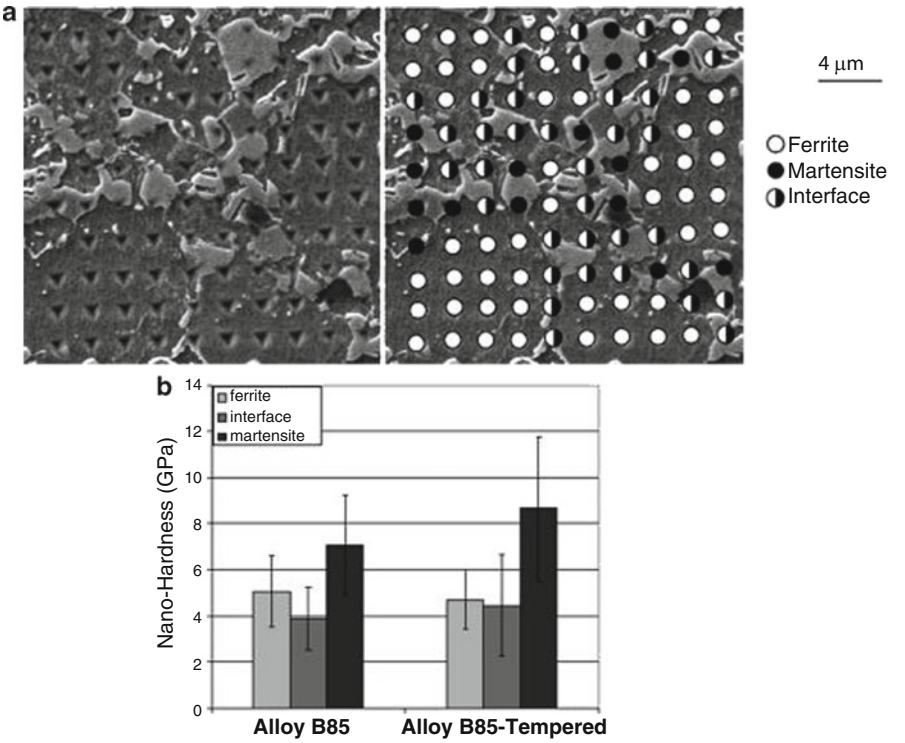


Fig. 1.60 Dual phase (Ferrite + Martensite) steel-Alloy B85. (**a**) Micro-indentations (maximum load: 2 mN, pitch: 2 μm) observed by SEM; (**b**) Nano-hardness in ferrite and martensite phases and at the interface in B85 steel and in tempered B85 steel (by courtesy of A. Dalloz)

1.3.5.4 Fracture Mechanics Tests

Based on the theory of linear elastic fracture mechanics (LEFM), fracture mechanics tests allow the determination of the fracture toughness of materials (ESIS 1992, ISO 2002b, 2005e). This is the fracture energy for crack propagation and is usually expressed in term of the critical stress intensity factor K_{Ic} in MPa$\sqrt{\mathrm{m}}$, which is related to it by a simple formula.

The test is performed on precracked test pieces. Precracking is usually achieved by fatigue. The fracture toughness thus measured is supposed to be size independent. However statistical theory based on weakest link concept has shown that there is a relation between the fracture toughness and the specimen thickness, B, such as ${K_{Ic}}^4 B = \text{constant}$ (see Volume II). This relation holds in the lower shelf and in the transition range of ferritic steels (see also ASTM 2011). Fracture toughness measurements are very useful for the determination of crack propagation in structures. The drawback of LEFM tests is the need, for validity, to avoid excessive plastic deformation at the crack tip. Thus for low yield strength materials extremely large test pieces should be used.

Elastoplastic fracture mechanics (EPFM) theory, not quite as well founded as LEFM, allowed nevertheless the development of similar tests, which differ essentially from the LEFM ones by the way to analyse the data. All this will be detailed in Sect. 2 of Volume II.

1.3.5.5 Impact Tests

The embrittling effect of the increase in strain rate led to the development of various impact tests (Frund 2008; François and Pineau 2002). The *Charpy*[15] *test* is the most common. This is an impact bending test (ISO 2009c) performed with a pendulum. The standard test piece is a square-section bar of length 55 mm and side 10 mm, with a notch cut at the mid-point of the length. This is a 45° V of depth 2 mm with a radius of 0.25 mm at the apex (Charpy V). Still sometimes a U of 5 mm depth with a radius of 1 mm at the base (Charpy U) is used. In the test the bar is broken by dropping a striker held at the end of a pendulum and the energy absorbed at the impact is measured; the result, expressed as joules per cm^2 of the remaining section, is called the KCV (or KCU) resilience. Applied to structural steels, the method has made it possible to show that in ferritic steels there is a transition from ductile fracture at high temperature (high energy) to brittle fracture at lower temperature (low energy).

The Charpy resilience is used mostly for control. Instrumentation of the pendulum by sticking strain gauges on the striker provides more information than the mere fracture energy. In particular it allows the determination of *dynamic fracture toughness* by testing precracked Charpy specimens (ISO 2000b).

[15] Augustin Georges Albert Charpy (1865–1945) was a French engineer, scientist and entrepreneur. He proposed the pendulum which bears his name in 1901.

1.3.5.6 Fatigue Tests

There are several tests for characterising fatigue, which fall into two main classes: *endurance* tests and *low-cycle fatigue* tests (Lieurade et al. 2008). The aim of the first is to find the stress corresponding to the *endurance limit* (number of cycles, $N > 10^6$); these are performed on smooth test pieces stressed in tension or in either plane or rotational bending and the test machines are usually mechanical (ISO 1975, 2006). The second aims not only to measure the endurance but also to study its constitutive law under cyclic loading with a much smaller number of cycles, from 10^2 to 10^5. The machines are strain servo-controlled and the stressing usually tension-compression (ISO 2003).

Fatigue crack propagation tests are based on LEFM. They require a measuring device to follow the crack advance as the test proceeds (ISO 2002c).

1.3.5.7 Creep Tests

Resistance to creep is most commonly measured by applying with the help of a lever a constant tensile load to a smooth specimen (ISO 2009d). The tests are performed at various temperatures. They must be closely controlled, to within $\pm$ 1°C. For each value of the load the strain rate $\mathrm{d}\varepsilon/\mathrm{d}t$ and the time t_R to fracture are measured. The first is used to give the constitutive law of the material, the second its resistance to creep fracture.

1.3.6 Measurements of Elastic Characteristics

1.3.6.1 Static Methods

Static methods simply use the previously described tests: tensile test, bend test and torsion test (Gadaud 2008). The deformation of the test piece must be measured with the best possible precision, for instance with strain gauges or with an extensometer. In each case there is a relation between this deformation and the applied load, which provides an elastic modulus: Young modulus in the case of the tensile and bend tests, shear modulus in the case of the torsion test. The Poisson ratio can also be determined in a tensile test by measuring the lateral contraction of the test piece.

Such measurements can be performed at temperatures different from ambient, subject to some adaptations of the deformation measuring devices.

1.3.6.2 Ultrasonic Methods

The propagation of plane sine waves is discussed in Sect. 2.3.5 of Volume I. Ultrasonic waves are but a particular case.

An ultrasonic wave propagates in an isotropic material with a longitudinal velocity $V_L = \left[\frac{1-\nu}{(1+\nu)(1-2\nu)}\frac{E}{\rho}\right]^{1/2}$; with a transversal wave velocity $V_T = \left(\frac{\mu}{\rho}\right)^{1/2}$, where E is the Young modulus, ν the Poisson ratio, μ the shear modulus and ρ the density. The measurement of these two velocities allows the determination of the elastic constants, two of which only are independent, provided the density is known.

A piezoelectric transducer is applied on the face of a plate of thickness e. It delivers an impulse. Depending on the orientation of the transducer the wave is longitudinal or transversal. Another transducer can be applied on the opposite face, acting as a receiver, or, more simply, the first transducer can act as well as a receiver. In any case the travel time of the wave impulse across the plate is measured, which provides the wave velocity.

Measurements at temperatures less than about 400°C are possible by the use of cements to fix the transducers. At higher temperatures, they must be kept outside the furnace, the wave being introduced in the specimen by a wave-guide.

For anisotropic materials there are more than two wave velocities. A convenient method of measurement consists in immerging in water the specimen held by a rotating support. The ultrasonic wave is then introduced in the specimen without contact. By refraction, it decomposes into three waves, one longitudinal and two transversal. The measurement of the travelling time from an emitter to a receiver provides the wave velocities in the sample, according to its orientation, by comparison with the travelling time in the absence of the specimen. A mathematical treatment makes it possible to go back to the different elastic constants.

1.3.6.3 Resonance Methods

There exist relations between the resonance frequencies of a specimen and the elastic constants. In these expressions always appears the ratio $(E/\rho)^{1/2}$ and the dependence on the Poisson ratio is geometry dependent.

In the resonance method, the specimen is excited by a device, which can be piezoelectric, electro-magnetic, electrostatic for instance. This last method has the advantage that it does not require a contact and it can be adapted to measurements at temperatures different from ambient. The specimen can be a rod or a disk. It can be a bar excited in bending or in torsion.

1.3.6.4 Instrumented Micro-indentation

The instrumented micro-hardness tests mentioned in Sect. 1.3.5.3 can be used to measure the elastic characteristics. The measurements must be performed during unloading to avoid plastic deformation. Numerical calculations are needed to determine precisely the local deformations and the relation between the penetration and the load to obtain the elastic moduli. The advantage of this method is that it allows measurement of local characteristics at a very small scale.

1.3.7 Tendencies in the Evolution of Methods for Studying Mechanical Properties of Materials

We have already, in Sect. 1.3.1, mentioned one tendency, that towards multi-axial stressing; here we mention some others. One, not indeed particularly recent, is the increasing interest in properties related to *viscosity*: – visco-elasticity of polymers, visco-plasticity of metal alloys at low and at high temperatures. For the first, the quantities of interest are the elastic moduli, real and complex, of the material and the energy dissipation under cyclic loading. Measurement here requires special equipment, which is now becoming available commercially. For the second, the quantities studied are those that characterise thermal activation of irreversible deformation, such as activation volume and energy; they can be found by making sudden changes of speed and/or temperature in the course of tests, for which specially-designed equipment is needed.

The development of instruments for observations at low scales (SEM Sect. 1.2.2.3, TEM Sect. 1.2.2.2) led to the possibility to perform mechanical tests within these microscopes. Commercialy available devices replaced the DIC. In the SEM various techniques are used to measure the strains, following the displacements of a grid inscribed on the test piece is one of the most widely developed. Performing bend tests on nano-tubes can be envisaged.

Another tendency, more recent, derives from the development of methods for calculating the response of structures, notably the finite-element method. It is now possible to design tests to be performed on structural test pieces, such as notched specimens, for which calculations have been made in advance, to study the influence of some parameter such as the maximum principal stress or the trace of the stress tensor in some specific type of damage such as intergranular cavitation or ductile fracture. The same approach enables important questions such as the directional aspects of damage to be investigated. Clearly, for this to be possible, the constitutive laws expressing the behaviour of the material must be known, and these can be established as a result of tests made on samples whose geometry is simple enough for the stresses to be deduced without knowing the constitutive equations. Otherwise, inverse methods must be employed.

1.4 Constitutive Equations: General Considerations

1.4.1 Modelling

So far in this introductory chapter we have discussed the field of Mechanics and Materials, the main classes into which materials can be grouped and the means for measuring and analysing their properties on various scales. We now consider the problem of modelling their mechanical behaviour. Modelling is needed to

establish the relation between the microstructure of a material and its macroscopic properties, with the aims of enabling the response to loading of a proposed construction to be predicted and of suggesting and guiding the development of new materials or improvements to existing ones. The laws provided by modelling must be realistic, practical and soundly based; two main types of approach can be envisaged:

1. phenomenological and inductive, operating in a thermo-dynamical framework and proceeding by experimenting on a macroscopic scale to identify parameters
2. deductive, seeking to deduce macroscopic properties from a description of the microstructure, the elementary mechanisms of deformation and damage and combining them into upscaling treatments

The first approach is generally more useable and more frequently used; we shall seek to explain it by relating the information obtained on the microstructural scale to the various types of behaviour considered in the following pages. The second is of course more complex, less well established and less practical, but it is richer and more soundly based physically. We shall take this up in several contexts, particularly in connexion with heterogeneous materials.

1.4.2 Constitutive Equations: Main Classes

The general equations of physics – conservation of mass, momentum, energy, ... – are not sufficient to determine the stress or displacement fields in a structure; we have to add the physical laws that, for the constituent material, relate the dynamic, geometric and kinematic variables. For example, the equilibrium equations for the stresses –

$$\begin{aligned} &\underline{\underline{\sigma}} = \underline{\underline{\sigma}}^{\mathrm{T}} \\ &\operatorname{div} \underline{\underline{\sigma}} + \underline{f} - \rho \underline{\gamma} = 0 \end{aligned} \tag{1.16}$$

provide only six equations for nine unknowns, and the equations of strain compatibility, $\mathbf{Inc}\,\underline{\underline{\varepsilon}} = 0$ (see Annex 3), only three independent equations for six unknowns. It is only the *constitutive equations* which, by linking stresses and deformations, enable the two sets of equations to be used together, and so, with the boundary conditions, provide enough equations to determine both stresses and displacements.

Deduction of a mechanical behaviour from an experiment poses a problem of principle, since the experiment can only be performed on a structure – even a test specimen is a structure, with its specific geometry, under a given loading – and we cannot deduce the stress field in such a loaded structure from the global forces, as is necessary for deriving the behaviour, unless we already know the constitutive

equations. Classically[16] we can only resolve this difficulty of principle by using very specific geometries for the test specimens and special forms of loading that enable us to deduce the local mechanical variables from global quantities, whatever the behaviour: for example, uniaxial traction of a long bar, torsion and internal pressure in a thin-walled tube (see Sect. 1.3). Under such conditions, in a test defined by the pair of parameters Q (load) and q (deformation), related by energy considerations expressing the work done in the deformation, we can distinguish by the form of the relation between Q and q three main classes of behaviour:

Elastic: this is exhibited by a structure recovering, at least partially, its deformation when the load is removed. If the recovery is complete and there is a bijective relation $f(Q, q) = 0$ we speak of *perfect elasticity.* If this relation is *linear and homogeneous* then the elasticity is *linear*: this is so for a linear spring, for which $Q = Eq$ or $q = JQ$. The law, Hooke's (see note 10) law, can be generalised to cover the case of several pairs of parameters: $Q_i = E_{ij}q_j$, $q_i = J_{ij}Q_j$, where, as in the following, the Einstein convention on repeated indices is used.

Viscous: here the response to the deformation varies with its rate $\dot{q}$; this viscous resistance means that, for given q, Q is an increasing function of $\dot{q}$. Viscosity is *pure* when there is a bijective relation between Q and $\dot{q}$, $g(Q,\dot{q}) = 0$; and if this is linear and homogeneous the viscosity is *linear*. This is the case for the linear dashpot, $Q = \eta\dot{q}$, which is Newton's[17] law, which can be generalised to $Q_i = \eta_{ij}\dot{q}_j$ for several parameters.

Plastic: this is characterised by a permanent deformation remaining when the load is removed, provided the load is great enough, that is, exceeds some threshold. Generally, this threshold rises with plastic flow – the phenomenon of *work-hardening*; if it remains constant the plasticity is said to be *perfect.* If further the deformation before the plastic threshold is negligible then the behaviour is described as *rigid-perfect-plastic*: this is the case of a friction slider with constant threshold, represented by $-k \leq Q \leq k, \dot{q}Q \geq 0$.

These three basic behaviours can combine in many ways, so that one can have, for example, visco-elasticity, elasto-plasticity, visco-plasticity, even elasto-visco-plasticity. The combinations can be analysed qualitatively by assembling, in series or in parallel, the rheological models for the basic constituents (spring, dashpot and friction slider).

[16]New methods based on full field strain measurements (see Sect. 1.2.2.9), numerical computations and inverse resolutions have been developed recently in order to perform a direct identification of the material constitutive equations from the experimental response of structures with arbitrary geometry. Information on these approaches can be found *e.g.* in the book edited by Grédiac and Hild 2011.

[17]Sir Isaac Newton (1643–1727) was an English physicist, mathematician, astronomer, natural philosopher, alchemist, and theologian. His "*Principia*", published in 1687, lays the groundwork for most of classical mechanics.

Damage: this is the progressive alteration of mechanical properties that usually derives from the formation and growth of microcracks and microcavities, which accompanies large deformation or cyclic loading for instance, and which can end with fracture. These phenomena are naturally coupled with those which determine the deformation itself, so that the laws governing damage must in general be coupled with those given above: the total assemblage determines the real behaviour of the material.

1.4.3 General Formulation of the Constitutive Equations

Apart from the case of perfect elasticity and pure viscosity, mechanical behaviour exhibits *hereditary* properties: this means that the response to a load depends not only on the load exerted at that moment but also on its entire previous history. According to the principle of *causality*, or *determinism*, this property can be described either by a functional relation between the dynamical and geometrical variables or by dependence, holding at that instant, on a set of *internal parameters* or *hidden variables* describing the state of the material, whose laws of evolution must be known.

Consider the functional representation as a finite transformation $\underline{T}(M,\tau)$ at point M and time τ with gradient $\nabla \underline{T} = \underline{\underline{F}} = \underline{\underline{U}}.\underline{\underline{S}}$ where $\underline{\underline{S}}$ is symmetric and $\underline{\underline{U}}$ orthogonal. Let $\underline{\underline{\sigma}}(M,t)$ and $\underline{\underline{\Pi}}(M,t)$ be the Cauchy and Piola-Kirchoff II stress tensors respectively (see Annex 3), with $\underline{\underline{\Pi}} = J\,\underline{\underline{F}}^{-1}.\underline{\underline{\sigma}}.(\underline{\underline{F}}^{\mathrm{T}})^{-1}$: the Jacobian $J = \det(\underline{\underline{F}})$ is equal to ρ_0/ρ, the ratio of the densities in the reference and current configurations respectively, according to the law of mass conservation.

The representation necessarily involves certain quantities **A**, mechanical variables and characteristics of the system, which satisfy the principle of tensorial invariance and guarantee independence of the coordinate system. Thus after a change of base defined by the matrix $[\alpha]$ or $[\beta]$ with $[\beta] = [\alpha]^{-1}$ the components of **A** are transformed according to the rule:

$$A'^{mn\ldots p}_{ij\ldots k} = \alpha^{\eta}_{i}\alpha^{\theta}_{j}\ldots\alpha^{\varsigma}_{k}\beta^{m}_{\mu}\beta^{n}_{\nu}\ldots\beta^{p}_{\pi}A^{\mu\nu\ldots\pi}_{\eta\theta\ldots\varsigma} \tag{1.17}$$

According as whether a Eulerian or a Lagrangian description is chosen, the functional dependence can be written, omitting temperature and physico-chemical variables,

$$\begin{aligned}\underline{\underline{\sigma}}(M,t) &= \underline{\underline{f}}\left[M,t,\underline{x}(N,\tau)^{t}_{-\infty}\right]\\ \underline{\underline{\Pi}}(M,t) &= \underline{\underline{\mathcal{F}}}\left[M,t,\underline{x}(N,\tau)^{t}_{-\infty}\right]\end{aligned} \tag{1.18}$$

where $\underline{x}(N,\tau)^{t}_{-\infty}$ represents the set of positions of all the points N of the structure at all instants τ in the interval $]-\infty, t]$. This long-range influence of the points N on M expresses *non-local behaviour*.

An important simplification results from assuming a *principle of local action*, according to which only interactions in the neighbourhood of M need to be taken into account: that is, the behaviour in the neighbourhood of M is the same, whatever the mechanical state at a finite distance. Eq. 1.18 then becomes:

$$\begin{aligned}\underline{\underline{\sigma}}(M,t) &= \underline{\underline{f}}\left[M,t,\nabla\underline{T}(M,\tau)_{-\infty}^{t},\nabla^2\underline{T}(M,\tau)_{-\infty}^{t},\nabla^3\underline{T}(M,\tau)_{-\infty}^{t},...\right]\\ \underline{\underline{\Pi}}(M,t) &= \underline{\underline{\mathcal{F}}}\left[M,t,\nabla\underline{T}(M,\tau)_{-\infty}^{t},\nabla^2\underline{T}(M,\tau)_{-\infty}^{t},\nabla^3\underline{T}(M,\tau)_{-\infty}^{t},...\right]\end{aligned} \tag{1.19}$$

A further simplification is made by considering only the first-order gradient $\nabla\underline{T} = \underline{\underline{F}}$; this is the assumption of *material simplicity,* which we shall take for granted from now on. We now have

$$\begin{aligned}\underline{\underline{\sigma}}(M,t) &= \underline{\underline{f}}\left[M,t,\underline{\underline{F}}(M,\tau)_{-\infty}^{t}\right]\\ \underline{\underline{\Pi}}(M,t) &= \underline{\underline{\mathcal{F}}}\left[M,t,\underline{\underline{F}}(M,\tau)_{-\infty}^{t}\right]\end{aligned} \tag{1.20}$$

This form can be specified by introducing the *principle of objectivity*, which expresses the independence of the constitutive equations of the movement of the observer and therefore of any change of frame of reference – that is, any rigid-body movements. It can be shown that the principle implies on the one hand that the time t cannot appear explicitly in the above relations (1.20) (considering a difference in clock time between two observers) and on the other that there is a very particular dependence of the stresses on the deformations $\underline{\underline{S}}$ and the rotations $\underline{\underline{U}}$, that is

$$\begin{aligned}\underline{\underline{\sigma}}(M,t) &= \underline{\underline{U}}(t).\underline{\underline{f}}\left[M,\underline{\underline{S}}(M,\tau)_{-\infty}^{t}\right].\underline{\underline{U}}^{\mathrm{T}}(t)\\ \underline{\underline{\Pi}}(M,t) &= \underline{\underline{\mathcal{F}}}\left[M,\underline{\underline{S}}(M,\tau)_{-\infty}^{t}\right]\end{aligned} \tag{1.21}$$

The point M can occur as an argument of the functional relations only in the case of heterogeneous materials, that is, materials where the properties vary from point to point. For homogeneous (and simple) materials we have, at every point

$$\begin{aligned}\underline{\underline{\sigma}}(t) &= \underline{\underline{U}}(t).\underline{\underline{f}}\left[\underline{\underline{S}}(\tau)_{-\infty}^{t}\right].\underline{\underline{U}}^{\mathrm{T}}(t)\\ \underline{\underline{\Pi}}(t) &= \underline{\underline{\mathcal{F}}}\left[\underline{\underline{\Delta}}(\tau)_{-\infty}^{t}\right]\end{aligned} \tag{1.22}$$

with $\underline{\underline{\Delta}}$ the Green-Lagrange strain tensor.

For an infinitesimal transformation the two descriptions coincide and both reduce to

$$\underline{\underline{\sigma}}(t) = \underline{\underline{f}}\left[\underline{\underline{\varepsilon}}(\tau)_{-\infty}^{t}\right] \tag{1.23}$$

1.4.4 Anisotropy and Heterogeneity

These are two important features of the properties of real materials, often related; we now look at them in turn, emphasising their interactions.

1.4.4.1 Heterogeneity and Homogenisation

In real materials there is usually, between the level of atomic/molecular structure (crystal lattice, point defects, dislocations, molecular chains, branchings, entanglements, etc.) and that of macroscopic treatments (the elementary volume in continuum mechanics), an intermediate level of microstructural heterogeneity (polycrystalline grains; different phases in metals, alloys and polymers; reinforcing particles; fibres . . .). The transition to the macroscopic level can be made more completely, and more quantitatively, from this intermediate level than from the atomic/molecular level.

What we have to do is to try to replace the real *heterogeneous* material by its *homogeneous equivalent* (HEM), in such a way that in the structures studied, and at the (macroscopic) scale at which they are studied, the stress, strain and other fields are the same in both. For this we must be able to take as elementary volume for the structure a *representative volume element* (RVE) of the material; this, whilst large with respect to the inhomogeneities of the microstructure, must be small enough on the macroscopic scale to be treated as a continuous medium. With d, ℓ and L the characteristic lengths of the inhomogeneities, the RVE and the structure, respectively, we must then have: $d \ll \ell \ll L$. In addition, ℓ must also be small with respect to λ, the length of fluctuation of the macroscopic boundary conditions ($\ell \ll \lambda$) and d must be large enough (say $d_0 \leq d$ where d_0 depends on the material and can range from a fraction of nm to a few microns) for the classical concepts, tools and laws of continuum mechanics to be relevant. In summary we assume the following conditions, which are necessary but generally not sufficient to guarantee the existence of a RVE and of the associated HEM

$$\begin{cases} d_0 \leq d \ll \ell \ll L \\ \ell \ll \lambda \end{cases} \tag{1.24}$$

In what follows we consider only *macro-homogeneous materials* for which such a RVE exists; thus we exclude structures that are heterogeneous on the macroscopic scale, for which there are characteristic size and scale effects. Even so, the choice of the representative volume can vary with the theory, the method used and the behaviour studied.

We shall use lower-case symbols for quantities defined on the microscopic scale and upper case for the macroscopic.

Note: In studying finite transformations it may be necessary to describe and take into account the way the microstructure itself changes with the deformation, for example in textured materials or stress-induced morphological changes.

Methodology. Three stages can be distinguished in all of the methods for treating heterogeneous materials:

Representation (or description): definition of the phases and description of their spatial distribution and mechanical behaviour.

Concentration (or localisation): mechanical modelling, and determining the relations between the local fields and the macroscopic quantities.

Homogenisation: averaging the properties and determining the equivalent homogeneous behaviour.

We now describe these in order.

(a) Representation/description. We must first decide on a level that characterises the heterogeneity and then determine the nature of the parameters that will enable us to describe the corresponding *phases*.

For a particulate composite, we can deal with two phases only, the particles and the matrix; but the particle phase can be decomposed itself into sub-phases if different kinds of particles are present or if the shape and orientation of the particles are taken into account. For a polycrystal, a phase can be defined by all the grains of the RVE which have the same crystal orientation. In addition, each phase can be decomposed into sets of grains with the same shape, size and orientation ...

For a two-phase polycrystalline material we could consider only two phases, each described globally; or we could note the granular and crystallographic nature of each *macroscopic phase* and take as elementary phases the grains of the same chemical composition and the same crystal orientation. In addition we could use the shape and size of the grains, or other features.

Having made this choice we have then to characterise the mechanical behaviour of each phase, considered as a homogeneous continuous medium. Various questions arise, for example how *not* to assimilate the average behaviour of polycrystal grains into that of a single crystal of the same orientation; or how to take account of the particular behaviour of the interfacial zone in a composite material consisting of fibres in a polymer matrix?

Finally, we have to describe the geometry of the microstructure, and apart from certain cases where we can give a deterministic description – periodic media for example – we have to resort to statistics.

Example 1: *crystallographic texture of a polycrystal*.

Let $g = (\psi, \theta, \varphi)$ be the Eulerian angles for the crystal orientation with respect to an external frame; we define the distribution function $f(g)$ for the crystal orientations (the *texture* function) by $\mathrm{d}V/V = f(g)\,\mathrm{d}g$, where $\mathrm{d}V/V$ is the volume fraction of grains with orientation g in an interval of $\mathrm{d}g$ at g. This definition makes no reference to the spatial distribution of the grains – no correlation between position and orientation is given or implied.

Example 2: statistical description of the spatial distribution of the elastic moduli $\mathbf{c}(\underline{x})$ in terms of correlation functions. Let $\mathbf{c}_1 = \mathbf{c}(\underline{x}_1)$ and let $P_1(\mathbf{c}_1)$ be the probability density, so that $P_1(\mathbf{c}_1)\,\mathrm{d}\mathbf{c}_1$ is the probability that $\mathbf{c}(\underline{x}_1)$ lies within an interval $\mathrm{d}\mathbf{c}_1$ at $\mathbf{c}_1$. Similarly $P_2(\mathbf{c}_1, \mathbf{c}_2)\mathrm{d}\mathbf{c}_1\mathrm{d}\mathbf{c}_2$ is the probability that $\mathbf{c}(\underline{x}_1)$ and $\mathbf{c}(\underline{x}_2)$ are both within the corresponding intervals. Note that $P_1 = \int P_2\,\mathrm{d}\mathbf{c}_2$.

We now define the ensemble averages for the functions $\mathrm{f}(\mathbf{c}_1)$, $\mathrm{g}(\mathbf{c}_1, \mathbf{c}_2)$, etc. as follows

$$< \mathrm{f} >= \int \mathrm{f}(\mathbf{c}_1)\, P_1(\mathbf{c}_1)\, \mathrm{d}\mathbf{c}_1, \quad < \mathrm{g} >= \int \mathrm{g}(\mathbf{c}_1, \mathbf{c}_2)\, P_2(\mathbf{c}_1, \mathbf{c}_2)\, \mathrm{d}\mathbf{c}_1\mathrm{d}\mathbf{c}_2, \quad \text{etc.} \tag{1.25}$$

NB: This infinite sequence of correlation functions of successive orders is equivalent to that of probability densities $P_\mathrm{n}(\mathbf{c}_1, \mathbf{c}_2, \ldots \mathbf{c}_\mathrm{n})$ as $n \to \infty$.

We shall see later that whilst a knowledge of a limited number of these correlation functions does not enable us to predict the values of the elastic moduli with certainty, it does allow us to set bounds to their values.

The *ergodic hypothesis*, implied in what follows, relates the ensemble and spatial averages:

$$<\mathbf{c}_1> = (1/V) \int_V \mathbf{c}(\mathrm{x}) dV \tag{1.26}$$

(b) Concentration/localisation. It consists in relating local fields and macroscopic variables: in general, heterogeneity means that $\underline{\underline{\sigma}} \neq \underline{\underline{\Sigma}}$, $\underline{\underline{\varepsilon}} \neq \underline{\underline{E}}$, etc.

Example: internal stresses in elasto-plastic materials.

Level 1: residual stresses $\underline{\underline{\sigma}}_\mathrm{I}$ in a plastically deformed component from which the load has been removed, where there is a field of incompatible plastic deformations (see Annex 3, Eq. 3.3). If $\mathbf{Inc}\underline{\underline{\varepsilon}}^\mathrm{p} \neq 0$, then elastic deformations $\underline{\underline{\varepsilon}}^\mathrm{e}$ must exist such that $\mathbf{Inc}\,\underline{\underline{\varepsilon}} = \mathbf{Inc}\,\underline{\underline{\varepsilon}}^\mathrm{e} + \mathbf{Inc}\,\underline{\underline{\varepsilon}}^\mathrm{p} = 0$, with $\underline{\underline{\varepsilon}}^\mathrm{e}$ and $\underline{\underline{\sigma}}_\mathrm{I}$ linked by the laws of elasticity.

Level 2: internal stresses $\underline{\underline{\sigma}}_\mathrm{II}$ on the scale of the grains, within a RVE, deriving from *inter*granular plastic incompatibilities (*i.e.,* between grains)

Level 3: internal stresses $\underline{\underline{\sigma}}_\mathrm{III}$ varying within the grains, deriving from *intra*granular plastic incompatibilities (dislocations, cell walls, ...)

NB: Within a grain $< \underline{\underline{\sigma}}_\mathrm{III} >= \underline{\underline{\sigma}}_\mathrm{II}$

Within a representative element of volume $< \underline{\underline{\sigma}}_\mathrm{II} >= \underline{\underline{\sigma}}_\mathrm{I}$

Within the structure $< \underline{\underline{\sigma}}_\mathrm{I} >= 0$ in the absence of load.

The concentration stage, which is the one that concerns mechanical modelling proper, is specific to each model. It should lead to a law of localisation or concentration which links $\underline{\underline{\sigma}}$ and $\underline{\underline{\Sigma}}$ or $\underline{\underline{\varepsilon}}$ and $\underline{\underline{E}}$, or their derivatives. For linear behaviour we have

$$\begin{cases} \underline{\underline{\varepsilon}}(\underline{x}) = \mathbf{A}(\underline{x}, \Omega, ...) : \underline{\underline{E}} \\ \underline{\underline{\sigma}}(\underline{x}) = \mathbf{B}(\underline{x}, \Omega, ...) : \underline{\underline{\Sigma}} \end{cases} \tag{1.27}$$

where **A** and **B**, (fourth-order) concentration tensors for strains and stresses respectively, depend on the position and on the parameters Ω that describe the phases being considered.

(c) Homogenisation. This is the stage in which the effective (that is, homogenised) behaviour of the material is determined, and possibly the updating of the initial microstructure; it makes use of the relation between the macroscopic variables and their averaged local counterparts. Here we restrict ourselves to infinitesimal transformations.

Averages: Macroscopic additive quantities such as volume mass, internal energy or specific entropy, dissipation, can be defined as the spatial averages of the corresponding local quantities. The same can only be derived for stress and strain tensors, namely

$$\underline{\underline{\Sigma}} = <\underline{\underline{\sigma}}> = 1/V \int_V \underline{\underline{\sigma}}\, \mathrm{d}V \qquad \underline{\underline{E}} = <\underline{\underline{\sigma}}> = 1/V \int_V \underline{\underline{\varepsilon}}\, \mathrm{d}V \tag{1.28}$$

for restricted *homogeneous* boundary conditions. Thus:

(a) if we impose the condition $\underline{\underline{\sigma}}(\underline{x}).\underline{n} = \underline{\underline{\Sigma}}.\underline{n} \quad \forall \underline{x} \in \partial V$, with $\underline{\underline{\Sigma}}$ constant, we get, since $\operatorname{div} \underline{\underline{\sigma}} = 0$ in V

$$\begin{aligned} \int_V \sigma_{ij}\, \mathrm{d}V &= \int_V \left(\sigma_{ik}\delta_{kj} + \sigma_{ik,k}x_j\right) \mathrm{d}V = \int_V \left(\sigma_{ik}x_j\right)_{,k} \mathrm{d}V = \int_{\partial V} \sigma_{ik}x_j n_k \mathrm{d}S \\ &= \Sigma_{ik} \int_{\partial V} x_j n_k \, \mathrm{d}S = \Sigma_{ik} \int_V x_{j,k} \, \mathrm{d}V = \Sigma_{ik} \int_V \delta_{jk} \, \mathrm{d}V = V \Sigma_{ij} \end{aligned} \tag{1.29}$$

(b) if we impose the displacement $\underline{u} = \underline{\underline{E}}.\underline{x}$ on ∂V we have

$$\begin{aligned} 2\int_V \varepsilon_{ij} \mathrm{d}V &= \int_V (u_{i,j} + u_{j,i})\, \mathrm{d}V = \int_{\partial V} \left(u_i n_j + u_j n_i\right) \mathrm{d}S \\ &= \left(E_{ik} \int_{\partial V} x_k n_j \, \mathrm{d}S + E_{jk} \int_{\partial V} x_k n_i \, \mathrm{d}S\right) \\ &= \left(E_{ik} \int_V \delta_{kj} \mathrm{d}V + E_{jk} \int_V \delta_{ki} \mathrm{d}V\right) \\ &= 2\, V\, E_{ij} \end{aligned} \tag{1.30}$$

Suppose now either a statically admissible (SA) $\underline{\underline{\sigma^*}}(\underline{x})$ stress field or a kinematically admissible (KA) $\underline{\underline{\varepsilon'}}(\underline{x})$ strain field obeys the above homogeneous conditions; we then have

$$V < \underline{\underline{\sigma^*}} : \underline{\underline{\varepsilon'}} > = \int_V \sigma^*_{ij} u'_{i,j} \, \mathrm{d}V = \int_V \left(\sigma^*_{ij} u'_i\right)_{,j} \mathrm{d}V = \int_{\partial V} \sigma^*_{ij} u'_i n_j \, \mathrm{d}S \tag{1.31}$$

whence, if $\underline{\underline{\sigma}}^*(\underline{x}).\underline{n} = \underline{\underline{\Sigma}}.\underline{n} \quad \forall \underline{x} \in \partial V$,

$$< \underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}' >= (1/V)\Sigma_{ij} \int_{\partial V} u'_i n_j \, \mathrm{d}S = (1/V)\Sigma_{ij} \int_V \varepsilon'_{ij} \, \mathrm{d}V = \underline{\underline{\Sigma}} :< \underline{\underline{\varepsilon}}' > \tag{1.32}$$

which implies especially (with $\underline{\underline{\varepsilon}}'(\underline{x})$ uniform) $< \underline{\underline{\sigma}}^* >= \underline{\underline{\Sigma}}$,and if $\underline{u}' = \underline{\underline{E}}.\underline{x}$ on ∂V

$$< \underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}' >= (1/V)E_{ik} \int_{\partial V} \sigma^*_{ij} x_k n_j \, \mathrm{d}S = (1/V)E_{ik} \int_V \left(\sigma^*_{ij} x_k\right)_{,j} \mathrm{d}V = \underline{\underline{E}} :< \underline{\underline{\sigma}}^* > \tag{1.33}$$

which implies especially (with $\underline{\underline{\sigma}}^*(\underline{x})$ uniform) $< \underline{\underline{\varepsilon}}' >= \underline{\underline{E}}$.

So, in all cases

$$< \underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}' >=< \underline{\underline{\sigma}}^* > :< \underline{\underline{\varepsilon}}' > \tag{1.34}$$

which is known as *Hill's*[18] *lemma* (Hill 1963). In (1.34), which is independent of the constitutive behaviour, $\underline{\underline{\sigma}}^*(\underline{x})$ and $\underline{\underline{\varepsilon}}'(\underline{x})$ do not need to be associated with one another; if they are, (1.34) means that the macroscopic work density $\underline{\underline{\Sigma}} : \underline{\underline{E}}$ equals the spatial average of the local (or microscopic) work density $\underline{\underline{\sigma}} : \underline{\underline{\varepsilon}}$. This property is usually referred to as "Hill-Mandel's[19] macro-homogeneity condition".

Effective behaviour. From the solution of the concentration problem, together with these relations between the means, we can determine the effective (*i.e.*, homogenised) behaviour of the equivalent medium; here we consider only *linear behaviour* (or linearised, which we would indicate by writing $\dot{\underline{\underline{\sigma}}}, \dot{\underline{\underline{\varepsilon}}}$, etc. in place of $\underline{\underline{\sigma}}, \underline{\underline{\varepsilon}}$, etc.)

From $\underline{\underline{\sigma}} = \mathbf{B} : \underline{\underline{\Sigma}}$ we deduce

$$< \underline{\underline{\sigma}} >=< \mathbf{B} : \underline{\underline{\Sigma}} >=< \mathbf{B} >: \underline{\underline{\Sigma}} = \underline{\underline{\Sigma}} \text{ so that } < \mathbf{B} >= I \tag{1.35}$$

and from $\underline{\underline{\varepsilon}} = \mathbf{A} : \underline{\underline{E}}$

$$< \underline{\underline{\varepsilon}} >=< \mathbf{A} : \underline{\underline{E}} >=< \mathbf{A} >: \underline{\underline{E}} = \underline{\underline{E}} \text{ so that } < \mathbf{A} >= I \tag{1.36}$$

[18] Rodney Hill (1921–) is an English applied mathematician and a former Professor of Mechanics of Solids at the University of Cambridge (UK). He was the founding editor of the Journal of the Mechanics and Physics of Solids, was elected a Fellow of the Royal Society in 1961 and won the Royal Medal in 1993 for his contribution to the theoretical mechanics of soil and the plasticity of solids.

[19] Jean Mandel (1907–1982) was a French mechanic and a Professor of Mechanics at Ecole Polytechnique and Ecole des Mines de Paris who made major contributions to the theoretical mechanics of solids.

If locally $\underline{\underline{\sigma}} = \underline{\underline{f}}(\underline{\underline{\varepsilon}})$ then

$$\underline{\underline{\Sigma}} =< \underline{\underline{\sigma}} >=< \underline{\underline{f}}(\underline{\underline{\varepsilon}}) >=< \underline{\underline{f}}(\mathbf{A} : \underline{\underline{E}}) >= \underline{\underline{F}}^{\text{eff}}(\underline{\underline{E}}) \tag{1.37}$$

and from $\underline{\underline{\varepsilon}} = \underline{\underline{g}}(\underline{\underline{\sigma}})$ we get

$$\underline{\underline{E}} =< \underline{\underline{\varepsilon}} >=< \underline{\underline{g}}(\underline{\underline{\sigma}}) >=< \underline{\underline{g}}(\mathbf{B} : \underline{\underline{\Sigma}}) >= \underline{\underline{G}}^{\text{eff}}(\underline{\underline{\Sigma}}) \tag{1.38}$$

Alternatively, the effective behaviour can be defined from energetic considerations, namely

$$< \underline{\underline{\sigma}} : \underline{\underline{\varepsilon}} >= \underline{\underline{E}} : \underline{\underline{F}}^{\text{eff}}(\underline{\underline{E}}) = \underline{\underline{\Sigma}} : \underline{\underline{G}}^{\text{eff}}(\underline{\underline{\Sigma}}) \tag{1.39}$$

which leads to

$$\left\{ \begin{array}{l} \underline{\underline{E}} : \underline{\underline{F}}^{\text{eff}}(\underline{\underline{E}}) =< (\underline{\underline{E}} : \mathbf{A}^{\mathrm{T}}) : f(\mathbf{A} : \underline{\underline{E}}) > \Rightarrow \underline{\underline{F}}^{\text{eff}}(\underline{\underline{E}}) =< \mathbf{A}^{\mathrm{T}} : f(\mathbf{A} : \underline{\underline{E}}) > \\ \underline{\underline{\Sigma}} : \underline{\underline{G}}^{\text{eff}}(\underline{\underline{\Sigma}}) =< (\underline{\underline{\Sigma}} : \mathbf{B}^{\mathrm{T}}) : g(\mathbf{B} : \underline{\underline{\Sigma}}) > \Rightarrow \underline{\underline{G}}^{\text{eff}}(\underline{\underline{E}}) =< \mathbf{B}^{\mathrm{T}} : g(\mathbf{B} : \underline{\underline{E}}) > \end{array} \right. \tag{1.40}$$

This definition coincides with (1.37) and (1.38): by Hill's lemma, we have

$$\begin{aligned} &< \mathbf{A}^{\mathrm{T}} : \underline{\underline{f}}(\mathbf{A} : \underline{\underline{E}}) >=< \mathbf{A}^{\mathrm{T}} >:< \underline{\underline{f}}(\mathbf{A} : \underline{\underline{E}}) >=< \underline{\underline{f}}(\mathbf{A} : \underline{\underline{E}}) > \\ &< \mathbf{B}^{\mathrm{T}} : \underline{\underline{g}}(\mathbf{B} : \underline{\underline{\Sigma}}) >=< \mathbf{B}^{\mathrm{T}} >:< \underline{\underline{g}}(\mathbf{B} : \underline{\underline{\Sigma}}) >=< \underline{\underline{g}}(\mathbf{B} : \underline{\underline{\Sigma}}) > \end{aligned} \tag{1.41}$$

since $< \mathbf{A}^{\mathrm{T}} >=< \mathbf{B}^{\mathrm{T}} >= \mathbf{I}$ from (1.35) and (1.36) and since $\varepsilon'^{(ij)}_{kl} = A^{\mathrm{T}}_{(ij)kl}$ and $\sigma^{*(ij)}_{kl} = B^{\mathrm{T}}_{(ij)kl}$ can be considered as kinematically and statically admissible trial fields, respectively, for any fixed value of indices (ij).

1.4.4.2 Anisotropy

Mechanical behaviour is generally *anisotropic*, meaning that the response to an applied load usually depends on the direction in which this is applied; this is a consequence of the anisotropic constitution of the material, which in turn is closely related to the heterogeneity of its microstructure. However, if the microstructure has certain symmetries so too will the anisotropy, and the constitutive equations will reflect this.

It is not possible to determine the physical origin of anisotropic behaviour in materials without studying the heterogeneity of their microstructure. If this structure is decomposed into homogeneous elements, the macroscopic anisotropy results from the combination of two forms: an *elemental anisotropy*, that of these

homogeneous elements, and an *organisational anisotropy*, resulting from their spatial arrangement.

Elemental anisotropy is determined by the anisotropy of the deformation mechanisms, which in turn is related to that of the internal structure of the elements, but is not fully determined by this: thus the elasticity of tungsten (BCC) and of magnesium (HCP) crystals is practically isotropic, because of their chemical bonds and their cohesion energies.

Elasticity. The anisotropy of interatomic bonds, depending on the atomic, molecular, electronic ... structures is expressed in that of the elastic moduli and compliances C_{ijkl}, S_{ijkl}; thus the value of Young's modulus for a crystal depends on the orientation of the tensile axis with respect to the crystal lattice.

Viscosity. The anisotropy of reversible movements of impurities (*e.g.* of carbon atoms in steel), of stacking faults in crystals, of relative slips of polymer chains, etc. are expressed in that of the viscosities η_{ijkl}.

Plasticity. Anisotropy arises from the fact that crystallographic slip can occur only on certain planes and/or in certain directions, both determined crystallographically; anisotropy of work-hardening arises from that of interactions between families of defects or between slip systems, of twinning or diffusion.

Organisational anisotropy results from heterogeneity in the microstructure, of which various kinds can be distinguished:

Morphological: the influence of the shapes of crystalline grains, of inclusions, of cavities, of phase domains, etc.; isotropy corresponds to spherical shape and the departure from this determines one contribution to the anisotropy.

Orientation (determined by the extent to which the elements themselves show anisotropic behaviour or have non-spherical shape): non-isotropic distribution of crystallographic orientations (textures); preferential orientations of inclusions, pores, micro-cracks, etc.

Distribution: related to the spatial distribution of the microstructural elements (bands of one phase within another, lamellar structures, oriented composites, alignment or segregation of inclusions, etc.)

To these should be added specific deformation mechanisms, related to the way in which the elements are organised with respect to each other: grain boundary sliding, decohesion at interfaces, delamination of composites, ...

Notes:

(i) The study of anisotropy must consider not only the initial state but also the way this changes with the deformation; this *induced anisotropy* is determined mainly by the changes to the organisation of the microstructure and the anisotropy resulting from these: for example, texture resulting from metal forming.

(ii) The anisotropy observed results from the combination of the two sources just described. Thus the overall behaviour of an *isotropic* aggregate of *anisotropic*

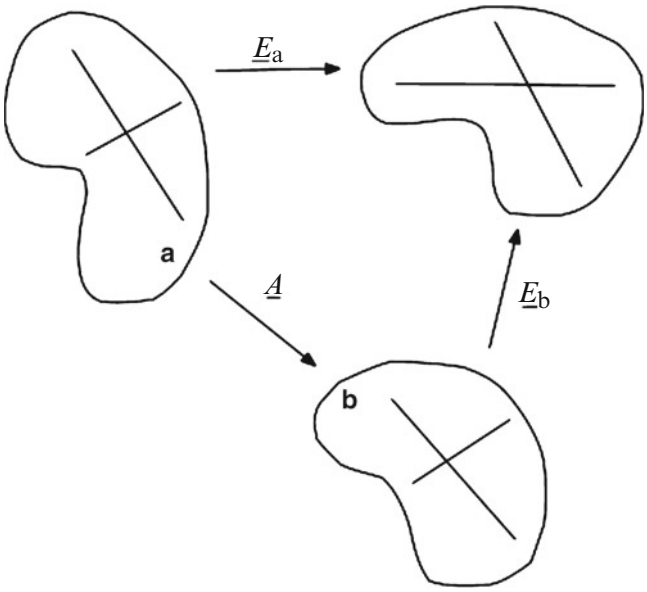


Fig. 1.61 Change of frame of reference

crystals will be *isotropic*, whilst an *anisotropic* organisation of heterogeneous components, which individually are *isotropic,* will behave *anisotropically*.

(iii) Since the anisotropy is expressing the dependence of properties on orientation, rotations play an especially important role.

(iv) Restriction to isotropic loads can mask anisotropic behaviour: thus application of a spherically symmetrical load (hydrostatic pressure) will not reveal the anisotropy in the elasticity of a cubic crystal, or even of an aggregate of such crystals.

The homogenisation techniques make it possible to specify how certain aspects of the anisotropy of micro-heterogeneous materials should be modelled (see later). Nevertheless, it is possible to lay down in advance the conditions that must be fulfilled by a set of anisotropic constitutive equations so that they are consistent with the symmetries of the anisotropy.

For this, consider a body referred to two reference configurations (a), (b) related by the orthogonal tensor $\underline{\underline{A}}$: symbolically (b) $=\underline{\underline{A}}$(a) (Fig. 1.61). The current state is described either by $\underline{\underline{F}}_a$ or by $\underline{\underline{F}}_b$, with $\underline{\underline{F}}_a = \underline{\underline{F}}_b.\underline{\underline{A}}$. The functional relation $\underline{\underline{f}}$ of (1.18) will depend on the particular reference configuration, say $\underline{\underline{f}}_a$ or $\underline{\underline{f}}_b$ with

$$\underline{\underline{f}}_a\left[\underline{\underline{F}}_a\right] = \underline{\underline{f}}_a\left[\underline{\underline{F}}_b.\underline{\underline{A}}\right] = \underline{\underline{f}}_b\left[\underline{\underline{F}}_b\right] \tag{1.42}$$

If $\underline{\underline{A}}$ belongs to the symmetry (or *indistinguishability*) group of the body, the two reference configurations (a), (b) are indistinguishable so far as this piece of material is concerned – for example, rotation through $\pi/2$ about the [001] axis for a cubic crystal. We then have

$$\underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{F}}_{\underline{\underline{a}}}\right] = \underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{F}}_{\underline{\underline{a}}}.\underline{\underline{A}}^{\mathrm{T}}\right] \quad \forall \underline{\underline{F}}_{\underline{\underline{a}}} \tag{1.43}$$

and therefore, since $\mathrm{A}^{\mathrm{T}} = \mathrm{A}^{-1}$ (and also belongs to the same symmetry group)

$$\underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{F}}_{\underline{\underline{a}}}\right] = \underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{F}}_{\underline{\underline{b}}}\right] \tag{1.44}$$

The polar decomposition of $\underline{\underline{F}}_{\underline{\underline{a}}}$ gives (see Annex 3)

$$\underline{\underline{F}}_{\underline{\underline{a}}} = \underline{\underline{U}}_{\underline{\underline{a}}}.\underline{\underline{S}}_{\underline{\underline{a}}} = \underline{\underline{F}}_{\underline{\underline{b}}}.\underline{\underline{A}} = \underline{\underline{U}}_{\underline{\underline{b}}}.\underline{\underline{S}}_{\underline{\underline{b}}}.\underline{\underline{A}} = \underline{\underline{U}}_{\underline{\underline{b}}}.\underline{\underline{A}}.\underline{\underline{A}}^{\mathrm{T}}.\underline{\underline{S}}_{\underline{\underline{b}}}.\underline{\underline{A}} \tag{1.45}$$

whence $\underline{\underline{U}}_{\underline{\underline{a}}} = \underline{\underline{U}}_{\underline{\underline{b}}}.\underline{\underline{A}}, \quad \underline{\underline{S}}_{\underline{\underline{a}}} = \underline{\underline{A}}^{\mathrm{T}}.\underline{\underline{S}}_{\underline{\underline{b}}}.\underline{\underline{A}}$.

Thus, from (1.22) and (1.45), we have

$$\begin{aligned}\underline{\underline{\sigma}} &= \underline{\underline{U}}_{\underline{\underline{a}}}.\underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{S}}_{\underline{\underline{a}}}\right].\underline{\underline{U}}_{\underline{\underline{a}}}{}^{\mathrm{T}} = \underline{\underline{U}}_{\underline{\underline{b}}}.\underline{\underline{A}}.\underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{A}}^{\mathrm{T}}.\underline{\underline{S}}_{\underline{\underline{b}}}.\underline{\underline{A}}\right].\underline{\underline{A}}^{\mathrm{T}}.\underline{\underline{U}}_{\underline{\underline{b}}}{}^{\mathrm{T}} \\ &= \underline{\underline{U}}_{\underline{\underline{b}}}.\underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{S}}_{\underline{\underline{b}}}\right].\underline{\underline{U}}_{\underline{\underline{b}}}{}^{\mathrm{T}}\end{aligned} \tag{1.46}$$

whence $\underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{S}}_{\underline{\underline{b}}}\right] = \underline{\underline{A}}.\underline{\underline{f}}_{\underline{\underline{a}}}\left[\underline{\underline{A}}^{\mathrm{T}}.\underline{\underline{S}}_{\underline{\underline{b}}}.\underline{\underline{A}}\right].\underline{\underline{A}}^{\mathrm{T}}$ and finally the expected condition

$$\underline{\underline{f}}\left[\underline{\underline{A}}.\underline{\underline{S}}.\underline{\underline{A}}^{\mathrm{T}}\right] = \underline{\underline{A}}.\underline{\underline{f}}\left[\underline{\underline{S}}\right].\underline{\underline{A}}^{\mathrm{T}}, \quad \forall \underline{\underline{S}} \tag{1.47}$$

This relation shows how the constitutive functional relation takes account of the symmetry of the body. For an isotropic material it must hold for any orthogonal tensor $\underline{\underline{A}}$ – expressing invariance with respect to rotation.

References

Argon AS, Mc Clintock FA (1966) Mechanical behaviour of materials. Addison Wesley, Reading

Ashby MF (1989) Materials selection in conceptual design. Mater Sci Technol 5:517–525

Ashby MF (1999) Materials selection in mechanical design. Reed Educational and Professional Pub, Oxford

ASTM (2011) Standards test method for determination of reference temperature T_0 for ferritic steels in the transition range. ASTM Int West Conshohocken PA

Behrenat S, Scherp M, Erdmann L, Kahlenhorn W, Moira F, Dereje C, Bleischwitz R, Delzeit R (2007) Rare metals. Measures and concepts for the solution of the problem of conflict aggravating rare material extraction. Report 363 01124 UBA_FB 000980 Federal Environmental Agency Dessau

Berger MH (2011) Private communication

Brooks CR (1982) Heat treatment; structure and properties of non-ferrous alloys. American Society for Metals, Metals Park

Clavel M, Pineau A (1982a) Fatigue behaviour of two Ni-base alloys. Part I: experimental results on low cycle fatigue, fatigue crack propagation and substructures. Mater Sci Eng 55:157–171

Clavel M, Pineau A (1982b) Fatigue behaviour of two Ni-base alloys. Part II: physical modelling of the fatigue crack propagation and substructures. Mater Sci Eng 55:173–180

Cozar R, Pineau A (1973) Morphology of γ' and γ'' precipitates and thermal stability of inconel 718 type alloys. Met Trans 4:47

de la Pena F, Berger MH, Hochepied JF, Dynys F, Stephan O, Walls M (2011) Mapping titanium and tin oxide phases using EELS: An application of independant component analysis. Ultramicroscopy 111:169–176

Egerton RF (1996) Electron energy-loss spectroscopy in the electron microscope. Springer, New York

Egerton RF (2005) Physical principles of electron microscopy. Springer, New York

ESIS (1992) Recommendation for determining the fracture toughness of materials. ESIS procedure P2/92

Flower HM, Gregson PJ (1987) Critical assessment: solid-state phase transformation in aluminum alloys containing lithium. In: Buschow KHJ (ed). Mater Sci Technol 3:81–90

François D (2008) Tensile test. In: François D (ed) Structural components. Wiley, Hoboken, pp 73–88

François D, Pineau A (2002) From Charpy to present impact tests. Elsevier, Amsterdam

Frund J-M (2008) Impact tests. In: François D (ed) Structural components. Wiley, Hoboken, pp 193–205

Gadaud P (2008) Measurements of elastic constants. In: François D (ed) Structural components. Wiley, Hoboken, pp 41–71

Gibson LT, Ashby MF (1988) Cellular solids: structures and properties. Penguin, London

Gourgues-Lorenzon AF (2007) Application of backscatter diffraction to the study of phase transformations. Int Mater Rev 128:52–65

Grediac M, Hild F (2011) Mesures de champs et identification en mécanique des solides. Hermes Science and Lavoisier, Paris

Gregson PJ, Dinsdale K, Harris SJ, Noble B (1987) Evolution of microstructure in Al-Li-Zn-Mg-Cu alloys. In: Buschow KHJ (ed). Mater Sci Technol 3:7–13

Haque MA, Saif MTA (2001) In situ tensile testing of nanoscale specimens in SEM and TEM. Exp Mech 42:123

Hernandez-Castillo LE, Rupin N, Boldetti C, Pinna C, Bornert M (2006) Evaluation of local strain fields in hot worked stainless steel. Photomechanics Clermond-Ferrand, Paris

Hill R (1963) Elastic properties of reinforced solids: some theoretical principles. J Mech Phys Solids 11:357–372

Hutchinson JW, Neale KW (1977) Influence of strain rate sensitivity on necking under uniaxial tension. Acta Met 25:839–846

ISO (1975) Metals rotating bar bending fatigue testing. ISO 1143

ISO (1979) Hard metals compression test. ISO 4506

ISO (1989) Steel conversion of hardness values to tensile strength values. ISO/TR 10108:1989

ISO (1997) Steel and steel products location and preparation of samples and test pieces for mechanical testing. ISO 377:1997

ISO (2000a) Metallic materials tensile testing at low temperature. ISO 15579:2000

ISO (2000b) Steel Charpy V-notch pendulum impact test instrumented test method. ISO 14556:2000

ISO (2002a) Metallic materials instrumented indentation test for hardness and materials parameters Part 1 test method. ISO 14577–1:2002
ISO (2002b) Metallic materials unified method of test for the determination of quasistatic fracture toughness. ISO 12135:2002
ISO (2002c) Metallic materials fatigue testing fatigue crack growth method. ISO 12108:2002
ISO (2003) Metallic materials fatigue testing axial-strain-controlled method. ISO 12106
ISO (2004) Metallic materials tensile testing in liquid helium. ISO 19819
ISO (2005a) Metallic materials Bend test. ISO 7438
ISO (2005b) Metallic materials Brinell hardness test Part 1 Test method. ISO 6506–1
ISO (2005c) Metallic materials Rockwell hardness test Part 1 Test method (scales A B C D E F G H K N T). ISO 6508–1
ISO (2005d) Metallic materials vickers hardness test Part 1 test method. ISO 6507–1
ISO (2005e) Metallic materials determination of plane strain fracture toughness. ISO 12737
ISO (2006) Metallic materials fatigue testing axial force-controlled method. ISO 1099
ISO (2009a) Metallic materials Tensile testing Part 1 method of test at room temperature. ISO 6892–1
ISO (2009b) Metallic materials tensile testing part 2 method of test at elevated temperature. ISO 6892–2
ISO (2009 c) Metallic materials Charpy pendulum impact test Part 1 test method. ISO 148–1
ISO (2009d) Metallic materials uniaxial creep testing in tension Method of test. ISO 204
Lankford WT, Snyder SC, Bausher J (1950) A New criteria for predicting the press performance of deep drawing sheets. Trans ASM 42:1197–1205
Lieurade H-P, Degallaix S, Degallaix G, Gauthier J-P (2008) Fatigue tests. In: François D (ed) Structural components. Wiley, Hoboken, pp 125–189
Miller MK (2000) Atom probe tomography: analysis at the atomic level. Kluwer Academic Plenum Pub, New York
Miller MK (2001) Contribution of atom probe tomography to the understanding of nickel-based superalloys. Micron 32:757–764
Motoyashiki Y, Brückner-Foig A, Sugeta A (2007) Investigation of a small crack behaviour under cyclic loading in a dual phase steel with an FIB tomography technique. Fatigue Fract Eng Mat Struct 30:556–564
Nanga S (2008) Comportement et tranformation martensitique de deux aciers inoxydables austénitiques: effet de la température, de la vitesse et du chargement. Phd thesis, Mines Paristech
Pinna C, Bornert M, Beynon JH, Sellars CM (2000) Experimental investigation and micro-mechanical modelling of the hot deformation of duplex stainless steels. In: Mathematical modeling in metal processing and manufacturing proceedings of the 39th conference
Polak J, Man J, Obrtlik K (2005) Atomic force microscopy study of the early fatigue damage. Mater Sci Forum 482:45–50
Pommier S (2008) Hardness tests. In: François D (ed) Structural components. Wiley, Hoboken, pp 89–121
Reimer I, Kohl H (2008) Transmission electron microscopy. Physics and image formation. Springer, Berlin

Chapter 2
Elastic Behaviour

Abstract Chapter 2 begins with the definition of the elastic potential from the strain energy and discusses the thermodynamic definition. Experimental observations concerning rubber elasticity are described followed by the explanation of models. Discussion of the cohesion energy allows classifying the various types of atomic bonds. Moduli and compliances in linear elasticity are described for anisotropic and then isotropic materials. Stability of the equilibrium is discussed and field equations are described and illustrated by the example of the propagation of plane sine waves. This is followed by the explanation of extremum theorems and an introduction to the finite elements method.

Chapter 2 continues with a discussion of homogenisation (estimating and bounding) for heterogeneous materials in linear elasticity, with explanations about the effective moduli and compliances, initial deformations, thermo-elasticity and the Voigt's and Reuss's bounds. A section is devoted to the problems of inclusions, ellipsoidal essentially.

Lastly, Chapter 2 is concerned with sharper bounds and improved estimates for the elastic moduli and compliances: Mori-Tanaka model and Hashin-Shtrikman bounds, self-consistent scheme. Finally an outline of the theory of elastic random media is given.

The elastic behaviour of materials is certainly the most traditional, the best known and the most widely used in structural design. Here we shall emphasise two aspects of this behaviour, often closely linked, which play an important role in current developments of composite materials: *anisotropy* and *heterogeneity*. After a descriptive introduction we distinguish two main classes of elastic behaviour, and then take up the study of linear elasticity in greater depth, introducing methods for the analysis and modelling of the behaviour of heterogeneous materials.

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1_2, © Springer Science+Business Media B.V. 2012

2.1 Elastic Potential

2.1.1 Strain Energy

In all that follows we shall assume perfect elasticity and infinitesimal strain. The strain tensor $\underline{\underline{\varepsilon}}$ is then a state variable, and determines the stress tensor $\underline{\underline{\sigma}}$ bijectively; the work done in a deformation path from one equilibrium state to another is independent of the path and can play the role of an *elastic potential*.

The work (possibly virtual) associated with the deformation path is defined as that done by the external and inertial forces, and from the fundamental theorem of dynamics we know that it is zero for rigid-body displacements; it can be expressed in two dual forms. In what follows the external and inertial forces are grouped together and denoted by their volume density $\underline{f}$.

Let $\underline{u'}(\underline{x})$ be a virtual kinematically admissible (KA) displacement field suffered by a volume V of a body whose behaviour is at this stage unspecified; $\underline{u'}(\underline{x})$ is assumed continuous and differentiable. The associated virtual deformation work is:

$$W'_{\text{def}} = \int_V \underline{f}.\underline{u'}\,\mathrm{d}V + \int_{\partial V} \underline{T}.\underline{u'}\,\mathrm{d}S \tag{2.1}$$

where $\underline{T} = \underline{\underline{\sigma}}.\underline{n}$ and $\underline{n}$ is the unit outward normal to the surface ∂V of V. Transforming the surface integral to a volume integral, taking into account the equilibrium of the stress field $\underline{\underline{\sigma}}(\underline{x})$, gives:

$$\begin{aligned}
W'_{\text{def}} &= \int_V \underline{f}.\underline{u'}\,\mathrm{d}V + \int_{\partial V} \underline{u'}.\underline{\underline{\sigma}}.\underline{n}\,\mathrm{d}S \left(= \int_V f_i.u'_i\,\mathrm{d}V + \int_{\partial V} (\sigma_{ij}u'_i)n_j\,\mathrm{d}S\right) \\
&= \int_V \left[\underline{f}.\underline{u'} + \mathrm{div}(\underline{\underline{\sigma}}.\underline{u'})\right]\mathrm{d}V \left(= \int_V f_i.u'_i + (\sigma_{ij}u'_i)_{,j}\,\mathrm{d}V\right) \\
&= \int_V \mathrm{Tr}\left[\underline{\underline{\sigma}}.\mathrm{grad}^T\,\underline{u'}\right]\mathrm{d}V \left(= \int_V \sigma_{ij}u'_{i,j}\,\mathrm{d}V\right) \\
&= \int_V \mathrm{Tr}\left[\underline{\underline{\sigma}}.\underline{\underline{\varepsilon'}}\right]\mathrm{d}V = \int_V \underline{\underline{\sigma}} : \underline{\underline{\varepsilon'}}\,\mathrm{d}V \left(= \int_V \sigma_{ij}\varepsilon'_{ij}\,\mathrm{d}V\right)
\end{aligned} \tag{2.2}$$

whence

$$\mathrm{d}W'_{\text{def}} = \underline{\underline{\sigma}} : \underline{\underline{\varepsilon'}}\,\mathrm{d}V \left(= \sigma_{ij}\varepsilon'_{ij}\,\mathrm{d}V\right) \tag{2.3}$$

and we have used the notation $\underline{\underline{a}} : \underline{\underline{b}} = \mathrm{Tr}\,(\underline{\underline{a}}.\underline{\underline{b}})$.

We can consider in the above, instead of the actual quantities $(\underline{\underline{\sigma}}, \underline{T}, \underline{f})$, a statically admissible (SA) stress field $\underline{\underline{\sigma}}^*$ in equilibrium with $\underline{f}^*$ in the body and $\underline{T}^* = \underline{\underline{\sigma}}^*.\underline{n}$ at the surface, and calculate the corresponding deformation work. A similar derivation then gives

$$W^*_{\text{def}} = \int_V \underline{f}^* . \underline{u}' \, \mathrm{d}V + \int_{\partial V} \underline{T}^* . \underline{u}' \, \mathrm{d}S = \int_V \underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}' \, \mathrm{d}V \tag{2.4}$$

whence

$$\mathrm{d}W^*_{\text{def}} = \underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}' \, \mathrm{d}V \left(= \sigma^*_{ij} \, \varepsilon'_{ij} \, \mathrm{d}V \right) \tag{2.5}$$

and the corresponding strain energy density is $\underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}'$. Going back to perfect elasticity, we can use this result for the definition of an elastic potential.

2.1.2 Elastic Potential and the Complementary Energy

We now consider perfectly elastic behaviour, and the change from one equilibrium state $(\underline{\underline{\sigma}}, \underline{\underline{\varepsilon}})$ to a neighbouring one $(\underline{\underline{\sigma}} + \underline{\underline{\mathrm{d}\sigma}}, \underline{\underline{\varepsilon}} + \underline{\underline{\mathrm{d}\varepsilon}})$. We can envisage different volume densities of elastic potential according to the state variable chosen. From (2.5), with $\underline{\underline{\sigma}}^* = \underline{\underline{\sigma}}$ and $\underline{\underline{\varepsilon}}' = \underline{\underline{\mathrm{d}\varepsilon}}$, the volume density $\pi(\underline{\underline{\varepsilon}})$ of elastic potential is defined by

$$\mathrm{d}\pi = \underline{\underline{\sigma}} : \underline{\underline{\mathrm{d}\varepsilon}} \left(= \sigma_{ij} \, \mathrm{d}\varepsilon_{ij} \right) \quad \text{with } \underline{\underline{\sigma}} = \partial\pi / \partial \underline{\underline{\varepsilon}} \tag{2.6}$$

Alternatively we can take $\underline{\underline{\varepsilon}}' = \underline{\underline{\varepsilon}}$ and $\underline{\underline{\sigma}}^* = \underline{\underline{\mathrm{d}\sigma}}$ and define the volume density of *complementary* potential $\varpi(\underline{\underline{\sigma}})$ by:

$$\mathrm{d}\varpi = \underline{\underline{\varepsilon}} : \underline{\underline{\mathrm{d}\sigma}} \left(= \varepsilon_{ij} \, \mathrm{d}\sigma_{ij} \right) \quad \text{with } \underline{\underline{\varepsilon}} = \partial\varpi / \partial \underline{\underline{\sigma}} \tag{2.7}$$

Note that $\mathrm{d}(\pi + \varpi) = \mathrm{d}(\underline{\underline{\sigma}} : \underline{\underline{\varepsilon}})$, so that if we start from a "natural" initial state (*i.e.,* free from stress and strain) we have

$$\pi + \varpi = \underline{\underline{\sigma}} : \underline{\underline{\varepsilon}} \text{ or } \varpi = \underline{\underline{\sigma}} : \underline{\underline{\varepsilon}} - \pi \tag{2.8}$$

Since the differential forms (2.6) and (2.7) can be integrated, the Cauchy conditions yield:

$$\frac{\partial \sigma_{ij}}{\partial \varepsilon_{kl}} = \frac{\partial \sigma_{kl}}{\partial \varepsilon_{ij}} \qquad \frac{\partial \varepsilon_{ij}}{\partial \sigma_{kl}} = \frac{\partial \varepsilon_{kl}}{\partial \sigma_{ij}} \tag{2.9}$$

2.1.3 Thermodynamic Definition

The foregoing definition of the elastic potential is a purely mechanical one and the function so defined cannot have any relation to the thermodynamic functions except under particular thermal conditions – essentially, isothermal or adiabatic. However, a thermodynamically rigorous definition of *perfect thermo-elasticity* at finite strain can be given (Salençon 2001), and the above results put in relation to this.

The thermodynamic framework is defined by two fundamental (conservation) principles, for which we give the local Eulerian expressions. For the configuration considered, ρ is the mass density, e the internal energy per unit mass, r the rate at which heat is input into, or generated in, unit volume and $\underline{q}$ the outward flux of heat. The first principle is expressed by:

$$\rho \dot{e} = \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} + r - \mathrm{div}\underline{q} \tag{2.10}$$

where $\underline{\underline{\sigma}}$ is the Cauchy[1] stress tensor and $\underline{\underline{\dot{\varepsilon}}}$ the Eulerian strain-rate tensor.[2] The net rate at which the finite volume V is gaining heat is then

$$Q' = \int_{\partial V} \underline{q}.\underline{n}\mathrm{d}S + \int_V r\mathrm{d}V \tag{2.11}$$

which is not in general the time-derivative of a function Q.

For the second principle, let s be the entropy per unit mass and T the absolute temperature; then

$$\rho \dot{s} + \mathrm{div}(\underline{q}/T) - r/T \geq 0 \tag{2.12}$$

which, using (2.10), can be written:

$$\underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} + \rho(T\dot{s} - \dot{e}) - (\underline{q}/T).\mathrm{grad}T \geq 0 \tag{2.13}$$

If now $f = e - Ts$ is the free energy per unit mass, (2.13) can be put in the form of the *Clausius*[3]*-Duhem*[4] *inequality* which expresses the fact that the *dissipation* Φ *per unit volume is non-negative*:

$$\Phi = \Phi_1 + \Phi_2 = \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \rho(\dot{f} + s\dot{T}) - (\underline{q}/T).\mathrm{grad}T \geq 0 \tag{2.14}$$

where $\Phi_1 = \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \rho(\dot{f} + s\dot{T})$ is the intrinsic dissipation $\Phi_2 = -(\underline{q}/T).\mathrm{grad}T$ is the thermal dissipation

---

[1] Augustin-Louis Cauchy (1789–1857) was a French mathematician.

[2] The definitions of the stress and strain tensors at finite strain are given in Appendix C and in Volume III.

[3] Rudolf Clausius (1822–1888) was a German physicist.

[4] Pierre Duhem (1861–1916) was a French physicist.

Thermodynamic reversibility is defined by $\Phi_1 = \Phi_2 = 0$; $\Phi_2 = 0$ holds for isothermal and adiabatic changes, and the vanishing of Φ_1 is associated with the definition of perfect elasticity.

The consequences of these definitions can be derived more easily by using a Lagrangian representation, in which the mass density in the reference configuration is ρ_0, the values of e, s, and f remain unchanged since mass is conserved, and the Piola[5]-Kirchhoff[6] stress tensor $\underline{\underline{\Pi}}$ and the Lagrange[7]-Green[8] deformation-rate tensor $\dot{\underline{\underline{\Delta}}}$ are used. The Lagrangian gradient ∇T is written $\nabla T = \mathrm{grad} T$. $\underline{\underline{F}}$ and we have:

$$(1/\rho)\underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}} = (1/\rho)\underline{\underline{\Pi}} : \dot{\underline{\underline{\Delta}}} \tag{2.15}$$

With a suitable definition of $\underline{q}_0$ the Clausius-Duhem inequality is now:

$$\underline{\underline{\Pi}} : \dot{\underline{\underline{\Delta}}} - \rho_0(\dot{f} + s\dot{T}) - (\underline{q}_0/T).\nabla T \geq 0 \tag{2.16}$$

Perfect thermo-elasticity can then be defined by stating that e, s and $\underline{\underline{\Pi}}$ are one-to-one functions of T and $\underline{\underline{\Delta}}$; in this case the inequality (2.16) holds for all values of ∇T and we have

$$\underline{\underline{\Pi}} : \dot{\underline{\underline{\Delta}}} - \rho_0(\dot{f} + s\dot{T}) \geq 0 \tag{2.17}$$

with $\dot{f} = \frac{\partial f}{\partial \underline{\underline{\Delta}}} : \dot{\underline{\underline{\Delta}}} + \frac{\partial f}{\partial T}\dot{T}$, whence

$$\left(\underline{\underline{\Pi}} - \rho_0 \frac{\partial f}{\partial \underline{\underline{\Delta}}}\right) : \dot{\underline{\underline{\Delta}}} - \rho_0 \left(\frac{\partial f}{\partial T} + s\right)\dot{T} \geq 0 \quad \forall \dot{\underline{\underline{\Delta}}}, \forall \dot{T} \tag{2.18}$$

This gives the constitutive equations (after symmetrisation):

$$s = -\frac{\partial f}{\partial T}, \quad \underline{\underline{\Pi}} = \rho_0 \frac{\partial f}{\partial \underline{\underline{\Delta}}} \tag{2.19}$$

It follows that $\Phi_1 = 0$, and the free energy f appears as a thermodynamic potential. It can be shown that this entails the convexity of f, and hence the one-to-one mapping between stress and strain and the reversibility of the deformations in the usual sense.

[5] Gabrio Piola (1794–1850) was an Italian physicist.

[6] Gustav Kirchhoff (1824–1887) was a German physicist.

[7] Joseph-Louis Lagrange (1736–1813) was an Italian born, French mathematician.

[8] George Green (1793–1841) was a British mathematician.

In these conditions we can also define the thermodynamic state by T and $\underline{\underline{\Pi}}$ and use the Legendre[9]-transformed ψ of f, namely:

$$\psi = (1/\rho_0)\,\underline{\underline{\Pi}} : \underline{\underline{\Delta}} - f \tag{2.20}$$

A similar reasoning, starting from $\psi(T, \underline{\underline{\Pi}})$ instead of $f(T, \underline{\underline{\Delta}})$ and from the relation

$$\dot{\psi} = (1/\rho_0)\,\dot{\underline{\underline{\Pi}}} : \underline{\underline{\Delta}} + (1/\rho_0)\,\underline{\underline{\Pi}} : \dot{\underline{\underline{\Delta}}} - \dot{f} = s\dot{T} + (1/\rho_0)\,\dot{\underline{\underline{\Pi}}} : \underline{\underline{\Delta}} \tag{2.21}$$

gives an equivalent expression for perfect thermo-elastic behaviour (after symmetrisation):

$$s = \frac{\partial \psi}{\partial T}, \quad \underline{\underline{\Delta}} = \rho_0 \frac{\partial \psi}{\partial \underline{\underline{\Pi}}} \tag{2.22}$$

where ψ also is a convex function of $\underline{\underline{\Pi}}$. The relations (2.19) and (2.22) lead to the remarkable properties:

$$\frac{\partial \Pi_{ij}}{\partial \Delta_{kl}} = \frac{\partial \Pi_{kl}}{\partial \Delta_{ij}}, \quad \frac{\partial \Delta_{ij}}{\partial \Pi_{kl}} = \frac{\partial \Delta_{kl}}{\partial \Pi_{ij}} \tag{2.23}$$

2.1.4 *Isothermal and Adiabatic Elastic Compliances*

We return now to the case of infinitesimal transformations; here the Eulerian and Lagrangian forms coincide, with ρ denoting the mass per unit volume. The volume density π of elastic potential, defined in (2.6), is related to e and s by

$$\rho\,\dot{e} = \dot{\pi} + \rho\,T\,\dot{s} \tag{2.24}$$

Thus in isothermal conditions

$$\dot{\pi} = \rho(\dot{e} - T_0\dot{s}) = \rho\,\dot{f} = \underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}} \tag{2.25}$$

whence

$$\underline{\underline{\sigma}} = \rho \left.\frac{\partial f}{\partial \underline{\underline{\varepsilon}}}\right|_T = \frac{\partial \pi}{\partial \underline{\underline{\varepsilon}}} \tag{2.26}$$

[9] Adrien-Marie Legendre (1752–1833) was a French mathematician.

In adiabatic conditions π is identified with the internal, not the free, energy since

$$\dot{\pi} = \rho \dot{e} = \underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}} \tag{2.27}$$

with

$$\underline{\underline{\sigma}} = \rho \left. \frac{\partial e}{\partial \underline{\underline{\varepsilon}}} \right|_s = \frac{\partial \pi}{\partial \underline{\underline{\varepsilon}}} \tag{2.28}$$

Thus the elastic constants of a material, which are defined by $\partial \underline{\underline{\varepsilon}} / \partial \underline{\underline{\sigma}}$ (or $\partial \underline{\underline{\sigma}} / \partial \underline{\underline{\varepsilon}}$), can differ according as the conditions are isothermal or adiabatic. This can be made explicit by starting with Eq. 2.21 written for small strains:

$$\mathrm{d}\psi = s\,\mathrm{d}T + (1/\rho)\varepsilon_{ij}\mathrm{d}\sigma_{ij} \tag{2.29}$$

where the thermodynamic state is defined by $\underline{\underline{\sigma}}$ and T; since ψ is a function of the state, the Cauchy conditions are satisfied, and in particular

$$\rho \left. \frac{\partial s}{\partial \sigma_{ij}} \right|_T = \left. \frac{\partial \varepsilon_{ij}}{\partial T} \right|_{\underline{\underline{\sigma}}} \tag{2.30}$$

If the deformations are purely thermal dilatations we have $\varepsilon_{ij}^{\mathrm{th}} = \alpha T \delta_{ij}$, where δ_{ij} is the Kronecker[10] delta and α the coefficient of thermal expansion, we get

$$\left. \frac{\partial \varepsilon_{ij}}{\partial T} \right|_{\underline{\underline{\sigma}}} = \frac{\partial \varepsilon_{ij}^{\mathrm{th}}}{\partial T} = \alpha \delta_{ij} = \rho \left. \frac{\partial s}{\partial \sigma_{ij}} \right|_T \tag{2.31}$$

On the other hand,

$$\mathrm{d}\varepsilon_{ij} = \frac{\partial \varepsilon_{ij}}{\partial T}\mathrm{d}T + \frac{\partial \varepsilon_{ij}}{\partial \sigma_{kl}}\mathrm{d}\sigma_{kl} \tag{2.32}$$

Under adiabatic conditions, $\mathrm{d}s = 0$ and we can use this to derive a relation between $\mathrm{d}\underline{\underline{\sigma}}$ and $\mathrm{d}T$:

$$\mathrm{d}s = \frac{\partial s}{\partial \sigma_{ij}}\mathrm{d}\sigma_{ij} + \frac{\partial s}{\partial T}\mathrm{d}T = 0 \Rightarrow \mathrm{d}T = -\frac{\left(\partial s \,/ \partial \sigma_{ij}\right)\mathrm{d}\sigma_{ij}}{\partial s \,/ \partial T} \tag{2.33}$$

[10] Leopold Kronecker (1823–1891) was a German mathematician.

Using (2.31) and (2.33), we may write (2.32) as

$$\mathrm{d}\varepsilon_{ij} = \left(\left.\frac{\partial \varepsilon_{ij}}{\partial \sigma_{kl}}\right|_T - \alpha \delta_{ij} \frac{\left.\frac{\partial s}{\partial \sigma_{kl}}\right|_T}{\frac{\partial s}{\partial T}} \right) \mathrm{d}\sigma_{kl} = \left(\left.\frac{\partial \varepsilon_{ij}}{\partial \sigma_{kl}}\right|_T - \alpha^2 \frac{\delta_{ij}\delta_{kl}}{\rho \left.\frac{\partial s}{\partial T}\right|_{\underline{\underline{\sigma}}}} \right) \mathrm{d}\sigma_{kl} \tag{2.34}$$

If C_p is the specific heat of the material per unit mass

$$T\,\mathrm{d}s = C_\mathrm{p}\mathrm{d}T \Rightarrow \left.\frac{\partial s}{\partial T}\right|_{\underline{\underline{\sigma}}} = \frac{C_\mathrm{p}}{T} \tag{2.35}$$

and (2.34) becomes

$$\mathrm{d}\varepsilon_{ij} = \left(\left.\frac{\partial \varepsilon_{ij}}{\partial \sigma_{kl}}\right|_T - \frac{\alpha^2 T \delta_{ij}\delta_{kl}}{\rho C_\mathrm{p}} \right) \mathrm{d}\sigma_{kl} \tag{2.36}$$

from which it follows that the isothermal (T constant) and adiabatic (s constant) compliances are related by:

$$\left.\frac{\partial \varepsilon_{ij}}{\partial \sigma_{kl}}\right|_s = \left.\frac{\partial \varepsilon_{ij}}{\partial \sigma_{kl}}\right|_T - \frac{\alpha^2 T \delta_{ij}\delta_{kl}}{\rho C_\mathrm{p}} \tag{2.37}$$

The second term on the right gives the difference between the two. With the values with which we are usually concerned in practice this seldom exceeds a few percent, but although small it has a theoretical importance since it provides a good illustration of the basic thermodynamics of perfect elasticity.

2.2 Two Major Classes of Elastic Behaviour

2.2.1 General Considerations

If we write (2.25) for isothermal conditions

$$\underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}} = \rho(\dot{e} - T_0 \dot{s}) = \rho \dot{f}\Big|_{T=T_0} \tag{2.38}$$

or equivalently

$$\underline{\underline{\sigma}} = \rho \left.\frac{\partial e}{\partial \underline{\underline{\varepsilon}}}\right|_T - \rho T_0 \left.\frac{\partial s}{\partial \underline{\underline{\varepsilon}}}\right|_T \tag{2.39}$$

we see at once that the origin of elasticity is twofold: one source is the variation of internal energy associated with the deformation, the other is that of the entropy, and according to the type of material these can be of very different orders of magnitude.

For crystalline materials the energy term is dominant; here small changes in the inter-atomic distances associated with the deformation bring about large changes in the internal energy, with very little change of entropy. For other materials, in particular elastomers in which chains of macromolecules can be greatly distorted, the entropy term dominates and gives what is called *rubber elasticity*.

The following section is devoted to this second type of behaviour. The first type is usually studied in terms of small strains, as *linear elasticity*. It will be treated afterwards.

2.2.2 Rubber Elasticity

2.2.2.1 Experimental Observations

Rubber elasticity is observed in a certain class of polymeric materials: *elastomeric* materials[11] (see Annex 1); it is characterised by a very low modulus of elasticity, of the order of 1 MPa, and extremely large reversible instantaneous elongations, which can reach 600–1,000%. Coupling between temperature and deformation induces shortening under load on heating, and heat production by elongation. The increase in temperature of natural rubber elongated to five times its original length can reach 5°C. The mechanism of rubber elasticity involves flexible molecular chains. They need to be interconnected so as to prevent gliding, which would induce plasticity. Such reticulation is obtained, for example, by vulcanisation, created by strong sulphur bonds.[12] The number of main chain atoms between two such cross-links is typically several hundreds. Rubber elasticity requires also a temperature high enough to ensure sufficient mobility of the molecular chains. The free energy of interaction between neighbouring molecular chains is independent of their configuration. Thus, the stored elastic energy is simply the sum of individual contributions of the chains.

The entropic nature of this type of elasticity is related to the large number of possible configurations for each molecular chain, which decreases when the material is stretched; that it really is entropic can be ascertained by studying how the stress needed to maintain a given strain varies with temperature.

[11]The word ‘rubber’ is derived from the ability of natural rubber to remove (rub off) marks from paper, which was noted by Joseph Priestley, an English scientist (1733–1804), in 1770.

[12]Vulcanisation was discovered in 1839 by Goodyear and Hyward.

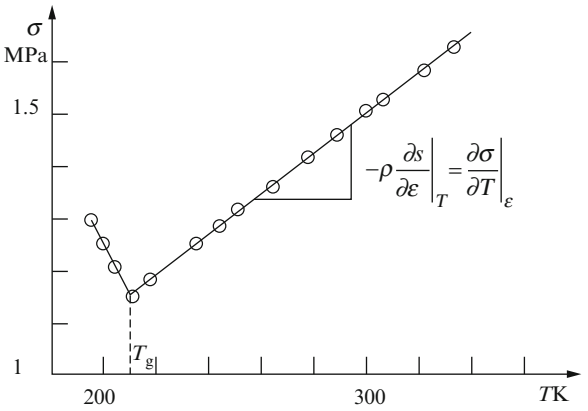


Fig. 2.1 Variation with temperature of the stress needed to maintain an elongation of 350% for an elastomer. T_g is the glass transition temperature (Adapted from Meyer and Ferri 1935)

Using the relation (2.19) for infinitesimal transformations, and the Cauchy conditions, we can write

$$-\rho \left.\frac{\partial s}{\partial \underline{\underline{\varepsilon}}}\right|_T = \left.\frac{\partial \underline{\underline{\sigma}}}{\partial T}\right|_{\underline{\underline{\varepsilon}}} \tag{2.40}$$

Experiments on elastomers show that above the glass transition temperature T_g (see Annex 1) the stress needed to maintain a given elongation varies almost linearly with temperature (Fig. 2.1): this was first observed by Ferry[13] in 1935 (Ferry 1980), for vulcanised natural rubber. For low elongations the slopes of the stress versus temperature plots are found to decrease with the elongation.

The slope of the line is $-\rho(\partial s/\partial \underline{\underline{\varepsilon}})_T$, which is thus found to remain almost constant. Extrapolating the line back to 0 K gives the value of $\rho(\partial e/\partial \underline{\underline{\varepsilon}})_T$; this is found to be very close to zero, showing that the contribution of the internal energy is indeed negligible.

Rubber elasticity can be modelled by counting the changing number of possible configurations for the molecular chains as the elastomer is stretched. It is assumed that the chain segments of the network deform independently and in the microscopic scale in the same way as the whole (affine deformation). The cross-links are assumed to be fixed in space at positions exactly defined by the specimen deformation ratio. This model is referred to as the "affine network model" (Flory and Rehner 1943; Treloar 1943, 1975; Flory 1953, 1985). On the other hand in the "phantom network model" the cross-links are allowed a certain fluctuation about their average position (James and Guth 1947). The affine model yields an upper bound modulus, while the phantom network model yields a lower one. We develop hereafter the affine network model only.

[13] John Douglas Ferry (1912–2002) was professor at the University of Wisconsin.

2.2.2.2 Affine Network Model

In order to obtain a constitutive stress-strain equation, in this model it is assumed that the molecular chains between cross-links can be represented by "phantom chains" having one extremity at the origin of coordinates and whose end to end distance r follows Gaussian statistics (Fig. 2.2). Thus the probability that the extremity of a phantom chain lies at a point of coordinates x_1, x_2, x_3 is:

$$p(x_i) = \frac{\exp\left(-{x_1}^2/2d^2\right)}{(2\pi d^2)^{1/2}} \frac{\exp\left(-{x_2}^2/2d^2\right)}{(2\pi d^2)^{1/2}} \frac{\exp\left(-{x_3}^2/2d^2\right)}{(2\pi d^2)^{1/2}}$$
$$= \frac{1}{(2\pi d^2)^{3/2}} \exp\left(-\frac{r^2}{2d^2}\right) \tag{2.41}$$

d being the standard deviation. Note that:

$$\int_{-\infty}^{+\infty} p(x_i)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = 1 \tag{2.42}$$

The probability that n phantom chains have their extremity at x_i is $[p(x_i)]^n$. If the total number of chains is N then $n = Np(x_i)$. The probability P of any one configuration is the product of the probabilities of the chains ending at the various points x_i, that is

$$P = \prod_i [p(x_i)]^{Np(x_i)} \tag{2.43}$$

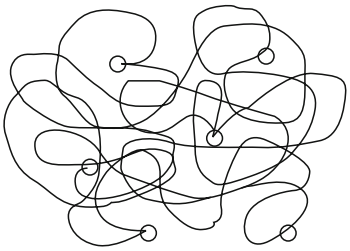
Finally, if the number of configurations of a molecular chain at x_i is C, the total number of configurations is CP and the entropy is

$$S = \mathrm{k}\log CP = \mathrm{k}\log C + \mathrm{k}N\sum_i p(x_i)\log p(x_i) \tag{2.44}$$

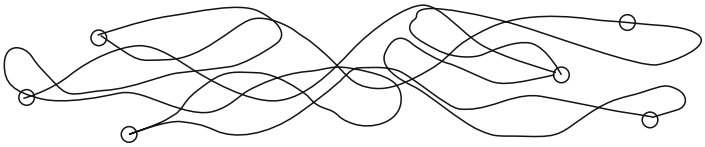
where k is the Boltzmann[14] constant. We can now evaluate the change in free energy $F = Q - TS$ when the material is stretched, that is, when the end-points of the phantom chains are moved from x_{i0} to x_i:

$$\Delta F = -\mathrm{k}T\sum n\log\left[\frac{p(x_i)}{p(x_{i0})}\right] \tag{2.45}$$

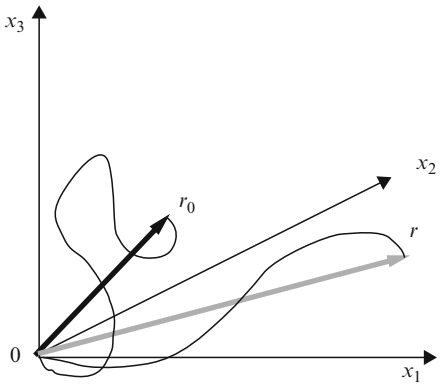
[14]Ludwig Boltzmann (1844–1906) was an Austrian physicist famous for his founding contributions in the fields of statistical mechanics and thermodynamics. The so-called "Boltzmann constant" was actually first introduced by the German physicist Max Planck (1858–1947), the founder of the quantum theory (Nobel Prize in Physics, 1918).



Molecular chain before extension



Molecular chain after extension



Phantom chain

Fig. 2.2 Molecular chains linked by vulcanisation and the phantom chains (affine network model)

Substituting for $p(x_i)$ we find:

$$\Delta F = \mathrm{k}TN \iiint \frac{x_1^2 - x_{10}^2 + x_2^2 - x_{20}^2 + x_3^2 - x_{30}^2}{2d^2} \left(\frac{1}{2\pi d^2}\right)^{3/2} \times \exp\left(-\frac{r_0^2}{2d^2}\right) \mathrm{d}x_{10}\mathrm{d}x_{20}\mathrm{d}x_{30} \tag{2.46}$$

where the integrals are over $\pm\infty$.

The elongation is measured by the ratios $\lambda_I = x_i/x_{i0}$. In terms of these we get:

$$\Delta F = \frac{kTN}{2}\left(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3\right) \tag{2.47}$$

and, for unit volume, N being then the number of chains per unit volume, we can write this as:

$$f_{\mathrm{v}} = \frac{\rho \mathrm{R}T}{2M_c}\left(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - 3\right) \tag{2.48}$$

ρ being the specific mass, M_{c} being the molecular weight and R the gas constant.

The isothermal stress-strain equation follows from the derivative of *the free energy*. For *uniaxial loading* $\lambda_2 = \lambda_3$ as the material is assumed to be isotropic and $\lambda_1 {\lambda_2}^2 = 1$ since the volume $S_0 L_0$ is assumed to remain constant. Hence the total free energy:

$$f_{\mathrm{t}} = S_0 L_0 \frac{\rho \mathrm{R}T}{2M_{\mathrm{c}}}\left(\lambda_1^2 + \frac{2}{\lambda_1}\right) \tag{2.49}$$

We deduce the stress σ_1:

$$\sigma_1 = \frac{1}{S_0}\lambda_1 \frac{\partial f_t}{\partial L} = \frac{\rho \mathrm{R}T}{M_{\mathrm{c}}}\left(\lambda_1^2 - \frac{1}{\lambda_1}\right) \tag{2.50}$$

L being the length of the test piece, while the engineering stress σ_{e1} is:

$$\sigma_{e1} = \frac{1}{S_0}\frac{\partial f_t}{\partial L} = \frac{\rho \mathrm{R}T}{M_c}\left(\lambda_1 - \frac{1}{\lambda_1^2}\right) \tag{2.51}$$

This relation fits well the experimentally observed nonlinear elastic behaviour of elastomers, at least for not too large deformations (Fig. 2.3). When the stretching becomes too great the model ceases to hold, since then some chains are fully extended and only a few behave according to the Gaussian distribution. Also crystallisation can occur under such conditions.

In the case of *pure shear*, if we repeat the calculation from (2.49) onwards, allowing for the increase in free energy, we find that, with $x_1 = x_{10} + \gamma\, x_{20}$, $x_2 = x_{20}$, $x_3 = x_{30}$, f is a function of γ^2 and therefore the shear stress $\tau = \partial f/\partial\sigma$ is proportional to γ. In pure shear elastomers display a linear behaviour.

The *Mooney-Rivlin* display of the results is often used to check the agreement between theory and experiments. It modifies the preceding Eq. 2.51 to read:

$$\sigma_{\mathrm{e1}} = 2\left(\lambda - \frac{1}{\lambda^2}\right)\left(\mathrm{C}_1 + \frac{\mathrm{C}_2}{\lambda}\right) \tag{2.52}$$

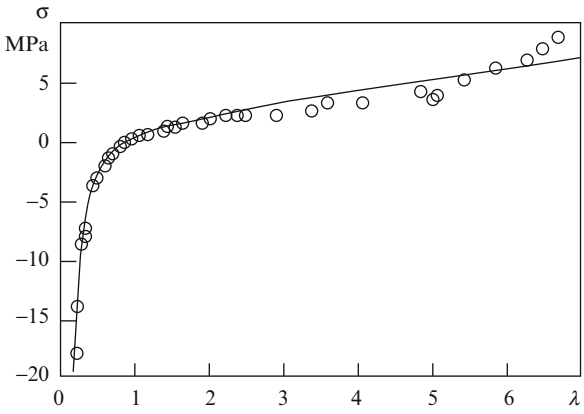


Fig. 2.3 Engineering stress-extension ratio for natural rubber and poly(dimethyl siloxane) (Higgs and Gaylord 1990). The data are scaled to fit the relation (2.51) near $\lambda = 1$

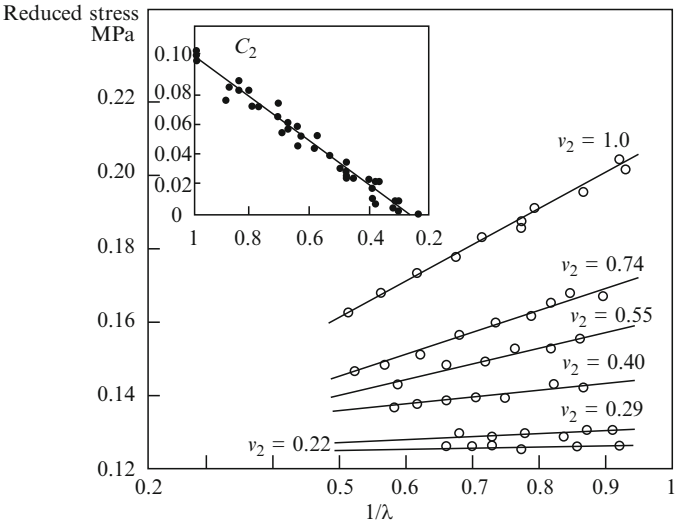


Fig. 2.4 Mooney diagram (reduced stress as a function of the reciprocal of the extension ratio) for natural rubber diluted in benzene. ν_2 is the volume fraction. The parameter C_2 of the Mooney equation decreases when the dilution increases (Grumbrell et al. 1953)

Data are then plotted as the reduced stress $\sigma_{e1} \big/ 2\left(\lambda - 1 \big/ \lambda^2\right)$ versus $1/\lambda$. The slope C_2 should be equal to zero if the affine network model is obeyed. The deviation is thought to be due to entanglements, which reduce the molecular chains mobility. This is in keeping with the decrease of the parameter C_2 when the polymer is diluted, providing more space between chains (Fig. 2.4).

Lastly the affine network model accounts well for the effect of molecular weight on the elastic modulus.

It is furthermore remarkable that, contrary to most materials, this modulus of rubber increases with temperature.

2.2.2.3 Phantom Network Model

As stated before this model, due to James and Guth (1947), considers that the position of the ends of the phantom molecular chains can fluctuate. As there is then a coupling between the movements of the various chains the calculation is involved and it will not be given here. The model predicts an elastic modulus which is half that of the affine network model. Experimental results fall in between the two. They are closer to the phantom network model results for large extensions.

2.2.3 Cohesion Energy, Elastic Constants

Now we consider crystalline materials whose elastic properties depend mainly on the interatomic binding energy and its sensitivity to small changes in interatomic distances. We therefore neglect the entropy term and consider only the equilibrium between the attractive binding forces and the repulsive forces; applying a mechanical load to the material results in small departures of the interatomic distances from their equilibrium values, with a return to those values when the load is removed.

The binding forces and their variation with distance vary greatly according to the nature of the chemical bond; briefly, the types are:

- ionic (*e.g.* NaCl): electrostatic, with a potential U of the form $-q^2/4\pi\varepsilon_0 r$, q being the charge of the ions, ε_0 the permittivity in empty space (8.85 10^{-12} F/mol) and r the interatomic distance. It is a very strong potential of the order of 600–1,500 kJ/mol.
- covalent (*e.g.* C = C), resulting from the sharing of electrons in the same orbital; also very strong (of the order of 400–800 kJ/mol), and of marked directionality.
- metallic, resulting from the interaction of the metal ions with the free-electron gas (of the order of 100–800 kJ/mol).
- van der Waals,[15] governing intermolecular attraction, as in polymers; varying like r^{-6}, weak, of the order of 7–15 kJ/mol.
- hydrogen, responsible for the dipole moment of the water molecule (of the order of 30–50 kJ/mol); significant in polymers, concrete, etc.

(recall that the thermal energy RT is 8.3 T J/mol)

[15] Johannes-Diderick van der Waals (1837–1923) was a Dutch physicist.

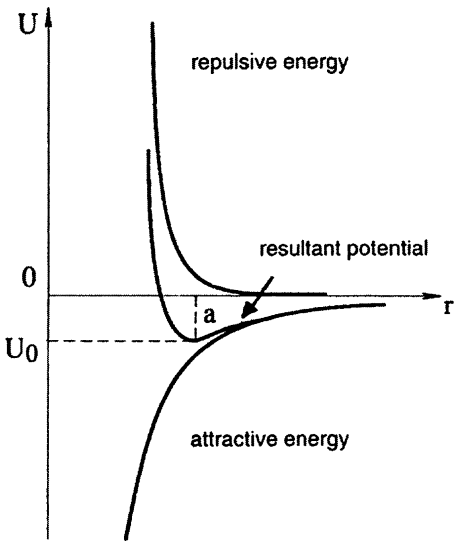


Fig. 2.5 Interatomic potential as a function of distance apart

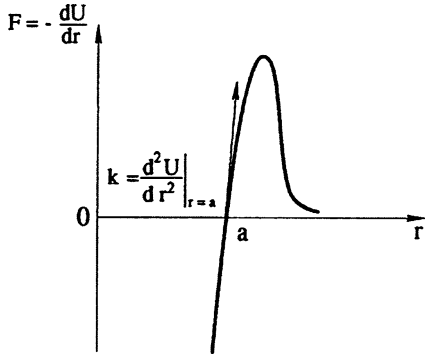


Fig. 2.6 Interatomic force as a function of distance apart

The equilibrium between the attractive and repulsive forces leads to a potential of the form:

$$U = -\frac{\mathrm{A}}{r^n} + \frac{\mathrm{B}}{r^m} \tag{2.53}$$

with $m > n$; differentiating, equilibrium is at $r = a$, where $a = (\mathrm{B}m/An)^{1/(m-n)}$. B is the Born[16] constant (Figs. 2.5 and 2.6).

The binding energy is the value U_0 of U at $r = a$. Being the energy needed to bring atoms from their equilibrium position in solids to far away distances, it corresponds to the energy of sublimation.

[16]Max Born (1882–1970) was a German physicist.

The so-called "elastic constants" are determined by the value k of $\mathrm{d}^2U/\mathrm{d}r^2$ in the neighbourhood of $r = a$. Linear elasticity uses only the value of this at $r = a$; beyond that anharmonic effects would have to be taken into account. In general, the binding energy U_0 and the moduli k vary in the same way from one material to another.

Modern computation tools allow the determination of the elastic constants by "*ab-initio*" calculations of the interatomic potentials. The moduli can vary by several orders of magnitude, as typical values for Young's modulus[17] show:

- 30–300 GPa for ceramics and metals: concrete 45, aluminium 70, steels 200
- 0.001–0.01 GPa for elastomers

Values for Young modulus and Poisson ratio[18] for a range of materials are given in Table 2.1.

Figure 2.7 shows the variation of Young modulus with temperature. This general regular decrease when the temperature is raised suffers anomalous behaviours owing to magnetic effects for some metals such as Fe64Ni36 called Invar, an alloy which has also no thermal expansion over a certain temperature range (Fig. 2.8).

The thermal expansion coefficient is also related to the binding energy, as far as a deep and narrow dip in the binding energy curve as a function of the interatomic distance corresponds to a lower displacement of the equilibrium distance when the temperature is raised than a shallow dip, such as the one found for van der Waals interactions.

2.3 Linear Elasticity

2.3.1 Elastic Moduli and Compliances

We now make this loose concept of "elastic constant" more precise for the case of isothermal elasticity for small deformations and with a natural initial state. Instead of scalar "constants", we need to use fourth-order tensors for the elastic moduli $\mathbf{C}$ and compliances $\mathbf{S}$ respectively, which are inverses of each other

$$\begin{cases} \underline{\underline{\sigma}} = \mathbf{C} : \underline{\underline{\varepsilon}} & \left(\sigma_{ij} = C_{ijkl}\varepsilon_{kl}\right) \\ \underline{\underline{\varepsilon}} = \mathbf{S} : \underline{\underline{\sigma}} & \left(\varepsilon_{ij} = S_{ijkl}\sigma_{kl}\right) \end{cases} \tag{2.54}$$

[17] Thomas Young (1773–1829) was an English physician and scientist, who mastered eight languages.

[18] Siméon Poisson (1781–1840) was a French mathematician.

Table 2.1 Bond dependent properties of some materials (the values coming from various sources are sometimes rather different and depend on the actual state of the material. If precise data is needed resort to bibliography or to experiments)

Material	Young modulus (MPa)	Poisson ratio	Coefficient of thermal expansion $K^{-1} \times 10^6$
Covalent	–	–	–
Carbon (C)			
Diamond	827,000	0.25	1.1
Graphite	30,000		7.9
Fibre	640,000		
Nanotube	1,100,000		
Silicon (Si)	188,000	0.38	2.6
Silicon carbide (SiC)	490,000	0.37	Cubic 3.6 Hexagonal 3.7 and 4.9
Titanium carbide (TiC)	440,000	–	–
Tungsten carbide (WC)	650,000	–	–
–			
Oxides	–	–	–
Alumina (Al_2O_3)			
Single crystal	441,000		
Polycrystal	378,000		5.4
Whisker	490,000		
Silica (SiO_2)	107,000	–	–
Magnesia (MgO)	250,000	–	–
Zirconia (ZrO_2)	205,000	–	–
Marble	26,000	–	5.5–14.1
Limestone	20,000–70,000		8
Granite	60,000		7.9
Glass	69,000	0.18–0.3	8.5
Glass E	72,000		
Glass S	87,000		
Glass R	86,000		
(see foot note)			
Concrete	20,000–50,000	0.20	13
–			
Metals	–	–	–
Beryllium (Be)	290,000	0.03	11.3
Magnesium (Mg)	40,000	0.29	8.2
Aluminum (Al)	69,000	0.34	23.1
2024	75,000		
Titanium (Ti)	118,000	0.32	8.6
Zirconium (Zr)	96,000	0.34	5.7
Tantalum (Ta)	189,000	0.35	6.3
Chromium (Cr)	252,000	0.21	4.9
Molybdenum (Mo)	330,000	0.31	4.8
Tungsten (W)	350,000	0.17	4.5

(continued)

Table 2.1 (continued)

Material	Young modulus (MPa)	Poisson ratio	Coefficient of thermal expansion $K^{-1} \times 10^6$
Iron (Fe)	200,000	0.28–0.30	11.8
Cast iron	83,000–170,000	0.21–0.26	10.8
Nickel (Ni)	225,000	0.31	13.4
Copper (Cu)	112,000	0.33	13.0
Silver (Ag)	72,000	0.37	18.9
Gold (Au)	80,000	0.44	14.2
Zinc (Zn)	90,000	0.25	30.2
Cadmium (Cd)	69,000	0.29	30.8
Lead (Pb)	18,000	0.44	28.9
Tin (Sn)	50,000	0.36	22.0
Uranium (U)	168,000	0.25	13.9
–			
Polymers	–	–	–
Polyethylene	200–700	–	200
Polyvinylchloride (PVC)	2,400	–	50.4
Epoxy	2,400	–	–
Polyester	5,000	–	–
Polycarbonate	2,300	–	–
Polystyrene	3,000–3,400	–	–
Plexiglas (methylmetacrylate)	2,450–3,500	–	–
Rubber	700–4,000	0.5	77
–			
Wood, organic and biomaterials	–	–	–
Mahogany	12,000	–	–
Bamboo	20,000	–	–
Oak	12,000	–	Parallel to cells 4.9 Perpendicular to cells 5.4
Spruce	13,000	–	–
Maple	10,000	–	–
Ash tree	10,000	–	–
Cork		0	–
Redwood	9,500	–	–
Paper	3,000–4,000	–	–
Femur	17,200	–	–
Vertebra	230	–	–
Hair	10,000	–	–
Spider silk	60,000	–	–

Fiberglass E (SiO_2 53–56%, Al_2O_3 12–16%, CaO and MgO 21–24%, B_2O_3 5–9%): general use; Fiberglass S (SiO_2 64–65%, Al_2O_3 24–25%, CaO and MgO 10–11%): high resistance; Fiberglass R (SiO_2 58–60%, Al_2O_3 23.5–25.5%, CaO and MgO 14–17%): high resistance (aeronautics)

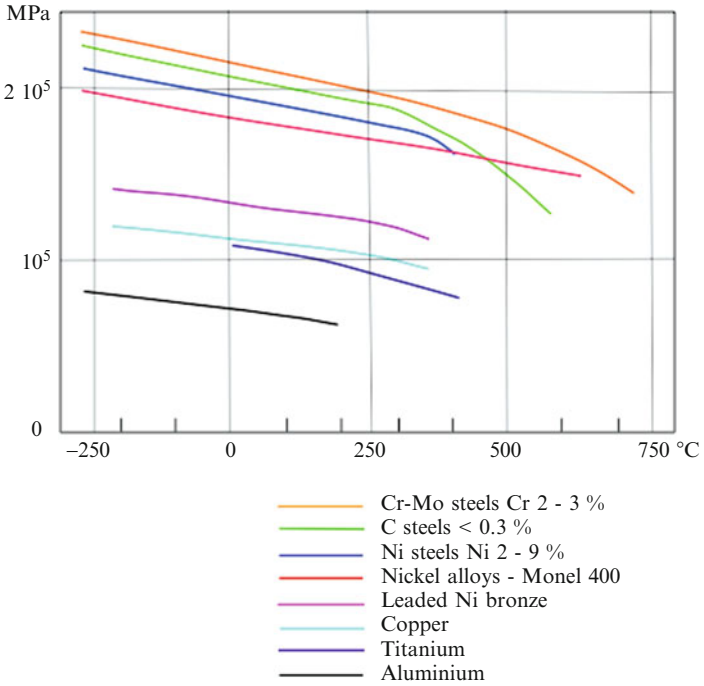


Fig. 2.7 Young modulus as a function of temperature for a number of metals

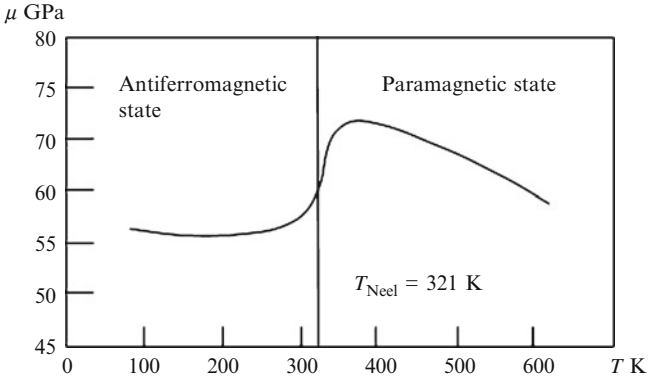


Fig. 2.8 Anomalous behaviour of the variation of Young modulus with temperature for Invar (By courtesy of Allain)

In addition to the obvious symmetry relations

$$C_{ijkl} = C_{jikl} = C_{ijlk}, S_{ijkl} = S_{jikl} = S_{ijlk} \tag{2.55}$$

they obey the "thermodynamical" relations (diagonal symmetry)

$$C_{ijkl} = C_{klij}, S_{ijkl} = S_{klij} \tag{2.56}$$

which follow from the Cauchy relations (2.23); it follows that **C** and **S** each have at most 21 independent components.

Two important properties follow from this:

(a) the quadratic form $\pi = (1/2)\underline{\underline{\varepsilon}} : \mathbf{C} : \underline{\underline{\varepsilon}} = (1/2)\underline{\underline{\sigma}} : \mathbf{S} : \underline{\underline{\sigma}} = \varpi$ is positive definite (positive $\forall \underline{\underline{\varepsilon}}, \forall \underline{\underline{\sigma}}$ and zero only if $\underline{\underline{\varepsilon}} = 0$ or $\underline{\underline{\sigma}} = 0$).
(b) owing to (2.56), there is a reciprocity between two states of load (1) and (2):

$$\mathrm{Tr}\left[\underline{\underline{\sigma}}^{(1)}.\underline{\underline{\varepsilon}}^{(2)}\right] = \mathrm{Tr}\left[\underline{\underline{\sigma}}^{(2)}.\underline{\underline{\varepsilon}}^{(1)}\right] \tag{2.57}$$

Note: *Hyper- and hypo-elasticity*.

The extension to finite strains can be made in different ways and in particular by distinguishing between hyper-elasticity ($\underline{\underline{\sigma}} = \partial\pi/\partial\underline{\underline{\varepsilon}}$) and hypo-elasticity ($\dot{\underline{\underline{\sigma}}} = \mathbf{C} : \dot{\underline{\underline{\varepsilon}}}$). Where small deformations are concerned, these definitions are equivalent and they coincide with (2.54).

2.3.2 Anisotropy

2.3.2.1 Voigt[19] Notation

We can use a matrix notation due to Voigt which takes account of the symmetry relations (2.45) and (2.46); this uses two indices I, J running from 1 to 6 according to the rule

(i,j):	1,1	2,2	3,3	2,3	3,1	1,2
I:	1	2	3	4	5	6

The notation C_{IJ} takes account of the symmetry of C_{ijkl} with respect to i,j on the one hand and to k,l on the other. In addition, (2.56) implies that $C_{IJ} = C_{JI}$, so that the 6×6 matrix C_{IJ} has at most 21 independent components; and similarly for S_{IJ}. Writing $\gamma_{12} = 2\varepsilon_{12}, \gamma_{23} = 2\varepsilon_{23}, \gamma_{31} = 2\varepsilon_{31}$ we have the following equations for $\underline{\underline{\sigma}}$ in terms of $\underline{\underline{\varepsilon}}$:

[19] Woldemar Voigt (1850–1919) was a German physicist.

$$\begin{cases} \sigma_{11} = C_{11}\varepsilon_{11} + C_{12}\varepsilon_{22} + C_{13}\varepsilon_{33} + C_{14}\gamma_{23} + C_{15}\gamma_{31} + C_{16}\gamma_{12} \\ \sigma_{22} = C_{21}\varepsilon_{11} + C_{22}\varepsilon_{22} + C_{23}\varepsilon_{33} + C_{24}\gamma_{23} + C_{25}\gamma_{31} + C_{26}\gamma_{12} \\ \sigma_{33} = C_{31}\varepsilon_{11} + C_{32}\varepsilon_{22} + C_{33}\varepsilon_{33} + C_{34}\gamma_{23} + C_{35}\gamma_{31} + C_{36}\gamma_{12} \\ \sigma_{23} = C_{41}\varepsilon_{11} + C_{42}\varepsilon_{22} + C_{43}\varepsilon_{33} + C_{44}\gamma_{23} + C_{45}\gamma_{31} + C_{46}\gamma_{12} \\ \sigma_{31} = C_{51}\varepsilon_{11} + C_{52}\varepsilon_{22} + C_{53}\varepsilon_{33} + C_{54}\gamma_{23} + C_{55}\gamma_{31} + C_{56}\gamma_{12} \\ \sigma_{12} = C_{61}\varepsilon_{11} + C_{62}\varepsilon_{22} + C_{63}\varepsilon_{33} + C_{64}\gamma_{23} + C_{65}\gamma_{31} + C_{66}\gamma_{12} \end{cases} \tag{2.58}$$

Similarly for $\underline{\underline{\varepsilon}}$ in terms of $\underline{\underline{\sigma}}$, but with new relations between S_{IJ} and S_{ijkl}

$$\begin{cases} \varepsilon_{11} = S_{11}\sigma_{11} + S_{12}\sigma_{22} + S_{13}\sigma_{33} + S_{14}\sigma_{23} + S_{15}\sigma_{31} + S_{16}\sigma_{12} \\ \varepsilon_{22} = S_{21}\sigma_{11} + S_{22}\sigma_{22} + S_{23}\sigma_{33} + S_{24}\sigma_{23} + S_{25}\sigma_{31} + S_{26}\sigma_{12} \\ \varepsilon_{33} = S_{31}\sigma_{11} + S_{32}\sigma_{22} + S_{33}\sigma_{33} + S_{34}\sigma_{23} + S_{35}\sigma_{31} + S_{36}\sigma_{12} \\ \gamma_{23} = S_{41}\sigma_{11} + S_{42}\sigma_{22} + S_{43}\sigma_{33} + S_{44}\sigma_{23} + S_{45}\sigma_{31} + S_{46}\sigma_{12} \\ \gamma_{31} = S_{51}\sigma_{11} + S_{52}\sigma_{22} + S_{53}\sigma_{33} + S_{54}\sigma_{23} + S_{55}\sigma_{31} + S_{56}\sigma_{12} \\ \gamma_{12} = S_{61}\sigma_{11} + S_{62}\sigma_{22} + S_{63}\sigma_{33} + S_{64}\sigma_{23} + S_{65}\sigma_{31} + S_{66}\sigma_{12} \end{cases} \tag{2.59}$$

The S_{IJ} are now related to the S_{ijkl} by:

$$\begin{aligned} &S_{11} = S_{1111}, S_{12} = S_{1122}, \text{ etc.} \\ &S_{14} = S_{41} = 2S_{1123}, S_{15} = S_{51} = 2S_{1131}, \text{ etc.} \\ &S_{44} = 4S_{2323}, S_{45} = 4S_{2331}, \text{ etc.} \end{aligned} \tag{2.60}$$

We shall assume these relations in what follows.

NB: One way in which anisotropy manifests itself is that hydrostatic pressure ($\underline{\underline{\sigma}} = -p\underline{\underline{\delta}}$) does not in general produce a pure dilatation (*i.e.*, $\varepsilon_{11} \neq \varepsilon_{22} \neq \varepsilon_{33}$ and the distortions γ_{12}, γ_{23}, γ_{31} are non-zero); and conversely, a pure dilatation gives rise to shear stresses.

2.3.2.2 Taking Account of Material Symmetries: General Relation

It follows from Eq. 1.47 of Chap. 1 that, if the frame of reference is not changed by the operation $\underline{\underline{A}}$, the symmetries of the material being taken into account, then the functional relation $\underline{\underline{f}}$ for the constitutive equation $\underline{\underline{\sigma}} = \underline{\underline{f}}\left[\underline{\underline{\varepsilon}}\right]$ satisfies the relation

$$\underline{\underline{f}}\left[\underline{\underline{A}}.\underline{\underline{\varepsilon}}.\underline{\underline{A}}^T\right] = \underline{\underline{A}}.\underline{\underline{f}}\left[\underline{\underline{\varepsilon}}\right].\underline{\underline{A}}^T, \quad \forall \underline{\underline{\varepsilon}} \tag{2.61}$$

For linear elasticity this becomes simply:

$$\underline{\underline{A}}.(\mathbf{C} : \underline{\underline{\varepsilon}}).\underline{\underline{A}}^{\mathrm{T}} = \mathbf{C} : (\underline{\underline{A}}.\underline{\underline{\varepsilon}}.\underline{\underline{A}}^{\mathrm{T}}), \quad \forall \underline{\underline{\varepsilon}} \tag{2.62}$$

or in index notation:

$$A_{ip} C_{pqmn} \varepsilon_{mn} A_{jq} = C_{ijkl} A_{km} \varepsilon_{mn} A_{ln}, \forall \varepsilon_{mn} \tag{2.63}$$

From the orthogonality of $\underline{\underline{A}}$, implying $\underline{\underline{A}}^{\mathrm{T}} = \underline{\underline{A}}^{-1}$ ($A_{km}A_{lm} = \delta_{kl}$), we can write (2.63) as

$$C_{ijkl} = A_{ip} A_{jq} A_{km} A_{ln} C_{pqmn} \tag{2.64}$$

NB: We can also derive (2.62) by expressing the fact that if we "rotate" $\underline{\underline{\varepsilon}}$ (that is, rotate its principal axes) by $\underline{\underline{A}}$ without changing the frame of reference, $\underline{\underline{\sigma}}$ "rotates" in the same way, the relation between $\underline{\underline{\sigma}}$ and $\underline{\underline{\varepsilon}}$ being still described by the same tensor $\mathbf{C}$:

$$\begin{aligned} &\underline{\underline{\sigma}} \rightarrow \underline{\underline{\sigma}}' = \underline{\underline{A}}.\underline{\underline{\sigma}}.\underline{\underline{A}}^{\mathrm{T}} \\ &\underline{\underline{\varepsilon}} \rightarrow \underline{\underline{\varepsilon}}' = \underline{\underline{A}}.\underline{\underline{\varepsilon}}.\underline{\underline{A}}^{\mathrm{T}} \\ &\underline{\underline{\sigma}} = \mathbf{C} : \underline{\underline{\varepsilon}} \quad \underline{\underline{\sigma}}' = \mathbf{C} : \underline{\underline{\varepsilon}}' \end{aligned} \tag{2.65}$$

thus

$$\underline{\underline{A}}.\left(\mathbf{C} : \underline{\underline{\varepsilon}}\right).\underline{\underline{A}}^{\mathrm{T}} = \mathbf{C} : \left(\underline{\underline{A}}.\underline{\underline{\varepsilon}}.\underline{\underline{A}}^{\mathrm{T}}\right) \tag{2.62}$$

Equation 2.64 enables us to study various cases of material symmetry, either for crystalline materials or heterogeneous materials reinforced in particular directions, such as composites with fibres or laminates (Fig. 2.9). We shall consider several examples.

2.3.2.3 Symmetry with Respect to One Co-ordinate Plane ($x_3 = 0$)

$$[A] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{2.66}$$

Equation 2.64 implies that all the components of C_{ijkl} with an odd number of indices of value three are zero; therefore, in the C_{IJ} notation

$$C_{14} = C_{24} = C_{34} = C_{64} = C_{15} = C_{25} = C_{35} = C_{65} = 0 \tag{2.67}$$

and there are only 13 independent coefficients.

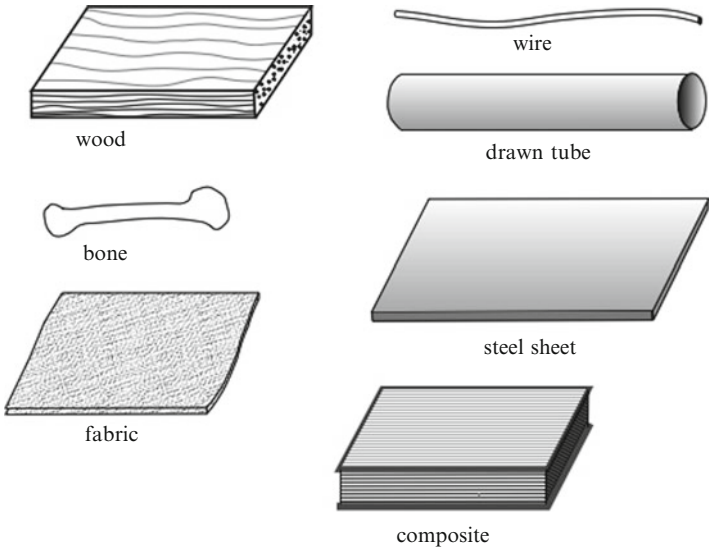


Fig. 2.9 Some anisotropic materials

2.3.2.4 Symmetry with Respect to a Pair of Orthogonal Planes ($x_3 = 0, x_1 = 0$)

The same argument shows that in addition to (2.67)

$$C_{16} = C_{26} = C_{36} = C_{45} = 0 \tag{2.68}$$

leaving only 9 independent coefficients. Note that for the elastic moduli **C** this shows that there is symmetry with respect to the third orthogonal plane, $x_2 = 0$, also; this is the case of *orthotropic* symmetry, corresponding to orthorhombic crystal structure.

2.3.2.5 Quadratic Symmetry

In addition to orthotropy there is the equivalence of the two axes of symmetry, say 1 and 2; this adds the further relations

$$C_{11} = C_{22}, C_{13} = C_{23}, C_{44} = C_{55} \tag{2.69}$$

reducing the number of independent coefficients to 6 (the tetragonal system).

2.3.2.6 Cubic Symmetry

In addition to the quadratic symmetry, there is the equivalence of all three axes of symmetry (3 in addition to 1 and 2):

$$C_{11} = C_{33}, C_{13} = C_{12}, C_{44} = C_{66} \tag{2.70}$$

leaving only three independent coefficients:

$$C_{11} = A(= C_{22} = C_{33}); \quad C_{12} = A'(= C_{13} = C_{23}); \quad C_{66} = D(= C_{44} = C_{55}) \tag{2.71}$$

with all others zero.

Note: In this case a hydrostatic pressure produces a pure dilatation.

2.3.2.7 Transverse Isotropy

This refers to invariance with respect to arbitrary rotation about a particular axis, for example about x_3 defined by

$$[A] = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \forall\alpha \tag{2.72}$$

This implies not only quadratic symmetry, and hence the relations (2.67–2.69), but also a further relation between different components. Applying (2.64) to C_{1212}, with $[A]$ the above matrix, and taking account of the vanishing of various components, we get

$$\begin{aligned} C_{1212} &= -\sin\alpha\cos\alpha A_{1p}A_{2q}C_{11pq} + \left(\cos^2\alpha - \sin^2\alpha\right) A_{1p}A_{2q}C_{12pq} \\ &\quad + \sin\alpha\cos\alpha A_{1p}A_{2q}C_{22pq} \\ &= -\sin\alpha\cos\alpha\left(-\sin\alpha\cos\alpha C_{1111} + \sin\alpha\cos\alpha C_{1122}\right) \\ &\quad + \left(\cos^2\alpha - \sin^2\alpha\right) C_{1212} \\ &\quad + \sin\alpha\cos\alpha\left(-\sin\alpha\cos\alpha C_{2211} + \sin\alpha\cos\alpha C_{2222}\right) \end{aligned} \tag{2.73}$$

Hence

$$C_{66} = \sin^2\alpha\cos^2\alpha(C_{11} + C_{22} - 2C_{12}) + (\cos^2\alpha - \sin^2\alpha)^2 C_{66} \tag{2.74}$$

i.e., $4\sin^2\alpha\cos^2\alpha\, C_{66} = 2\sin^2\alpha\cos^2\alpha\,(C_{11} - C_{12})$, $\forall\alpha$ and

$$C_{66} = (1/2)(C_{11} - C_{12}) \tag{2.75}$$

Thus there are at most five independent components; this is the case for *hexagonal crystal systems,* or *honeycomb* structures in composites.

Note: For the compliances **S** the relation corresponding to (2.75), taking (2.60) into account, is

$$S_{66} = 2(S_{11} - S_{12}) \tag{2.76}$$

2.3.2.8 Isotropy

We can now consider isotropy as the effect of a combination of cubic symmetry with a transverse isotropy with respect to one of the axes of the cube. Both (2.70) and (2.74) now apply and we have therefore

$$D = (1/2)\,(A - A') \tag{2.77}$$

and there are only two independent elastic moduli. Putting $\lambda = A'$, $\mu = D$, the Lamé[20] coefficients (so that $A = \lambda + 2\,\mu$), we get the tensor relation, valid in any system of axes,

$$\underline{\underline{\sigma}} = \lambda \mathrm{Tr}(\underline{\underline{\varepsilon}})\,\underline{\underline{\delta}} + 2\mu\underline{\underline{\varepsilon}} \tag{2.78}$$

and hence for the traces and the deviatoric parts

$$\mathrm{Tr}(\underline{\underline{\sigma}}) = (3\lambda + 2\mu)\mathrm{Tr}(\underline{\underline{\varepsilon}}) \qquad \underline{\underline{s}} = 2\mu\underline{\underline{e}} \tag{2.79}$$

Stresses and strains have the same principal axes; if we define the bulk modulus k as the ratio of the mean stress $\sigma_{\mathrm{m}} = (1/3)\ \mathrm{Tr}(\underline{\underline{\sigma}})$ to the volume dilatation $\theta = \mathrm{Tr}(\underline{\underline{\varepsilon}})$ then

$$k = \lambda + 2\mu/3 \tag{2.80}$$

Introducing Young modulus E and Poisson ratio ν, which refer to the analysis of the tensile test, we can put (2.78) in the form

$$\underline{\underline{\varepsilon}} = \frac{1+\nu}{E}\underline{\underline{\sigma}} - \frac{\nu}{E}\mathrm{Tr}(\underline{\underline{\sigma}})\underline{\underline{\delta}} \tag{2.81}$$

[20]Gabriel Lamé (1795–1870) was a French mathematician.

Table 2.2 Relation for passing from any pair of elastic constants to any other

–	E, ν	E, μ	k, ν	k, μ	λ, μ
E	E	E	$3(1-2\nu)k$	$\frac{9k}{1+3k/\mu}$	$\frac{\mu(3+2\mu/\lambda)}{1+\mu/\lambda}$
ν	ν	$-1+E/2\mu$	ν	$\frac{1-2\mu/3k}{2+2\mu/3k}$	$\frac{1}{2(1+\mu/\lambda)}$
μ	$\frac{E}{2(1+\nu)}$	μ	$\frac{3(1-2\nu)k}{2(1+\nu)}$	μ	μ
k	$\frac{E}{3(1-2\nu)}$	$\frac{E}{3(3-E/\mu)}$	k	k	$\lambda+2\mu/3$
λ	$\frac{E\nu}{(1+\nu)(1-2\nu)}$	$\frac{E(1-2\mu/E)}{3-E/\mu}$	$\frac{3k\nu}{1+\nu}$	$k-2\mu/3$	λ

E Young modulus
ν Poisson ratio
μ shear (Coulomb) modulus
k bulk (compressibility) modulus
λ, μ Lamé constants

with the following relations:

$$\nu = \frac{\lambda}{2(\lambda+\mu)} \qquad E = \frac{\mu(3\lambda+2\mu)}{\lambda+\mu} \qquad k = \frac{E}{3(1-2\nu)} \tag{2.82}$$

Incompressibility corresponds to $k \to \infty$ or $\nu = 1/2$.

Note: (2.76) does not hold for the cubic symmetry, for which the departure from isotropy can be characterised by a single dimensionless parameter a defined by

$$a = \frac{2(S_{11} - S_{12})}{S_{44}} \tag{2.83}$$

Values of this are: tungsten 1.0 (which is thus isotropic, despite its cubic structure); aluminium 1.4 (almost isotropic); iron 2.4; copper 3.3 (very anisotropic); β brass 8.75.

The relations between the elastic constants E, ν, μ, k, and λ are given in Table 2.2.

2.3.3 *Stability of the Equilibrium*

This is associated with the result that, as we saw in Sect. 2.3.1, the volume density of elastic energy is a positive definite quadratic form ($\underline{\underline{\varepsilon}} : \mathbf{C} : \underline{\underline{\varepsilon}}$ and $\underline{\underline{\sigma}} : \mathbf{S} : \underline{\underline{\sigma}}$ positive $\forall \underline{\underline{\varepsilon}}, \forall \underline{\underline{\sigma}}$ and zero only if $\underline{\underline{\varepsilon}} = 0$ or $\underline{\underline{\sigma}} = 0$). For the isotropic case it is easily shown,

by decomposing $\underline{\underline{\varepsilon}}$ and $\underline{\underline{\sigma}}$ into their deviatoric and spherical parts, that we must have

$$\mu \geq 0, \lambda + 2\mu/3 \geq 0, E \geq 0, -1 \leq \nu \leq 1/2 \tag{2.84}$$

For cubic symmetry we find

$$C_{11} \geq 0, C_{44} \geq 0, {C_{11}}^2 - {C_{12}}^2 \geq 0, C_{11} + 2C_{12} \geq 0 \tag{2.85}$$

2.3.4 Field Equations

Solving a problem in elasticity involves that all the following equations and conditions are satisfied simultaneously:

(a) equilibrium $\mathrm{div}\underline{\underline{\sigma}} + \underline{f} - \rho\underline{\gamma} = 0$
(b) compatibility $\mathbf{Inc}\,(\underline{\underline{\varepsilon}}) = 0$
(c) constitutive equations $\underline{\underline{\sigma}} = \mathbf{C} : \underline{\underline{\varepsilon}}$ or $\underline{\underline{\varepsilon}} = \mathbf{S} : \underline{\underline{\sigma}}$
(d) the boundary conditions

If the problem is solved in terms of displacements then (b) is satisfied automatically and in rectangular cartesian coordinates (a) and (c) are combined in the Lamé-Clapeyron[21] (or Navier[22]) equations:

$$C_{ijkl} u_{k,lj} + f_i - \rho\, \partial^2 u_i / \partial t^2 = 0 \tag{2.86}$$

If on the other hand the solution is in terms of stresses (b) and (c) can be combined into

$$\in_{ikl} \in_{jmn} S_{kmpq} \sigma_{pq,ln} = 0 \tag{2.87}$$

where $\in_{ikl}$ is the permutation tensor, with value 0 if any pair of suffices are equal and ± 1 according as (i,k,l) is an even or odd permutation of (1,2,3).

2.3.5 Example: Propagation of Plane Sine Waves

2.3.5.1 General Equation

We are looking for a solution to the Lamé-Clapeyron Eq. 2.86 with $f_i = 0$ of the form

$$u_i = u_i^0 \exp\left[\mathrm{i}(K_k x_k - \omega t)\right] \tag{2.88}$$

[21] Émile Clapeyron (1799–1864) was a French physicist.
[22] Henri Navier (1785–1836) was a French physicist.

We have

$$u_{k,jl} = -u_i^0 K_j K_l \exp\left[\mathrm{i}(K_m x_m - \omega t)\right] \tag{2.89}$$

$$\rho \frac{\partial^2 u_i}{\partial t^2} = -\rho\omega^2 u_i^0 \exp\left[\mathrm{i}(K_m x_m - \omega t)\right] \tag{2.90}$$

Hence

$$\left(C_{ijkl} K_j K_l - \rho\omega^2 \delta_{ki}\right) u_k^0 = 0 \tag{2.91}$$

and for a solution other than ${u_k}^0 = 0$ we must have

$$\det(C_{ijkl} K_j K_l - \rho\omega^2 \delta_{ki}) = 0 \tag{2.92}$$

2.3.5.2 Isotropic Materials

We now have

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{2.93}$$

whence

$$\begin{aligned} C_{ijkl} K_j K_l &= \lambda K_i K_k + \mu(K^2 \delta_{ik} + K_i K_k) \\ &= (\lambda + \mu) K_i K_k + \mu K^2 \delta_{ik} \end{aligned} \tag{2.94}$$

Equation 2.92 now becomes

$$\begin{vmatrix} (\lambda+\mu)\,K_1^2 + \mu K^2 - \rho\omega^2 & (\lambda+\mu)\,K_1 K_2 & (\lambda+\mu)\,K_1 K_3 \\ (\lambda+\mu)\,K_1 K_2 & (\lambda+\mu)\,K_2^2 + \mu K^2 - \rho\omega^2 & (\lambda+\mu)\,K_2 K_3 \\ (\lambda+\mu)\,K_1 K_3 & (\lambda+\mu)\,K_2 K_3 & (\lambda+\mu)\,K_3^2 + \mu K^2 - \rho\omega^2 \end{vmatrix} = 0 \tag{2.95}$$

Taking $\underline{x}$ in the direction of $\underline{K}$ ($K_2 = K_3 = 0$) we get

$$(\mu K^2 - \rho\omega^2)^2 [(\lambda + 2\mu) K^2 - \rho\omega^2] = 0 \tag{2.96}$$

The wave propagation velocity is $c = \omega/K$; thus for transverse propagation (for which $u_1^0 = 0, \quad u_2^0 \neq 0, \quad u_3^0 \neq 0$) we get $c_{\mathrm{T}} = \sqrt{\mu/\rho}$ and for longitudinal propagation (for which $u_1^0 \neq 0, \ u_2^0 = 0, \ u_3^0 = 0$) $c_{\mathrm{L}} = \sqrt{(\lambda + 2\mu)/\rho}$.

2.3.5.3 Cubic Symmetry (in the Axes of the Cube)

With the symmetry relations (2.67), (2.68) and (2.71), the determinantal Eq. 2.92 becomes

$$\begin{vmatrix} C_{11}K_1^2 + C_{44}(K_2^2 + K_3^2) - \rho\omega^2 & (C_{12} + C_{44})\,K_1K_2 & (C_{12} + C_{44})\,K_1K_3 \\ (C_{12} + C_{44})\,K_1K_2 & C_{11}K_2^2 + C_{44}(K_1^2 + K_3^2) - \rho\omega^2 & (C_{12} + C_{44})\,K_2K_3 \\ (C_{12} + C_{44})\,K_1K_3 & (C_{12} + C_{44})\,K_2K_3 & C_{11}K_3^2 + C_{44}(K_1^2 + K_2^2) - \rho\omega^2 \end{vmatrix} = 0 \tag{2.97}$$

For propagation in the direction [100], for example, this is

$$\begin{vmatrix} C_{11}K^2 - \rho\omega^2 & 0 & 0 \\ 0 & C_{44}K^2 - \rho\omega^2 & 0 \\ 0 & 0 & C_{44}K^2 - \rho\omega^2 \end{vmatrix} = 0 \tag{2.98}$$

and the transverse and longitudinal propagations velocities are

$$c_{\mathrm{T}} = \sqrt{C_{44}/\rho}, \quad c_{\mathrm{L}} = \sqrt{C_{11}/\rho} \tag{2.99}$$

For other directions the velocities can be different (see Exercises in Volume III).

2.4 Variational Methods: Introduction to the Finite Elements Method

Modern numerical methods for structural calculations are based on variational principles; the general aim is to find a set of values for the variables of the problem that minimise or maximise a certain functional relation that relates to the complete structure. This is done by considering many admissible fields, which partially satisfy the boundary conditions of the problem, and searching for the one which is the best approximation to the solution. The basic tools used in constructing variational methods are (1) a method for the construction of many admissible fields, (2) the virtual work theorem, applied to this set of admissible fields, and (3) a fundamental inequality deriving from the behaviour of the material of which the structure is made, which can be used for the comparison of the trial fields with the (unknown) solution. For the case of linear elasticity, considered here, this inequality is a consequence of the positive definite nature of the quadratic form associated with the tensor of the elastic moduli **C** or compliances **S** (see for example the discussion of elastic stability in Sect. 2.3.3).

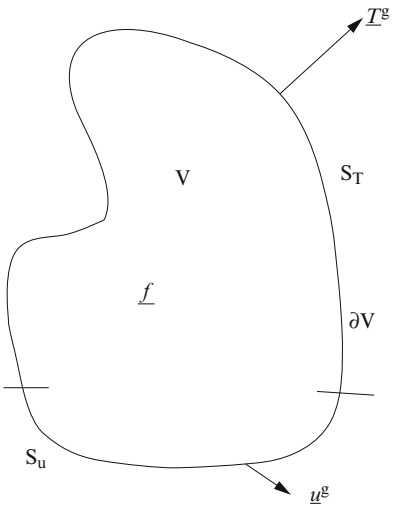


Fig. 2.10 Description of the boundary value problem being discussed

2.4.1 *Extremal Theorems in Linear Elasticity*

We use the virtual work theorem in the form (2.4), implying both a SA stress field $\underline{\underline{\sigma}}^*$ (associated with $\underline{f}^*$ and $\underline{T}^* = \underline{\underline{\sigma}}^*.\underline{n}$) and a KA strain field $\underline{\underline{\varepsilon}}'(\underline{u}')$ deriving from the displacement field $\underline{u}'$ (associated with $\underline{\underline{\varepsilon}}'$)

Consider now the problem of an elastic body subjected to body forces $\underline{f}$ and surface conditions of the type

prescribed displacements $\underline{u}^{g}$ over an area S_u of ∂V
prescribed forces $\underline{T}^{g}$ over the remaining part S_T (S_u or S_T can be null) (Fig. 2.10).

Let us adopt now a more restricted definition of the admissible fields $\underline{\underline{\sigma}}^*$ and $\underline{u}'$: they are statically (SA') and kinematically (KA') admissible *for the problem under consideration*, respectively, which means that $\underline{\underline{\sigma}}^*$ must be SA with $\underline{f}$, *i.e.*, $\mathrm{div}\underline{\underline{\sigma}}^* + \underline{f} = 0$, and must satisfy the boundary conditions over S_T, *i.e.*, $\underline{T}^* = \underline{\underline{\sigma}}^*.\underline{n} = \underline{T}^{g}$ over S_T whereas $\underline{\underline{\varepsilon}}'(\underline{u}')$ is compatible and $\underline{u}'$ must satisfy the boundary conditions over S_u , *i.e.*, $\underline{u}' = \underline{u}^{g}$ over S_u.

NB: $\underline{\underline{\varepsilon}}^*$ deduced from $\underline{\underline{\sigma}}^*$ by $\underline{\underline{\varepsilon}}^* = \mathbf{S} : \underline{\underline{\sigma}}^*$ need not be compatible, nor does $\underline{\underline{\sigma}}'$ deduced from $\underline{\underline{\varepsilon}}'$ by $\underline{\underline{\sigma}}' = \mathbf{C} : \underline{\underline{\varepsilon}}'$ need to be in equilibrium.

Application of the virtual work theorem (2.4) to $\underline{\underline{\sigma}}^*$ and $\underline{u}'$ gives at once

$$\int_V \underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}' \mathrm{d}V = \int_V \underline{f}.\underline{u}' \mathrm{d}V + \int_{S_T} \underline{T}^{g}.\underline{u}' \mathrm{d}S + \int_{S_u} \underline{T}^*.\underline{u}^{g} \mathrm{d}S \qquad (2.100)$$

We now use the fact that the quadratic forms $\underline{\underline{\varepsilon}} : \mathbf{C} : \underline{\underline{\varepsilon}}$, $\underline{\underline{\sigma}} : \mathbf{S} : \underline{\underline{\sigma}}$ are positive definite, *i.e.,*

$$\underline{\underline{\varepsilon}} : \mathbf{C} : \underline{\underline{\varepsilon}} \geq 0 \quad \forall \underline{\underline{\varepsilon}}, \quad \underline{\underline{\varepsilon}} : \mathbf{C} : \underline{\underline{\varepsilon}} = 0, \quad \text{iff } \underline{\underline{\varepsilon}} = 0 \tag{2.101a}$$

$$\underline{\underline{\sigma}} : \mathbf{S} : \underline{\underline{\sigma}} \geq 0 \quad \forall \underline{\underline{\sigma}}, \quad \underline{\underline{\sigma}} : \mathbf{S} : \underline{\underline{\sigma}} = 0, \quad \text{iff } \underline{\underline{\sigma}} = 0 \tag{2.101b}$$

In (2.101b) let us take, for example, $(\underline{\underline{\sigma}}^* - \mathbf{C} : \underline{\underline{\varepsilon}}')$ as stress field, where $\underline{\underline{\sigma}}^*$ and $\underline{\underline{\varepsilon}}'$ are SA' and KA', respectively. Note that in general $\underline{\underline{\sigma}}^*$ and $\underline{\underline{\varepsilon}}'$ are not related, that is, $\underline{\underline{\sigma}}^* \neq \mathbf{C} : \underline{\underline{\varepsilon}}'$; only the fields $\underline{\underline{\sigma}}$ and $\underline{\underline{\varepsilon}}$, the solutions to the problem, are both admissible and related.

It follows that the quadratic form $\varphi(\underline{\underline{\sigma}}^*, \underline{u}') = (1/2)(\underline{\underline{\sigma}}^* - \mathbf{C} : \underline{\underline{\varepsilon}}') : \mathbf{S} : (\underline{\underline{\sigma}}^* - \mathbf{C} : \underline{\underline{\varepsilon}}')$ is positive definite and is zero only for $\underline{\underline{\sigma}}$ and $\underline{u}$, the solutions to the problem, *i.e.*,

$$\varphi(\underline{\underline{\sigma}}, \underline{u}) = 0 \tag{2.102}$$

We find, easily:

$$\varphi(\underline{\underline{\sigma}}^*, \underline{u}') = (1/2)\, \underline{\underline{\sigma}}^* : \mathbf{S} : \underline{\underline{\sigma}}^* - \underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}' + (1/2)\, \underline{\underline{\varepsilon}}' : \mathbf{C} : \underline{\underline{\varepsilon}}' \tag{2.103}$$

The functional $\Phi(\underline{\underline{\sigma}}^*, \underline{u}')$ defined by $\Phi(\underline{\underline{\sigma}}^*, \underline{u}') = \int_V \varphi(\underline{\underline{\sigma}}^*, \underline{u}') \mathrm{d}V$ is similarly positive, and zero only for $\underline{\underline{\sigma}}$ and $\underline{u}$; using (2.100) we can write it in the form

$$\begin{aligned} \Phi(\underline{\underline{\sigma}}^*, \underline{u}') = \frac{1}{2} \int_V \underline{\underline{\sigma}}^* : \mathbf{S} : \underline{\underline{\sigma}}^* \mathrm{d}V - \int_V \underline{f} . \underline{u}' \, \mathrm{d}V - \int_{S_\mathrm{T}} \underline{T}^\mathrm{g} . \underline{u}' \, \mathrm{d}S - \int_{S_\mathrm{u}} \underline{T}^* . \underline{u}^\mathrm{g} \, \mathrm{d}S \\ + \frac{1}{2} \int_V \underline{\underline{\varepsilon}}' : \mathbf{C} : \underline{\underline{\varepsilon}}' \mathrm{d}V \end{aligned} \tag{2.104}$$

which shows that it can be expressed as the sum of two *independent* functionals:

$$\Phi(\underline{\underline{\sigma}}^*, \underline{u}') = \Phi_1(\underline{\underline{\sigma}}^*) + \Phi_2(\underline{u}') \quad \geq 0, \ \forall \underline{\underline{\sigma}}^*, \underline{u}' \tag{2.105}$$

where

$$\begin{cases} \Phi_1(\underline{\underline{\sigma}}^*) = \frac{1}{2} \int_V \underline{\underline{\sigma}}^* : \mathbf{S} : \underline{\underline{\sigma}}^* \, \mathrm{d}V - \int_{S_\mathrm{u}} \underline{T}^* . \underline{u}^\mathrm{g} \, \mathrm{d}S \\ \Phi_2(\underline{u}') = \frac{1}{2} \int_V \underline{\underline{\varepsilon}}' : \mathbf{C} : \underline{\underline{\varepsilon}}' \, \mathrm{d}V - \int_V \underline{f} . \underline{u}' \, \mathrm{d}V - \int_{S_\mathrm{T}} \underline{T}^\mathrm{g} . \underline{u}' \, \mathrm{d}S \end{cases} \tag{2.106}$$

Since Φ_1 and Φ_2 vary independently and their sum is minimal for $\underline{\underline{\sigma}}^* = \underline{\underline{\sigma}}$ and $\underline{u}' = \underline{u}$, it follows that Φ_1 is minimal for $\underline{\underline{\sigma}}^* = \underline{\underline{\sigma}}$ and Φ_2 is minimal for $\underline{u}' = \underline{u}$; so we have

$$\Phi_1(\underline{\underline{\sigma}}) \leq \Phi_1(\underline{\underline{\sigma}}^*) \text{ minimum complementary energy}$$

$$\Phi_2(\underline{u}) \leq \Phi_2(\underline{u}') \text{ minimum potential energy}$$

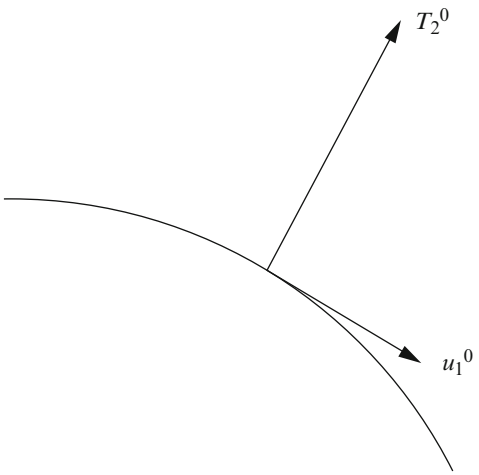


Fig. 2.11 Example of extended boundary conditions

where the equalities hold only for $\underline{\underline{\sigma}}^* = \underline{\underline{\sigma}}$ and $\underline{u'} = \underline{u}$, respectively. From (2.102) it follows

$$-\Phi_1(\underline{\underline{\sigma}}^*) \leq -\Phi_1(\underline{\underline{\sigma}}) = \Phi_2(\underline{u}) \leq \Phi_2(\underline{u'}) \tag{2.107}$$

This summarises the two extremal theorems we were seeking, each leading to a particular numerical method: to find, as the best approximation, among the set of SA' fields for a given problem, the one that minimises $\Phi_1(\underline{\underline{\sigma}}^*)$; or, among the set of KA' fields for this problem, the one that minimises $\Phi_2(\underline{u'})$.

Note: The class of problems that can be treated by these methods can be extended by allowing more general boundary conditions. We can assume only that at every point of the surface ∂V there is given, for the direction of each axis of the frame of reference, a component of either force or displacement. Thus Fig. 2.11 illustrates, in two dimensions, conditions which can be prescribed at the boundary. The above relations (2.104) then hold with $\underline{T}^{\mathrm{g}} = (0,\ T_2{}^0)$ and $\underline{u}^{\mathrm{g}} = (u_1{}^0,\ 0)$, provided that the surface integrals are taken over the whole surface ∂V, since there

$$\underline{T}^*.\underline{u'} = T_1^* u_1^0 + T_2^0 u_2' = \underline{T}^*.\underline{u}^{\mathrm{g}} + \underline{T}^{\mathrm{g}}.\underline{u'} \tag{2.108}$$

2.4.2 Principle of the Finite Element Method

Without going into the details of the method we show how the principle on which it is based follows from the above theorems, once the structure to be studied has been

discretised into geometrically simple *finite elements*. We give the method in terms of displacements, using the form $\Phi_2(\underline{u}')$.

Let $\{\underline{\varphi}_i\}$ be a base set of n vector functions, each attached to a node i of the finite-element mesh, in terms of which we can express a trial displacement field as

$$\underline{u}' = \sum_{i=1}^{n} X_i \underline{\varphi}_i \ \ (= X_i \underline{\varphi}_i) \tag{2.109}$$

The coefficients X_i are often the components of the displacements at the nodes and the $\underline{\varphi}_i$ (polynomial, trigonometric functions, etc.) interpolation functions that enable the value of $\underline{u}'$ to be determined at any point from the values at the nodes. The boundary conditions $\underline{u}' = \underline{u}^{\mathrm{g}}$ on S_{u} are given by p relations between the X_i; for the sake of simplicity we consider these as setting the first p values of X_i, from X_1 to X_p.

Since (2.109) is linear in the X_i ,

$$\underline{\underline{\varepsilon}}'(X_i \underline{\varphi}_i) = X_i \underline{\underline{\varepsilon}}'(\underline{\varphi}_i) \tag{2.110}$$

We then have

$$\begin{aligned}
\Phi_2(\underline{u}') &= \frac{1}{2}\int_V \underline{\underline{\varepsilon}}'(X_i\underline{\varphi}_i) : \mathbf{C} : \underline{\underline{\varepsilon}}'(X_i\underline{\varphi}_i)\,\mathrm{d}V - \int_V \underline{f}.X_i\underline{\varphi}_i\,\mathrm{d}V - \int_{S_{\mathrm{T}}} \underline{T}^{\mathrm{g}}.X_i\underline{\varphi}_i\,\mathrm{d}S \\
&= \frac{1}{2}X_i\left[\int_V \underline{\underline{\varepsilon}}'(\underline{\varphi}_i) : \mathbf{C} : \underline{\underline{\varepsilon}}'(\underline{\varphi}_j)\,\mathrm{d}V\right]X_j - X_i\left[\int_V \underline{f}.\underline{\varphi}_i\,\mathrm{d}V + \int_{S_{\mathrm{T}}} \underline{T}^{\mathrm{g}}.\underline{\varphi}_i\,\mathrm{d}S\right] \\
&= \frac{1}{2}X_i R_{ij} X_j - X_i F_i \quad \text{with } R_{ij} = R_{ji}
\end{aligned} \tag{2.111}$$

where in general it is arranged that each component of the matrix R_{ij} (or of F_i) involves only one, or at most a small number, of the finite elements. If we put

$$[X] = \left(X_1 \cdots X_p\ X_{p+1} \cdots X_n\right)^{\mathrm{T}} = \begin{pmatrix} X_1 \\ \cdots \\ X_{\mathrm{p}} \\ X_{\mathrm{p+1}} \\ \cdots \\ X_{\mathrm{n}} \end{pmatrix} = \begin{pmatrix} [X_1] \\ [X_2] \end{pmatrix} \begin{matrix} p \\ n-p \end{matrix}$$

$$[F] = \begin{pmatrix} [F_1] \\ [F_2] \end{pmatrix} \begin{matrix} p \\ n-p \end{matrix}$$

$$[R] = \begin{pmatrix} [R_{11}]\ [R_{12}] \\ [R_{21}]\ [R_{22}] \end{pmatrix} \begin{matrix} p \\ n-p \end{matrix} \tag{2.112}$$

then with $[R_{12}] = [R_{21}]^{\mathrm{T}}$ we have

$$\begin{aligned}\Phi_2(\underline{u'}) = \Phi_2([X_1], [X_2]) &= \frac{1}{2}[X]^{\mathrm{T}}.[R].[X] - [X]^{\mathrm{T}}.[F] \\ &= \frac{1}{2}\left([X_1]^{\mathrm{T}}.[R_{11}].[X_1] + [X_1]^{\mathrm{T}}.[R_{12}].[X_2] + [X_2]^{\mathrm{T}}.[R_{21}].[X_1]\right) + \ldots \\ &\ldots + \frac{1}{2}\left([X_2]^{\mathrm{T}}.[R_{22}].[X_2]\right) - [X_1]^{\mathrm{T}}.[F_1] - [X_2]^{\mathrm{T}}.[F_2] \end{aligned} \tag{2.113}$$

The minimisation of Φ_2 must then be done with respect to $[X_2]$, with $[X_1]$ fixed:

$$\begin{aligned}\mathrm{d}\Phi_2 &= \frac{1}{2}\left([X_1]^{\mathrm{T}}.[R_{12}].[\mathrm{d}X_2] + [\mathrm{d}X_2]^{\mathrm{T}}.[R_{21}].[X_1] + [\mathrm{d}X_2]^{\mathrm{T}}.[R_{22}].[X_2]\right)\ldots \\ &\ldots + \frac{1}{2}\left([X_2]^{\mathrm{T}}.[R_{22}].[\mathrm{d}X_2]\right) - [\mathrm{d}X_2]^{\mathrm{T}}.[F_2] \\ &= [\mathrm{d}X_2]^{\mathrm{T}}.\left\{\frac{1}{2}\left([R_{12}]^{\mathrm{T}}.[X_1] + [R_{21}].[X_1] + [R_{22}].[X_2] + [R_{22}]^{\mathrm{T}}.[X_2]\right) - [F_2]\right\} \\ &= [\mathrm{d}X_2]^{\mathrm{T}}.([R_{21}].[X_1] + [R_{22}].[X_2] - [F_2]) \end{aligned} \tag{2.114}$$

Setting $\mathrm{d}\Phi_2 = 0$, we have the *linear* system for X_2 to solve:

$$[R_{22}].[X_2] = [F_2] - [R_{21}].[X_1] \tag{2.115}$$

Here $[R_{22}]$ is the *stiffness matrix* of the structure.
We have passed over a number of questions that arise, in particular

- what shape to choose for the finite elements
- how to specify the nodes of the mesh, and what functions $\underline{\varphi_i}$ to choose as base
- how to enumerate the nodes so as to condition the matrix $[R]$ optimally
- how actually to construct the overall stiffness matrix
- how to solve the linear system (2.115).

We leave the reader to consult specialist works on these points.

2.5 Heterogeneous Materials: Estimates and Bounds in Linear Elasticity

The variational methods presented above have been mainly applied to classical situations of structural design where the geometry, the loading and the material properties are defined in a deterministic way. They can also be applied to RVEs

of random heterogeneous materials where we have only an incomplete knowledge of the spatial distribution of the different constitutive phases so that we cannot determine the overall behaviour of the HEM unambiguously. However, with the help of energy theorems, which lead to a variational approach to the problem, we may be able to establish *bounds* for the quantities that define the effective behaviour and even to find optimal bounds, that is, the closest possible consistent with the available information: the actual moduli for a material for which we have only this partial information can then be *anywhere* inside the domain defined by these optimal bounds. Remind that in the following we are only concerned with random materials. So, the specific and powerful methods that have been developed for materials with periodic microstructure are not discussed here. For an introduction to them, the interested reader may consult *e.g.* Suquet (1987). Laminates are not treated despite their important applications for the design of composite structures. Simple cases only are discussed as exercises in Volume III.

2.5.1 Effective Moduli and Compliances

Here we shall limit the treatment to the case of linear elasticity, with homogeneous boundary conditions on forces or displacements. The initial state is first supposed to be "natural" (*i.e.,* without any initial strains or stresses). Equations 1.37–38 then lead to the following relations for the effective moduli $\mathbf{C}^{\text{eff}}$ and compliances $\mathbf{S}^{\text{eff}}$

$$\begin{cases} \mathbf{C}^{\text{eff}} = <\mathbf{c} : \mathbf{A}> = <\mathbf{A}^{\text{T}} : \mathbf{c} : \mathbf{A}> \text{ with } <\mathbf{A}> = <\mathbf{A}^{\text{T}}> = \text{I} \\ \mathbf{S}^{\text{eff}} = <\mathbf{s} : \mathbf{B}> = <\mathbf{B}^{\text{T}} : \mathbf{s} : \mathbf{B}> \text{ with } <\mathbf{B}> = <\mathbf{B}^{\text{T}}> = \text{I} \end{cases} \tag{2.116}$$

Remember that $\mathbf{A}$ and $\mathbf{B}$ represent the strain and stress concentration tensors, respectively.

For a composite material with n phases, they read

$$\begin{cases} \mathbf{C}^{\text{eff}} = <\mathbf{c} : \mathbf{A}> = \sum_{r=1}^{n} f_r \mathbf{c}_r : \mathbf{A}_r, \ \mathbf{A}_r = <\mathbf{A}>_{(r)} \text{ with } \sum_{r=1}^{n} f_r \mathbf{A}_r = I \\ \mathbf{S}^{\text{eff}} = <\mathbf{s} : \mathbf{B}> = \sum_{r=1}^{n} f_r \mathbf{s}_r : \mathbf{B}_r, \ \mathbf{B}_r = <\mathbf{B}>_{(r)} \text{ with } \sum_{r=1}^{n} f_r \mathbf{B}_r = I \end{cases} \tag{2.117}$$

These relations show that $\mathbf{C}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$ have all the required symmetries of elastic moduli and compliances, including the diagonal symmetry ($C_{ijkl} = C_{klij}$ and $S_{ijkl} = S_{klij}$). They also imply that the quadratic forms associated with $\mathbf{C}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$ are positive definite as soon as this property holds for $\mathbf{c}$ and $\mathbf{s}$. For homogeneous boundary conditions and macrohomogeneity ($d \ll \ell$, see Chap. 1, Sect. 1.4.4.1), the homogenisation procedure yields the same results for homogeneous stress and strain boundary conditions, so that $\mathbf{C}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$ are inverse tensors.

2.5.2 *The Case of Initial Deformations*

In some cases the response of a heterogeneous material will involve not only the elastic behaviour but also deformations of different natures, such as thermal, plastic or resulting from phase changes. We shall consider the situation in which such an additional deformation can be regarded as given initially and not developing under the mechanical loading: the associated strains are then "stress-free strains" or "eigenstrains"; to this extent, when later on in this section we deal with "elastoplasticity" we shall exclude true elastoplastic flow, which will need to be treated by other methods. However, some problems, treated below, can be attacked by the same methods whether or not there is plastic flow; this is the case for the concepts of mean values and stored energy, which we shall treat for instance within the frameworks as well of elastoplasticity as of thermoelasticity, without calling on any description of plastic flow. In particular they are relevant to the important study of the dilatometric behaviour of composites whose constituents have different elastic moduli and different coefficients of thermal expansion.

2.5.2.1 Mean Values

For a heterogeneous material with eigenstrains, whilst the overall macroscopic stresses and (total) strains, and their rates of change, are still the means of local values, the same does not hold for the elastic and plastic parts of the strain field separately, because of their generally incompatible character (Mandel 1972). The total deformation is a combination of an elastic part $\underline{\underline{\varepsilon}}^{\mathrm{e}}(\underline{x})$ and a part having a physically different origin – thermal, plastic, phase change, or other, say an eigenstrain.

There is the local partition $\underline{\underline{\varepsilon}} = \underline{\underline{\varepsilon}}^{\mathrm{e}} + \underline{\underline{\varepsilon}}^{\mathrm{F}}$ (where F stands for "free") and we can also write $\underline{\underline{\sigma}} = \underline{\underline{\sigma}}' + \underline{\underline{\sigma}}^{\mathrm{r}}$ where $\underline{\underline{\sigma}}'$ is the stress field that would be set up if no eigenstrain was present in the heterogeneous medium, under the same loading, and $\underline{\underline{\sigma}}^{\mathrm{r}}$ is therefore the residual stress field. It follows that $<\underline{\underline{\sigma}}'> = <\underline{\underline{\sigma}}> = \underline{\underline{\Sigma}} \Rightarrow <\underline{\underline{\sigma}}^{\mathrm{r}}> = 0$ (self-equilibrated field).

As we know, $\underline{\underline{\sigma}}' = \mathbf{B} : \underline{\underline{\Sigma}}$, where $\mathbf{B}$ is the solution of the *purely elastic* problem (*i.e.*, without eigenstrains) of stress concentration. These relations follow:

$$
\begin{aligned}
<\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}> &= <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{e}}> + <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{F}}> = <\mathbf{B}^{\mathrm{T}} : \mathbf{s} : \underline{\underline{\sigma}}> + <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{F}}> \\
&= <\underline{\underline{\sigma}} : \mathbf{s} : \mathbf{B}> + <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{F}}> \quad (\text{since } \mathbf{s} \text{ is symmetric}) \\
&= <\underline{\underline{\sigma}}> : <\mathbf{s} : \mathbf{B}> + <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{F}}> \quad (\text{Hill's lemma, since } \mathbf{s} : \mathbf{B} \text{ is KA}) \\
&= <\underline{\underline{\sigma}}> : \mathbf{S}^{\mathrm{eff}} + <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{F}}> = \underline{\underline{E}}^{\mathrm{e}} + <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{F}}> = \underline{\underline{E}}^{\mathrm{e}} + <\underline{\underline{\varepsilon}}^{\mathrm{F}} : \mathbf{B}> \\
&= <\mathbf{B}^{\mathrm{T}}> : <\underline{\underline{\varepsilon}}> = <\underline{\underline{\varepsilon}} : \mathbf{B}> = \underline{\underline{E}} \quad (\text{Hill's lemma, since } \mathbf{B}^{\mathrm{T}} \text{ is SA})
\end{aligned}
\tag{2.118}
$$

Note that the fourth order tensors $\mathbf{s} : \mathbf{B}$ and $\mathbf{B}^{\mathrm{T}}$ have been treated here as second-order tensors by considering two of their indices as fixed: they can be then assimilated to a kinematically admissible strain field and a statically admissible stress field, respectively, so that Hill's lemma (1.34) can be used.

From that we have

$$\begin{aligned}
&\underline{\underline{E}} = <\underline{\underline{\varepsilon}}> \\
&\underline{\underline{E}}^{\mathrm{e}} = <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{e}}> = <\underline{\underline{\varepsilon}}^{\mathrm{e}} : \mathbf{B}> \ (\neq <\underline{\underline{\varepsilon}}^{\mathrm{e}}> \text{ in general}) \\
&\underline{\underline{E}}^{\mathrm{F}} = <\mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{F}} > = <\underline{\underline{\varepsilon}}^{\mathrm{F}} : \mathbf{B}> \ (\neq <\underline{\underline{\varepsilon}}^{\mathrm{F}}> \text{ in general})
\end{aligned} \tag{2.119}$$

while $<\underline{\underline{\sigma}}> = \underline{\underline{\Sigma}}$, $<\underline{\underline{\sigma'}}> = \mathbf{B} : \underline{\underline{\Sigma}} = \underline{\underline{\Sigma}}$, $<\underline{\underline{\sigma}}^{\mathrm{r}}> = 0$.

Let us briefly mention here without explicit development a "dual" analysis, which can be performed for initial stresses instead of initial strains. Denoting $\underline{\underline{p}}(\underline{x})(= -\mathbf{c} : \underline{\underline{\varepsilon}}^{\mathrm{F}})$ and $\underline{\underline{P}}(= -\mathbf{C}^{\mathrm{eff}} : \underline{\underline{E}}^{\mathrm{F}})$ the local and macroscopic *eigenstresses*, respectively, we have through similar arguments

$$\begin{aligned}
&\underline{\underline{\Sigma}} = <\underline{\underline{\sigma}}> \\
&\mathbf{C}^{\mathrm{eff}} : \underline{\underline{E}} = <\mathbf{A}^{\mathrm{T}} : (\mathbf{c} : \underline{\underline{\varepsilon}})> = <(\underline{\underline{\varepsilon}} : \mathbf{c}) : \mathbf{A}> \ (\neq <\mathbf{c} : \underline{\underline{\varepsilon}}> \text{ in general}) \\
&\underline{\underline{P}} = <\mathbf{A}^{\mathrm{T}} : \underline{\underline{p}}> = <\underline{\underline{p}} : \mathbf{A}> \ (\neq <\underline{\underline{p}}> \text{ in general})
\end{aligned} \tag{2.120}$$

2.5.2.2 Elastic Stored Energy

The internal energy per unit volume is given by $\frac{1}{2} <\underline{\underline{\sigma}} : \mathbf{s} : \underline{\underline{\sigma}}>$, the average of the local internal energy which can be assimilated to the microscopic elastic energy. From the definition of the residual stresses $\underline{\underline{\sigma}}^{\mathrm{r}}$, we have

$$\begin{aligned}
\frac{1}{2} <\underline{\underline{\sigma}} : \mathbf{s} : \underline{\underline{\sigma}}> &= \frac{1}{2} <(\underline{\underline{\sigma'}} + \underline{\underline{\sigma}}^{\mathrm{r}}) : \mathbf{s} : (\underline{\underline{\sigma'}} + \underline{\underline{\sigma}}^{\mathrm{r}})> \\
&= \frac{1}{2} <\underline{\underline{\sigma'}} : \mathbf{s} : \underline{\underline{\sigma'}}> + <\underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \underline{\underline{\sigma'}}> + \frac{1}{2} <\underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \underline{\underline{\sigma}}^{\mathrm{r}}> \ (\mathbf{s} = \mathbf{s}^{\mathrm{T}}) \\
&= \frac{1}{2} \underline{\underline{\Sigma}} :< \mathbf{B}^{\mathrm{T}} : \mathbf{s} : \mathbf{B}> : \underline{\underline{\Sigma}} + <\underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \underline{\underline{\sigma'}}> + \frac{1}{2} <\underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \underline{\underline{\sigma}}^{\mathrm{r}}> \\
&= \frac{1}{2} \underline{\underline{\Sigma}} : \mathbf{S}^{\mathrm{eff}} : \underline{\underline{\Sigma}} + \frac{1}{2} <\underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \underline{\underline{\sigma}}^{\mathrm{r}}> \\
&\quad \text{(Hill's lemma, since } (\mathbf{s} : \underline{\underline{\sigma'}}) \text{ is KA, } \underline{\underline{\sigma}}^{\mathrm{r}} \text{ is SA. and } <\underline{\underline{\sigma}}^{\mathrm{r}}> = 0)
\end{aligned} \tag{2.121}$$

Thus the internal energy comprises the macroscopic elastic energy and in addition ($< \underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \underline{\underline{\sigma}}^{\mathrm{r}} >$ cannot be negative since the quadratic form associated with $\mathbf{s}$ is positive definite) the elastic stored energy corresponding to the residual stresses.

2.5.2.3 Linear Thermoelasticity

Application to linear thermoelasticity of inhomogeneous materials is straightforward. Each phase (r) is characterised by its elastic moduli $\mathbf{c}_r$ and its coefficients of thermal expansion $\underline{\underline{\alpha}}_r$ where $\underline{\underline{\alpha}}_r$ is a symmetric second-order tensor (for an isotropic behaviour, $\underline{\underline{\alpha}}_r = \alpha_r \underline{\underline{\delta}}$) such that a local temperature variation $\Delta T(\underline{x})$ gives rise to the thermal strain $\underline{\underline{\varepsilon}}_r^{\mathrm{th}}(\underline{x}) = \underline{\underline{\alpha}}_r \Delta T(\underline{x})$. An equivalent description of the dilatometric behaviour can be achieved through the thermal coefficients $\underline{\underline{\kappa}}_r = \mathbf{c}_r : \underline{\underline{\alpha}}_r$, such that $\underline{\underline{\sigma}}_r = \mathbf{c}_r : (\underline{\underline{\varepsilon}}_r - \underline{\underline{\varepsilon}}_r^{\mathrm{th}}) = \mathbf{c}_r : \underline{\underline{\varepsilon}}_r - \underline{\underline{\kappa}}_r \Delta T$, where $(-\underline{\underline{\kappa}}_r \Delta T)$ is analogous to an eigenstress.

From now on, we assume the temperature change ΔT to be *uniform* throughout the RVE. The effective dilatometric macroscopic behaviour is defined by the effective coefficients of thermal expansion $\underline{\underline{\alpha}}^{\mathrm{eff}}$ (or by the effective thermal coefficients $\underline{\underline{\kappa}}^{\mathrm{eff}}$) such that the macroscopic thermal strain $\underline{\underline{E}}^{\mathrm{th}}$ reads

$$\begin{aligned} \underline{\underline{E}}^{\mathrm{th}} &= \underline{\underline{\alpha}}^{\mathrm{eff}} \Delta T \\ \text{with } \underline{\underline{E}}^{\mathrm{th}} &= \underline{\underline{E}} - \mathbf{S}^{\mathrm{eff}} : \underline{\underline{\Sigma}} \text{ or } \underline{\underline{\Sigma}} = \mathbf{C}^{\mathrm{eff}} : \underline{\underline{E}} - \underline{\underline{\kappa}}^{\mathrm{eff}} \Delta T \end{aligned} \tag{2.122}$$

From (2.118), we get

$$\begin{aligned} \underline{\underline{E}}^{\mathrm{th}} &= < \mathbf{B}^{\mathrm{T}} : \underline{\underline{\varepsilon}}^{\mathrm{th}} > = < \mathbf{B}^{\mathrm{T}} : \underline{\underline{\alpha}} \Delta T > = \underline{\underline{\alpha}}^{\mathrm{eff}} \Delta T \\ &\Rightarrow \underline{\underline{\alpha}}^{\mathrm{eff}} = < \mathbf{B}^{\mathrm{T}} : \underline{\underline{\alpha}} > = < \underline{\underline{\alpha}} : \mathbf{B} > \end{aligned} \tag{2.123}$$

Similarly, from (2.120) we would have

$$\underline{\underline{\kappa}}^{\mathrm{eff}} = < \mathbf{A}^{\mathrm{T}} : \underline{\underline{\kappa}} > = < \underline{\underline{\kappa}} : \mathbf{A} > \tag{2.124}$$

According to (2.118), these relations read also

$$\begin{aligned} \underline{\underline{\alpha}}^{\mathrm{eff}} &= \sum_r f_r \mathbf{B}_r^{\mathrm{T}} : \underline{\underline{\alpha}}_r = \sum_r f_r \underline{\underline{\alpha}}_r : \mathbf{B}_r \\ \underline{\underline{\kappa}}^{\mathrm{eff}} &= \sum_r f_r \mathbf{A}_r^{\mathrm{T}} : \underline{\underline{\kappa}}_r = \sum_r f_r \underline{\underline{\kappa}}_r : \mathbf{A}_r \end{aligned} \tag{2.125}$$

The main consequence of this analysis is that there is a close connexion between the homogenisation of the elastic and of the dilatometric behaviour of heterogeneous materials since the same concentration tensors $\mathbf{A}_r$ and $\mathbf{B}_r$ are used in both cases. This conclusion is to be reminded in Sect. 2.6 when dealing with estimations.

This connexion is still amplified in the case of two-phase materials. From (2.117) and (2.125), we have, with $f = f_2$ and $f_1 = (1 - f)$

$$\begin{cases} \underline{\underline{\alpha}}^{\text{eff}} = (1-f)\,\underline{\underline{\alpha}}_1 : \mathbf{B}_1 + f\,\underline{\underline{\alpha}}_2 : \mathbf{B}_2 \\ \mathbf{S}^{\text{eff}} = (1-f)\,\mathbf{s}_1 : \mathbf{B}_1 + f\,\mathbf{s}_2 : \mathbf{B}_2 \\ \mathbf{I} = (1-f)\,\mathbf{B}_1 + f\,\mathbf{B}_2 \end{cases} \tag{2.126}$$

so that $\mathbf{B}_1$ and $\mathbf{B}_2$ can be eliminated and yield a relation (the so-called Levin's[23] theorem) between $\underline{\underline{\alpha}}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$ which can be written as

$$\underline{\underline{\alpha}}^{\text{eff}} = <\underline{\underline{\alpha}}> + (\mathbf{S}^{\text{eff}} - <\mathbf{s}>) : (\mathbf{s}_2 - \mathbf{s}_1)^{-1} : (\underline{\underline{\alpha}}_2 - \underline{\underline{\alpha}}_1) \tag{2.127}$$

So, if we measure or bound or estimate the effective compliance, (2.127) yields at once the necessary value or bounds or estimate for the effective coefficients of thermal expansion. Of course, a similar relation holds between $\underline{\underline{\kappa}}^{\text{eff}}$ and $\mathbf{C}^{\text{eff}}$.

2.5.3 Bounds for the Overall Moduli and Compliances

The relations (2.116) are only theoretical definitions for $\mathbf{C}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$ since $\mathbf{c}(\underline{x})$ or $\mathbf{s}(\underline{x})$ are generally not known unambiguously at each point $\underline{x}$. In order to bound $\mathbf{C}^{\text{eff}}$ or $\mathbf{S}^{\text{eff}}$ for all the materials which correspond to a given partial information on the spatial distribution of the phases (or the moduli), we can use (2.107) in a new way: the solution $(\underline{\underline{\sigma}}, \underline{u})$ is no longer unique, but if we can find admissible fields $\underline{\underline{\sigma}}^*$ or $\underline{u}'$ such that $\Phi_1(\underline{\underline{\sigma}}^*)$ or $\Phi_2(\underline{u}')$ can be calculated, despite the limited nature of the available information, then (2.107) yields lower or upper bounds for $\Phi_1(\underline{\underline{\sigma}})$ or $\Phi_2(\underline{u})$ and thus for $\mathbf{C}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$. We do this first by using uniform strain and stress fields to derive the so-called Voigt's and Reuss's bounds.

2.5.3.1 Voigt's Bound (Uniform Strains)

Let $\underline{\underline{\varepsilon}}'(\underline{u}')$ be a kinematically admissible field, with $u'_i = E_{ij}\,x_j$ on ∂V; in the absence of inertial and body forces and with homogeneous strain boundary conditions $\underline{\underline{E}}$, the potential energy theorem states that any solution $\underline{\underline{\varepsilon}}$ minimises the elastic energy:

[23] Valery M. Levin is a Russian Professor of Mechanics at Petrozavodsk State University (Russia).

$$(1/2)\int_V \underline{\underline{\varepsilon}} : \mathbf{c} : \underline{\underline{\varepsilon}}\, \mathrm{d}V \leq (1/2)\int_V \underline{\underline{\varepsilon'}} : \mathbf{c} : \underline{\underline{\varepsilon'}}\, \mathrm{d}V \tag{2.128}$$

In particular, taking $\underline{\underline{\varepsilon'}} = \underline{\underline{E}}$, which is admissible, we have

$$< \underline{\underline{\varepsilon}} : \mathbf{c} : \underline{\underline{\varepsilon}} > \leq \underline{\underline{E}} :< \mathbf{c} >: \underline{\underline{E}} \tag{2.129}$$

Now $< \underline{\underline{\varepsilon}} : \mathbf{c} : \underline{\underline{\varepsilon}} >= \underline{\underline{E}} : \mathbf{C}^{\mathrm{eff}} : \underline{\underline{E}}$ whence

$$\underline{\underline{E}} : (< \mathbf{c} > -\mathbf{C}^{\mathrm{eff}}) : \underline{\underline{E}} \geq 0, \ \forall \underline{\underline{E}} \tag{2.130}$$

2.5.3.2 Reuss's Bound (Uniform Stresses)

Let $\underline{\underline{\sigma}}^*$ be a statically admissible stress field for the same problem, with homogeneous strain boundary conditions again. The complementary energy theorem gives

$$\frac{1}{2}\int_V \underline{\underline{\sigma}} : \mathbf{s} : \underline{\underline{\sigma}}\, \mathrm{d}V - \int_{\partial V} (\underline{\underline{\sigma}}.\underline{n}).\underline{u}^{\mathrm{g}} \mathrm{d}S \leq \frac{1}{2}\int_V \underline{\underline{\sigma}}^* : \mathbf{s} : \underline{\underline{\sigma}}^* \mathrm{d}V - \int_{\partial V} (\underline{\underline{\sigma}}^*.\underline{n}).\underline{u}^{\mathrm{g}} \mathrm{d}S \tag{2.131}$$

Now

$$\int_V \sigma_{ij} n_j u_i^{\mathrm{g}}\, \mathrm{d}S = \int_V (\sigma_{ij} E_{ik} x_k)_{,j}\, \mathrm{d}V = \int_V \sigma_{ij} E_{ij}\, \mathrm{d}V = V \Sigma_{ij} E_{ij} \tag{2.132}$$

with $\underline{\underline{\Sigma}} =< \underline{\underline{\sigma}} >$. If we choose $\underline{\underline{\sigma}}^* = \underline{\underline{\Sigma}}$ we have, similarly,

$$\int_{\partial V} \sigma_{ij}^* n_j u_i^{\mathrm{g}}\, \mathrm{d}S = \Sigma_{ij} \int_{\partial V} n_j u_i^{\mathrm{g}}\, \mathrm{d}S = V \Sigma_{ij} E_{ij} \tag{2.133}$$

from which it follows that $\int_V \underline{\underline{\sigma}} : \mathbf{s} : \underline{\underline{\sigma}}\, \mathrm{d}V \leq \int_V \underline{\underline{\Sigma}} : \mathbf{s} : \underline{\underline{\Sigma}}\, \mathrm{d}V$ that is, $< \underline{\underline{\sigma}} : \mathbf{s} : \underline{\underline{\sigma}} > \leq \underline{\underline{\Sigma}} :< \mathbf{s} >: \underline{\underline{\Sigma}}$.

Therefore, since $< \underline{\underline{\sigma}} : \mathbf{s} : \underline{\underline{\sigma}} >= \underline{\underline{\Sigma}} : \mathbf{S}^{\mathrm{eff}} : \underline{\underline{\Sigma}}$, we have finally

$$\underline{\underline{\Sigma}} : (< \mathbf{s} > -\mathbf{S}^{\mathrm{eff}}) : \underline{\underline{\Sigma}} \geq 0, \quad \forall \underline{\underline{\Sigma}} \tag{2.134}$$

Symbolically, the two inequalities can be summarised as

$$< \mathbf{c}^{-1} >^{-1} \leq (\mathbf{S}^{\mathrm{eff}})^{-1} = \mathbf{C}^{\mathrm{eff}} \leq < \mathbf{c} > \tag{2.135}$$

2.5.3.3 Voigt's and Reuss's Bounds for Scalar Moduli

The inequalities of Voigt and Reuss concern only the quadratic forms associated with $\mathbf{S}^{\text{eff}}$ and $\mathbf{C}^{\text{eff}}$; by choosing the fields $\underline{\underline{E}}$ and $\underline{\underline{\Sigma}}$ suitably we can deduce from them inequalities for the components of $\mathbf{S}^{\text{eff}}$ and $\mathbf{C}^{\text{eff}}$, or for combinations of these. For example, for an isotropic composite with isotropic constituents, with $k = \lambda + 2\mu/3$ and $e_{ij} = \varepsilon_{ij} - (1/3)\varepsilon_{kk}\delta_{ij}$, we have

$$\begin{aligned} c_{ijkl}\varepsilon_{ij}\varepsilon_{kl} &= \left[(k - 2\mu/3)\varepsilon_{kk}\delta_{ij} + 2\mu\varepsilon_{ij}\right]\varepsilon_{ij} \\ &= (k\varepsilon_{kk}\delta_{ij} + 2\mu e_{ij})(e_{ij} + (1/3)\varepsilon_{ll}\delta_{ij}) = 2\mu e_{ij}e_{ij} + k\varepsilon_{kk}^2 \end{aligned} \tag{2.136}$$

Hence

$$2 < \mu > e_{ij}e_{ij} + < k > \varepsilon_{kk}^2 \;\geq\; 2\mu^{eff} e_{ij}e_{ij} + k^{eff}\varepsilon_{kk}^2, \quad \forall e_{ij}, \; \forall \varepsilon_{kk} \tag{2.137}$$

implying $< k > \geq k^{\text{eff}}$ and $< \mu > \geq \mu^{\text{eff}}$ (Voigt's bound for the bulk and shear moduli).

On the other hand, with $s_{ij} = \sigma_{ij} - (1/3)\sigma_{kk}\delta_{ij}$ (deviator) and

$$\begin{aligned} s_{ijkl}\sigma_{kl} = \varepsilon_{ij} &= \frac{1+\nu}{E}\sigma_{ij} - \frac{\nu}{E}\sigma_{kk}\delta_{ij} \\ &= \frac{1}{2\mu}\sigma_{ij} - \frac{3k - 2\mu}{18k\mu}\sigma_{kk}\delta_{ij} \end{aligned} \tag{2.138}$$

we have

$$\begin{aligned} s_{ijkl}\sigma_{ij}\sigma_{kl} &= \frac{1}{2\mu}\left(\sigma_{ij} - \frac{3k - 2\mu}{9k}\sigma_{kk}\delta_{ij}\right)\sigma_{ij} \\ &= \frac{1}{2\mu}\left(s_{ij} + \frac{2\mu}{9k}\sigma_{kk}\delta_{ij}\right)\left(s_{ij} + \frac{1}{3}\sigma_{ll}\delta_{ij}\right) \\ &= (1/2\mu)s_{ij}s_{ij} + (1/9k)\sigma_{kk}^2 \end{aligned} \tag{2.139}$$

Hence

$$\begin{aligned} &< 1/2\mu > s_{ij}s_{ij} + (1/9) < 1/k > \sigma_{kk}^2 \\ &\quad \geq \; (1/2\mu^{\text{eff}})s_{ij}s_{ij} + (1/9k^{\text{eff}})\sigma_{kk}^2, \quad \forall s_{ij}, \; \forall \sigma_{kk} \end{aligned} \tag{2.140}$$

implying $<1/\mu> \geq 1/\mu^{\text{eff}}$, $<1/k> \geq 1/k^{\text{eff}}$ (Reuss's bound for the bulk and shear moduli).[24]

[24] Note that the above calculation could have been strongly shortened by the use of the fourth-order unit tensors $\mathbf{J}$ and $\mathbf{K}$ defined by $\mathrm{J}_{ijkl} = (1/3)\delta_{ij}\delta_{kl}$ and $\mathbf{K} + \mathbf{J} = \mathbf{I}$. They are such that $\mathbf{c} =$

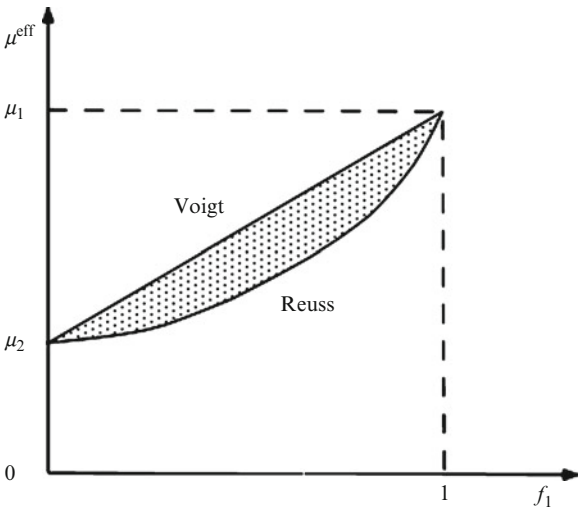


Fig. 2.12 The Hill domain for the shear modulus of an isotropic two-phase composite, as a function of the volume fraction of one component

For a two-phase composite (volume fractions f_1, f_2 with $f_1 + f_2 = 1$), one gets

$$\begin{cases} 1/(f_1/\mu_1 + f_2/\mu_2) \leq \mu^{\text{eff}} \leq f_1\mu_1 + f_2\mu_2 \\ 1/(f_1/k_1 + f_2/k_2) \leq k^{\text{eff}} \leq f_1k_1 + f_2k_2 \end{cases} \tag{2.141}$$

We thus have upper and lower bounds for the effective moduli; together they define the *Hill domain* (see footnote 20 of Chap. 1), shown in Fig. 2.12 for μ^{eff}.

Note that the "laws of mixtures", which estimate any overall parameter as the mean value of the corresponding local ones, are only valid here as bounding properties for the bulk and shear moduli. This is not even true for other elastic constants such as the Young modulus and the Poisson ratio and it is completely wrong for anisotropic elasticity.

2.5.3.4 Higher Order Bounds

The calculation of $<\mathbf{c}>$ or $<\mathbf{s}>$ in (2.135) requires only a knowledge of the volume fractions of the phases; it follows that if this is the only information we have we cannot set sharper bounds to the effective moduli than those given by these inequalities. In other words, if all we know for an elastic composite is

---

$2\mu\mathbf{K} + 3\,k\mathbf{J}$ and $\mathbf{s} = (1/2\,\mu)\mathbf{K} + (1/3\,\text{k})\mathbf{J}$. When acting on a symmetric second-order tensor such as ε_{ij}, they yield directly $\text{K}_{ijkl}\varepsilon_{kl} = e_{ij}$ (deviator) and $\text{J}_{ijkl}\varepsilon_{kl} = (1/3)\varepsilon_{kk}\delta_{ij}$ (see Exercises of Volume III for more applications).

the volume fraction of each constituent phase (and, of course, the elastic moduli for each phase) then the values of the effective moduli can lie *anywhere* within the Voigt/Reuss bounds: only more information about the spatial distribution of the phases will enable us to set closer bounds. In particular, we show later that the Hashin-Shtrikman bounds are sharper, but they hold in more restrictive conditions. This question is discussed later (Sect. 2.8).

2.6 Elasticity of Heterogeneous Materials: Estimates for the Overall Moduli and Compliances

Instead of looking for bounds for the overall moduli, it can be found easier to derive *estimates* by introducing additional information, assumptions or approximations so as to be able to propose possible solutions to the general problem which, as stressed before, has an infinite number of solutions. From (2.84), estimates for $\mathbf{C}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$ can be derived straightforwardly from estimates for $\mathbf{A}_r$ and $\mathbf{B}_r$, say $\mathbf{A}_r^{\text{est}}$ and $\mathbf{B}_r^{\text{est}}$, which yield estimates for the strain and stress averages per phase

$$\begin{aligned}
&\underline{\underline{\varepsilon}}_r^{\text{est}} = <\underline{\underline{\varepsilon}}^{\text{est}}>_{(r)} = \mathbf{A}_r^{\text{est}} : \underline{\underline{E}} \text{ with } \sum_{r=1}^{n} f_r \mathbf{A}_r^{\text{est}} = \mathbf{I} \\
&\underline{\underline{\sigma}}_r^{\text{est}} = <\underline{\underline{\sigma}}^{\text{est}}>_{(r)} = \mathbf{B}_r^{\text{est}} : \underline{\underline{\Sigma}} \text{ with } \sum_{r=1}^{n} f_r \mathbf{B}_r^{\text{est}} = \mathbf{I}
\end{aligned} \tag{2.142}$$

through the relations

$$\begin{cases}
\mathbf{C}^{\text{est}} = \sum\limits_{r=1}^{n} f_r \mathbf{c}_r : \mathbf{A}_r^{\text{est}} \quad \text{with} \quad \sum\limits_{r=1}^{n} f_r \mathbf{A}_r^{\text{est}} = \mathbf{I} \\
\mathbf{S}^{\text{est}} = \sum\limits_{r=1}^{n} f_r \mathbf{s}_r : \mathbf{B}_r^{\text{est}} \quad \text{with} \quad \sum\limits_{r=1}^{n} f_r \mathbf{B}_r^{\text{est}} = \mathbf{I}
\end{cases} \tag{2.143}$$

As already mentioned in Sect. 2.5.2.3, estimates for the overall elastic moduli or compliances (through $\mathbf{A}_r^{\text{est}}$ or $\mathbf{B}_r^{\text{est}}$) yield at once estimates for the overall thermal coefficients which, according to (2.90), must read

$$\begin{cases}
\underline{\underline{\alpha}}^{\text{est}} = \sum\limits_{r} f_r \mathbf{B}_r^{\text{estT}} : \underline{\underline{\alpha}}_r = \sum\limits_{r} f_r \underline{\underline{\alpha}}_r : \mathbf{B}_r^{\text{est}} \\
\underline{\underline{\kappa}}^{\text{est}} = \sum\limits_{r} f_r \mathbf{A}_r^{\text{estT}} : \underline{\underline{\kappa}}_r = \sum\limits_{r} f_r \underline{\underline{\kappa}}_r : \mathbf{A}_r^{\text{est}}
\end{cases} \tag{2.144}$$

where $\mathbf{A}_r^{\text{est}}$ and $\mathbf{B}_r^{\text{est}}$ have to be *the same* as in (2.143).

The simplest choice for $\underline{\underline{\varepsilon}}_r^{\text{est}}$ or $\underline{\underline{\sigma}}_r^{\text{est}}$ is to assume uniform average strain or stress per phase, namely $\mathbf{A}_r^{\text{est}} = \mathbf{I}, \forall r$ or $\mathbf{B}_r^{\text{est}} = \mathbf{I}, \forall r$, which leads to the estimates $\mathbf{C}^{\text{est}} =< \mathbf{c} >$ or $\mathbf{S}^{\text{est}} =< \mathbf{s} >$, *i.e.*, to the "direct" or "inverse" laws of

mixtures, respectively. Despite the advantage of simplicity, which is connected with convenient images of "parallel" and "series" assemblages, these empirical laws exhibit obvious drawbacks: they predict isotropic overall moduli or compliances for isotropic constituents whatever their spatial distribution, contrary to what is observed *e.g.* for long fibre reinforced composites or laminates; they either lead to an overall rigid behaviour as soon as one constituent is rigid, whatever its volume fraction, or to vanishing overall moduli for any porous material, whatever its porosity, etc.

This shows that, as well as for bounds, improved estimates are needed. In both cases, the Green's[25] techniques have proved to be the key for further progress. They are briefly described in what follows through the resolution of the inclusion problem.

2.7 The Inclusion Problem

This is a fundamental problem in the mechanics of heterogeneous materials because it concerns the basic situation of the heterogeneity between a zone in the material (an inclusion) and its environment (the matrix). Further, it is important in itself because its study enables us to analyse particular situations such as the presence of precipitates, voids and pores, and to do so for a range of scales.

We shall follow Eshelby[26] (1957) and treat the case of an elastic inclusion in an infinite elastic matrix, dealing in detail with the situation in which the inclusion is ellipsoidal and the matrix elasticity is isotropic.

2.7.1 Inclusion with Uniform Stress-Free Strain in a Load-Free Infinite Matrix

Let the prescribed uniform "stress-free strain" in the inclusion I be $\underline{\underline{\varepsilon}}^{\mathrm{F}}$ and the elastic matrix with uniform moduli $\mathbf{C}$ be load-free at infinity; by *stress-free strain* (or *eigenstrain*) we mean a deformation which, if it were present in an isolated element, would not be accompanied by any stress – it could be due, for example, to thermal effects or to a phase change. We can write $\underline{\underline{\varepsilon}}^{\mathrm{F}}(\underline{x}) = \underline{\underline{\varepsilon}}^{\mathrm{F}}\delta_0(\mathrm{I})$ where $\delta_0(\mathrm{I})$ is the characteristic function for the inclusion I; hence $\sigma_{ij} = C_{ijkl}[u_{l,k} - \varepsilon^{\mathrm{F}}_{kl}\ \delta_0(\mathrm{I})]$ and by the equilibrium equations

$$C_{ijkl}u_{l,kj} + C_{ijkl}\varepsilon^{\mathrm{F}}_{kl}n_j\delta(S) = 0 \tag{2.145}$$

[25] George Green (1793–1841) was a British mathematician.

[26] John Douglas Eshelby (1916–1988) was a British scientist.

where $\delta(S)$ is the Dirac function applied to the interface S between the inclusion and the matrix, at which the unit outward normal is $\underline{n}$. These are the equations for the equilibrium of a homogeneous material acted on at the surface S by a force distribution

$$f_i = C_{ijkl}\varepsilon^{\mathrm{F}}_{kl} n_j \tag{2.146}$$

Now let G_{ij} be the Green tensor for the infinite medium with elastic moduli $\mathbf{C}$; this is solution of the equation

$$C_{ijkl}G_{km,lj} + \delta_{im}\delta(\underline{r}, \underline{r'}) = 0 \tag{2.147}$$

The Green tensor gives the displacement at a point $\underline{x}$ due to a force $\underline{F}(\underline{x'})$ applied at $\underline{x'}$ as $u_i(\underline{x'}) = G_{ij}\ (\underline{x}, \underline{x'})\ F_j(\underline{x'})$. For an infinite homogeneous medium it depends only on the distance $\rho = |\underline{x} - \underline{x'}|$, and in the present case we get

$$u_i(\underline{x}) = \int_S G_{ij} C_{jklm}\varepsilon^{\mathrm{F}}_{lm} n_k \mathrm{d}S' = \int_{\mathrm{I}} G_{ij,k'} C_{jklm}\varepsilon^{\mathrm{F}}_{lm} \mathrm{d}V' \quad (G_{ij,k'} = \partial G_{ij}/\partial x'_k) \tag{2.148}$$

where the notations S' and V' show that the integration variable is $\underline{x'}$. Hence

$$u_{i,n} = C_{jklm}\varepsilon^{\mathrm{F}}_{lm} \int_{\mathrm{I}} G_{ij,k'n}\ \mathrm{d}V' \tag{2.149}$$

and by symmetrisation with respect to i and n

$$\begin{aligned} \varepsilon_{in}(\underline{x}) &= \left(\int_{\mathrm{I}} G_{ij,k'n_{(in)}}(\underline{x} - \underline{x'}) \mathrm{d}V' \right) C_{jklm}\varepsilon^{\mathrm{F}}_{lm} \\ &= -\left(\int_{\mathrm{I}} G_{ij}(\underline{x} - \underline{x'}) \mathrm{d}V' \right)_{,kn_{(in)(jk)}} C_{jklm}\varepsilon^{\mathrm{F}}_{lm} \end{aligned} \tag{2.150}$$

Omitting some mathematical difficulties arising from the singularity of the second-order derivatives of the Green tensor at $\underline{x} = \underline{x'}$ and from the infinite extent of the body, we define the so-called "modified strain Green operator" $\tilde{\mathbf{\Gamma}}$ by

$$\Gamma_{injk} = G_{ij,k'n_{(in)(jk)}} \tag{2.151}$$

so that we have

$$\varepsilon_{in}(\underline{x}) = P_{injk}(\underline{x})\ C_{jklm}\varepsilon^{\mathrm{F}}_{lm} \quad \text{where} \quad P_{injk}(\underline{x}) = \int_{\mathrm{I}} \Gamma_{injk}(\underline{x} - \underline{x'})\ \mathrm{d}V' \tag{2.152}$$

Note: This result can be obtained by working through the following sequence of fictitious operations, shown in Fig. 2.13.

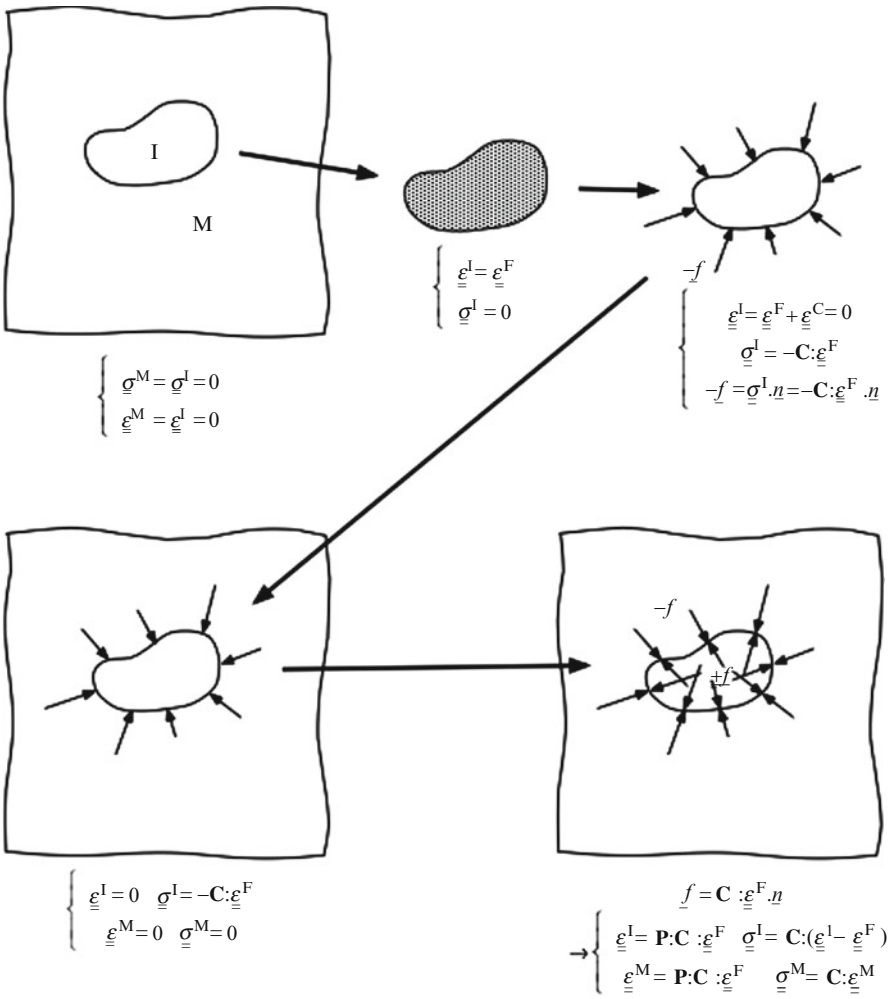


Fig. 2.13 The Eshelby procedure for solving the inclusion problem. 1 Extract the inclusion out of the matrix; 2 Let the strain in the inclusion be the eigenstrain; 3 Apply distributed forces on the surface of the inclusion to bring it back to its initial shape; 4 Insert back the inclusion in the matrix; 5 Apply distributed forces on the surface of the inclusion to cancel the ones applied at step 3

At this stage, the fourth-order tensor **P** is expected to depend on $\underline{x}$. We now accept the main following result due to Eshelby (see Volume III for the demonstration):

if I is ellipsoidal and $\underline{x} \in \mathbf{I}$, then **P** does not depend on $\underline{x}$.

It follows that, in this case, $\underline{\underline{\varepsilon}}^{I}$ is uniform and can be written

$$\underline{\underline{\varepsilon}}^{\mathrm{I}} = \mathbf{P} : \mathbf{C} : \underline{\underline{\varepsilon}}^{\mathrm{F}} = \mathbf{S}^{\mathrm{Esh}} : \underline{\underline{\varepsilon}}^{\mathrm{F}} \tag{2.153}$$

where the Eshelby tensor $\mathbf{S}^{\mathrm{Esh}}$ only depends on the moduli $\mathbf{C}$, the shape (aspect ratios) and orientation of the ellipsoidal inclusion. Note that, contrary to $\mathbf{S}^{\mathrm{Esh}}$, the so-called Hill tensor $\mathbf{P} = \mathbf{S}^{\mathrm{Esh}} : \mathbf{S}$ has the advantage of exhibiting diagonal symmetry.

2.7.2 The Case of Isotropic Elasticity

The Green tensor has no simple analytical expression except for the case of isotropic elasticity; we shall consider only this. We have

$$G_{ij}(\underline{x}, \underline{x}') = G_{ij}(|\underline{x} - \underline{x}'|) = G_{ij}(\rho) = \frac{1}{4\pi\mu}\left(\frac{\delta_{ij}}{\rho} - \frac{\rho_{,ij}}{4(1-\nu)}\right) \tag{2.154}$$

with $\rho_{,ij} = \frac{1}{\rho}(\delta_{ij} - e_i e_j)$ and $\underline{e} = \frac{\underline{x}-\underline{x}'}{\rho}$. Since $f_j = 2\mu\left(\varepsilon^{\mathrm{F}}_{jk} + \frac{\nu}{1-2\nu}\varepsilon^{\mathrm{F}}_{ll}\delta_{jk}\right)n_k$ and $G_{ij,k'} = -G_{ij,k}$, we get

$$u_i(\underline{x}) = -2\mu\left(\varepsilon^{\mathrm{F}}_{jk} + \frac{\nu}{1-2\nu}\varepsilon^{\mathrm{F}}_{ll}\delta_{jk}\right)\left[\int_{\mathrm{I}} G_{ij}\,\mathrm{d}V'\right]_{,k} \tag{2.155}$$

Putting $\varphi(\underline{x}) = \int_{\mathrm{I}} \frac{\mathrm{d}V'}{\rho}$, $\psi(\underline{x}) = \int_{\mathrm{I}} \rho\,\mathrm{d}V'$ we have $\Delta\psi = 2\varphi$ and $\Delta\varphi = 0$ in I, $\Delta\varphi = 4\pi$ outside I. Thus

$$\begin{aligned} u_i(\underline{x}) &= -2\mu\left(\varepsilon^{\mathrm{F}}_{jk} + \frac{\nu}{1-2\nu}\varepsilon^{\mathrm{F}}_{ll}\delta_{jk}\right)\left(\frac{\delta_{ij}}{4\pi\mu}\varphi_{,k} - \frac{\psi_{,ijk}}{16\pi\mu(1-\nu)}\right) \\ &= -\frac{1}{2\pi}\varepsilon^{\mathrm{F}}_{ik}\varphi_{,k} - \frac{\nu}{4\pi(1-\nu)}\varepsilon^{\mathrm{F}}_{ll}\varphi_{,i} + \frac{1}{8\pi(1-\nu)}\varepsilon^{\mathrm{F}}_{jk}\psi_{,ijk} \end{aligned} \tag{2.156}$$

In the rest of this discussion we shall assume that the inclusion is *ellipsoidal*. It can be shown that u_i is then linear inside I and that $\underline{\underline{\sigma}}$ and $\underline{\underline{\varepsilon}}$ are therefore constant there; we have, in I, $\varphi = \frac{1}{2}[(a^2-x^2)I_a + (b^2-y^2)I_b + (c^2-z^2)I_c]$ where *a, b, c* are the semi-axes of the ellipsoid and I_a, I_b, I_c are constants (elliptic integrals) depending on the geometry of the inclusion. The resulting deformation within the inclusion can then be written $\varepsilon^{\mathrm{I}}_{ij} = S^{\mathrm{Esh}}_{ijkl}\,\varepsilon^{\mathrm{F}}_{kl}$ where $\mathbf{S}^{\mathrm{Esh}}$ is the (constant) Eshelby tensor, of which the only non-zero components are those with indices 1111, 1122, 2211, 1133, 3311, 2222, 2233, 3322, 3333, 1212, 2323, 1313.

We thus have

$$\sigma^{\mathrm{I}}_{ij} = 2\mu\left[S^{\mathrm{Esh}}_{ijkl}\,\varepsilon^{\mathrm{F}}_{kl} - \varepsilon^{\mathrm{F}}_{ij} + \frac{\nu}{1-2\nu}\left(S^{\mathrm{Esh}}_{mmkl}\,\varepsilon^{\mathrm{F}}_{kl} - \varepsilon^{\mathrm{F}}_{mm}\right)\delta_{ij}\right] \tag{2.157}$$

Table 2.3 Eshelby's tensor for some extreme shapes of ellipsoidal inclusion (isotropic elasticity)

Elliptic cylinder$(c \to \infty)$

$$S^{\text{Esh}}_{1111} = \frac{1}{2(1-\nu)}\left[\frac{b^2+2ab}{(a+b)^2} + \frac{b(1-2\nu)}{a+b}\right]$$

$$S^{\text{Esh}}_{2222} = \frac{1}{2(1-\nu)}\left[\frac{a^2+2ab}{(a+b)^2} + \frac{a(1-2\nu)}{a+b}\right]$$

$$S^{\text{Esh}}_{1122} = \frac{1}{2(1-\nu)}\left[\frac{b^2}{(a+b)^2} - \frac{b(1-2\nu)}{a+b}\right]$$

$$S^{\text{Esh}}_{2211} = \frac{1}{2(1-\nu)}\left[\frac{a^2}{(a+b)^2} - \frac{a(1-2\nu)}{a+b}\right]$$

$$S^{\text{Esh}}_{1212} = \frac{1}{2(1-\nu)}\left[\frac{a^2+b^2}{2(a+b)^2} + \frac{1-2\nu}{2}\right]$$

$$S^{\text{Esh}}_{1133} = \frac{1}{2(1-\nu)}\frac{2b\nu}{a+b} \quad S^{\text{Esh}}_{2233} = \frac{1}{2(1-\nu)}\frac{2a\nu}{a+b}$$

$$S^{\text{Esh}}_{2323} = \frac{a}{2(a+b)} \quad S^{\text{Esh}}_{3131} = \frac{b}{2(a+b)}$$

$$S^{\text{Esh}}_{3311} = S^{\text{Esh}}_{3322} = S^{\text{Esh}}_{3333} = 0$$

Penny shape$(a = b \gg c)$

$$S^{\text{Esh}}_{1111} = S^{\text{Esh}}_{2222} = \frac{\pi c(13-8\nu)}{32a(1-\nu)}$$

$$S^{\text{Esh}}_{1122} = S^{\text{Esh}}_{2211} = \frac{\pi c(8\nu-1)}{32a(1-\nu)}$$

$$S^{\text{Esh}}_{1212} = \frac{\pi c(7-8\nu)}{32a(1-\nu)} \; S^{\text{Esh}}_{3333} = 1 - \frac{\pi c(1-2\nu)}{4a(1-\nu)}$$

$$S^{\text{Esh}}_{1133} = S^{\text{Esh}}_{2233} = \frac{\pi c(2\nu-1)}{8a(1-\nu)}$$

$$S^{\text{Esh}}_{3311} = S^{\text{Esh}}_{3322} = \frac{\nu}{1-\nu}\left(1 - \frac{\pi c(4\nu+1)}{8\nu a}\right)$$

$$S^{\text{Esh}}_{3131} = S^{\text{Esh}}_{2323} = \frac{1}{2}\left(1 + \frac{\pi c(\nu-2)}{4a(1-\nu)}\right)$$

$$\text{when } c \to 0 : S^{\text{Esh}}_{3131} = S^{\text{Esh}}_{2323} = \frac{1}{2}$$

$$S^{\text{Esh}}_{3311} = S^{\text{Esh}}_{3322} = \frac{\nu}{(1-\nu)}$$

$$S^{\text{Esh}}_{3333} = 1 \quad \text{all other } S^{\text{Esh}}_{ijkl} = 0$$

Other shapes – cylinders, plates, etc. – can be treated as special or limiting cases of the ellipsoid; here we give the results for a sphere. Please refer to Table 2.3 for partial results for other shapes (Mura 1987; Nemat-Nasser and Hori 1999).

Spherical inclusion For a sphere of radius a (Eshelby (1961)):

$$\varphi(r) = \frac{4}{3}\pi a^2\left(\frac{3}{2} - \frac{r^2}{2a^2}\right), \quad r < a; \quad \varphi(r) = \frac{4}{3}\frac{\pi a^3}{r}, \quad r > a$$

$$\psi(r) = \frac{4}{3}\pi a^4\left(\frac{3}{4} + \frac{r^2}{2a^2} - \frac{r^4}{20a^4}\right), \quad r < a; \quad \psi(r) = \frac{4}{3}\pi a^4\left(\frac{a}{5r} + \frac{r}{a}\right), \quad r > a$$

from which we find

$$\varepsilon^{\mathrm{I}}_{il} = \frac{2(4-5\nu)}{15(1-\nu)}\varepsilon^{\mathrm{F}}_{il} + \frac{5\nu-1}{15(1-\nu)}\varepsilon^{\mathrm{F}}_{kk}\delta_{il} \tag{2.158}$$

Thus for the spherical (hydrostatic) and deviatoric parts of $\underline{\underline{\varepsilon}}^{\mathrm{I}}$ and $\underline{\underline{\varepsilon}}^{\mathrm{F}}$

$$\varepsilon^{\mathrm{I}}_{kk} = \alpha\varepsilon^{\mathrm{F}}_{kk}, \quad \alpha = \frac{1+\nu}{3(1-\nu)}; \quad e^{\mathrm{I}}_{ij} = \beta e^{\mathrm{F}}_{ij}, \quad \beta = \frac{2(4-5\nu)}{15(1-\nu)} \tag{2.159}$$

Finally

$$\varepsilon^{\mathrm{I}}_{ij} = \frac{\alpha-\beta}{3}\varepsilon^{\mathrm{F}}_{kk}\delta_{ij} + \beta\varepsilon^{\mathrm{F}}_{ij}, \quad \sigma^{I}_{ij} = -2\mu\left[(1-\beta)\varepsilon^{\mathrm{F}}_{ij} + \frac{\beta+2\alpha-1}{3}\varepsilon^{\mathrm{F}}_{kk}\delta_{ij}\right] \tag{2.160}$$

Comments: – when ν varies from -1 to .5, β decreases from 3/5 to 2/5 and α increases from -1 to 1.

- the solution for stresses and strains is independent of the radius of the sphere. Far from the inclusion, $\varphi \approx 1/r$, $\psi \approx r$, giving $|u_i| \approx 1/r^2$, $|\sigma_{ij}| \approx |\varepsilon_{ij}| \approx 1/r^3$.
- the solution just obtained is valid only for perfect contact between the matrix and the inclusion, without any sliding or debonding.
- if $\underline{\underline{\varepsilon}}^{\mathrm{F}}$ is an isochoric (*e.g.* plastic) strain, then $\underline{\underline{\sigma}}^{\mathrm{I}} \approx -\mu\,\underline{\underline{\varepsilon}}^{\mathrm{F}}$ since $\beta \approx 1/2$.

2.7.3 Other Problems Concerning Ellipsoidal Inclusions

Many other problems can be treated as applications of the solution just obtained, in particular by using the method of the "fictitious equivalent inclusion" with the fictitious eigenstrain $\underline{\underline{\varepsilon}}^{\mathrm{F}*}$ (Eshelby 1957). We now run quickly over a selection.

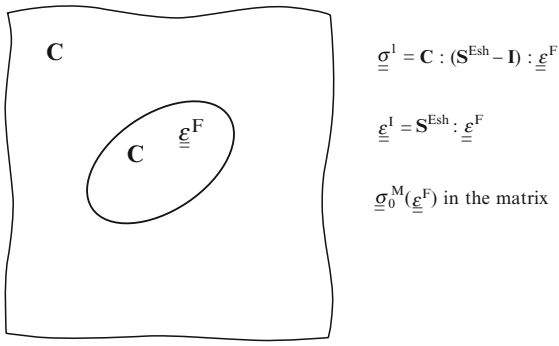


Fig. 2.14 Ellipsoidal inclusion with uniform stress-free strain in an infinite unloaded matrix

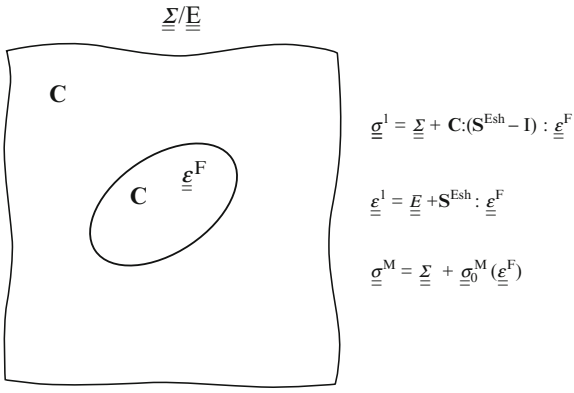


Fig. 2.15 Ellipsoidal inclusion with uniform stress-free strain in an infinite matrix uniformly loaded at infinity

2.7.3.1 The Reference Problem

See Fig. 2.14.

2.7.3.2 Infinite Matrix Uniformly Loaded at Infinity

See Fig. 2.15.

2.7.3.3 Cavity in a Uniformly Loaded Infinite Matrix

See Fig. 2.16.

2.7.3.4 Elastic Heterogeneity

In Fig. 2.17, $\underline{\underline{\varepsilon}}$ and $\underline{\underline{\sigma}}$ must be the same in the two ellipsoids with, at the same time,

$$\underline{\underline{\sigma}}^{H} = \underline{\underline{\sigma}}^{I} = \underline{\underline{\Sigma}} + \mathbf{C} : (\mathbf{S}^{Esh} - \mathbf{I}) : \underline{\underline{\varepsilon}}^{F*}, \quad \underline{\underline{\varepsilon}}^{H} = \underline{\underline{\varepsilon}}^{I} = \underline{\underline{E}} + \mathbf{S}^{Esh} : \underline{\underline{\varepsilon}}^{F*} \tag{2.161}$$

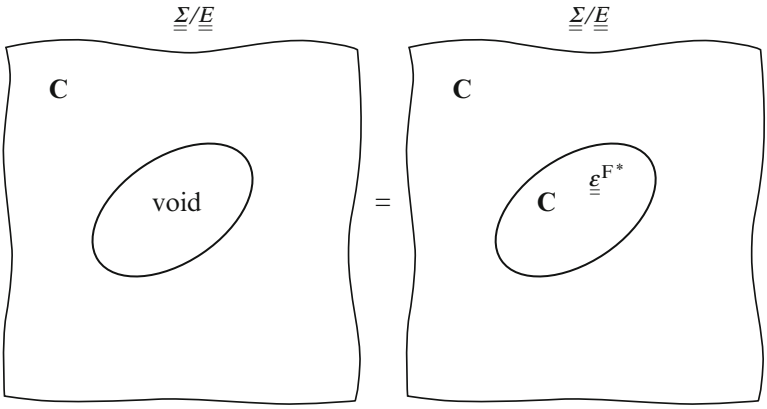


Fig. 2.16 Cavity in a uniformly loaded infinite matrix

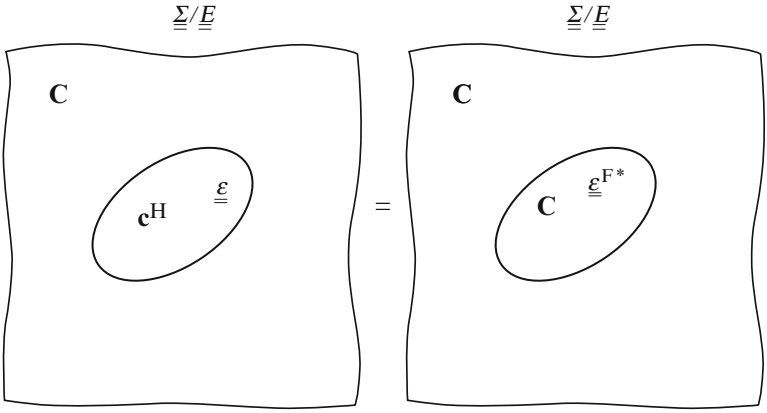


Fig. 2.17 Elastic heterogeneity

as in Sect. 2.7.3.2 and

$$\underline{\underline{\sigma}}^{\mathrm{H}} = \mathbf{c}^{\mathrm{H}} : \underline{\underline{\varepsilon}}^{\mathrm{H}}, i.e., \underline{\underline{\Sigma}} + \mathbf{C} : (\mathbf{S}^{\mathrm{Esh}} - \mathbf{I}) : \underline{\underline{\varepsilon}}^{\mathrm{F*}} = \mathbf{c}^{\mathrm{H}} : (\underline{\underline{E}} + \mathbf{S}^{\mathrm{Esh}} : \underline{\underline{\varepsilon}}^{\mathrm{F*}}). \tag{2.162}$$

Thus, since $\underline{\underline{\Sigma}} = \mathbf{C} : \underline{\underline{E}}$, we must have

$$\underline{\underline{\varepsilon}}^{\mathrm{H}} = \underline{\underline{E}} + \mathbf{S}^{\mathrm{Esh}} : [(\mathbf{c}^{\mathrm{H}} - \mathbf{C}) : \mathbf{S}^{\mathrm{Esh}} + \mathbf{C}]^{-1} : (\mathbf{C} - \mathbf{c}^{\mathrm{H}}) : \underline{\underline{E}} \tag{2.163}$$

giving finally after some algebra and with the definition (2.153)

$$\underline{\underline{\varepsilon}}^{\mathrm{H}} = (\mathbf{I} + \mathbf{P} : \delta\mathbf{c}^{\mathrm{H}})^{-1} : \underline{\underline{E}} \quad \text{with } \mathbf{P} = \mathbf{S}^{\mathrm{Esh}} : \mathbf{C}^{-1}, \quad \delta\mathbf{c}^{\mathrm{H}} = \mathbf{c}^{\mathrm{H}} - \mathbf{C} \tag{2.164}$$

For a sphere, the strain deviator $\underline{\underline{e}}^{\mathrm{H}}_{\mathrm{Sph}}$ and the trace $\theta^{\mathrm{H}}_{\mathrm{Sph}}$ are given by

$$\underline{\underline{e}}^{\mathrm{H}}_{\mathrm{Sph}} = \left(1 + \frac{\beta\,\delta\mu^{\mathrm{H}}}{\mu}\right)^{-1} \underline{\underline{\mathcal{E}}} \qquad \theta^{\mathrm{H}}_{\mathrm{Sph}} = \left(1 + \frac{\alpha\,\delta k^{\mathrm{H}}}{k}\right)^{-1} \Theta$$
$$\text{with } \underline{\underline{E}} = \underline{\underline{\mathcal{E}}} + (\Theta/3)\,\underline{\underline{\delta}} \tag{2.165}$$

When $\nu = \nu^{\mathrm{H}} = ½$, we have

$$\varepsilon^{\mathrm{H}}_{ij} = \frac{5\mu}{3\mu + 2\mu^{\mathrm{H}}} E_{ij}, \qquad \sigma^{\mathrm{H}}_{ij} = \frac{5\mu}{3\mu + 2\mu^{\mathrm{H}}} \Sigma_{ij} \tag{2.166}$$

2.7.3.5 Elastic Heterogeneity with Stress-Free Strain ("Heterogeneous Inclusion")

Using the same method we write (see Fig. 2.18)

$$\begin{aligned} \underline{\underline{\sigma}}^{\mathrm{HI}} &= \underline{\underline{\Sigma}} + \mathbf{C} : (\mathbf{S}^{\mathrm{Esh}} - \mathbf{I}) : \underline{\underline{\varepsilon}}^{\mathrm{F}*} \\ \underline{\underline{\varepsilon}}^{\mathrm{HI}} &= \underline{\underline{E}} + \mathbf{S}^{\mathrm{Esh}} : \underline{\underline{\varepsilon}}^{\mathrm{F}*} \\ \underline{\underline{\sigma}}^{\mathrm{HI}} &= \mathbf{c}^{\mathrm{HI}} : (\underline{\underline{\varepsilon}}^{\mathrm{HI}} - \underline{\underline{\varepsilon}}^{\mathrm{F}}) \end{aligned} \tag{2.167}$$

from which

$$\begin{aligned} &\underline{\underline{\Sigma}} + \mathbf{C} : (\mathbf{S}^{\mathrm{Esh}} - \mathbf{I}) : \underline{\underline{\varepsilon}}^{\mathrm{F}*} = \mathbf{c}^{\mathrm{HI}} : (\underline{\underline{E}} + \mathbf{S}^{\mathrm{Esh}} : \underline{\underline{\varepsilon}}^{\mathrm{F}*} - \underline{\underline{\varepsilon}}^{\mathrm{F}}) \\ &\Rightarrow \quad \underline{\underline{\varepsilon}}^{\mathrm{F}*} = [\mathbf{C} - \delta\mathbf{c}^{\mathrm{HI}} : \mathbf{S}^{\mathrm{Esh}}]^{-1} : [\mathbf{c}^{\mathrm{HI}} : \underline{\underline{\varepsilon}}^{\mathrm{F}} - \delta\mathbf{c}^{\mathrm{HI}} : \underline{\underline{E}}] \end{aligned} \tag{2.168}$$

and then, from $\underline{\underline{\varepsilon}}^{\mathrm{HI}} = \underline{\underline{E}} + \mathbf{S}^{\mathrm{Esh}} : \underline{\underline{\varepsilon}}^{\mathrm{F}*}$ and after some algebra

$$\underline{\underline{\varepsilon}}^{\mathrm{HI}} = (\mathbf{I} + \mathbf{P} : \delta\mathbf{c}^{\mathrm{HI}})^{-1} : (\underline{\underline{E}} + \mathbf{P} : \mathbf{c}^{\mathrm{HI}} : \underline{\underline{\varepsilon}}^{\mathrm{F}}) \tag{2.169}$$

If the matrix is subjected to the uniform stress-free strain $\underline{\underline{E}}^{\mathrm{F}}$, (2.169) becomes simply

$$\underline{\underline{\varepsilon}}^{\mathrm{HI}} = (\mathbf{I} + \mathbf{P} : \delta\mathbf{c}^{\mathrm{HI}})^{-1} : \left[\underline{\underline{E}} + \mathbf{P} : (\mathbf{c}^{\mathrm{HI}} : \underline{\underline{\varepsilon}}^{\mathrm{F}} - \mathbf{C} : \underline{\underline{E}}^{\mathrm{F}})\right] \tag{2.170}$$

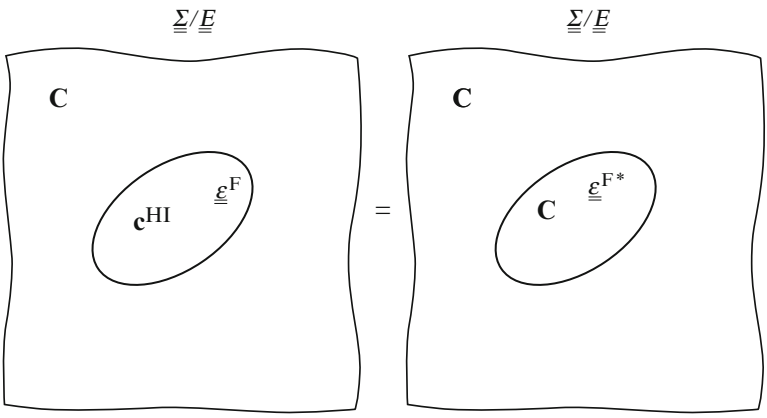


Fig. 2.18 Elastic heterogeneity with stress-free strain

Comment.

In all these problems we have

$$\begin{cases} \underline{\underline{\sigma}}^{\mathrm{I}} = \underline{\underline{\Sigma}} + \mathbf{C}^* : (\underline{\underline{E}} - \underline{\underline{\varepsilon}}^{\mathrm{I}}) \\ \mathbf{C}^* = \mathbf{C} : \left[(\mathbf{S}^{\mathrm{Esh}})^{-1} - \mathbf{I} \right] \end{cases} \tag{2.171}$$

$\mathbf{C}^*$ is Hill's *constraint tensor*; the relation in which it appears has the status of a *law of interaction*. $\mathbf{C}^*$ depends on the elastic moduli $\mathbf{C}$ of the matrix and on the shape (aspect ratios) and orientation of the inclusion.

2.7.4 Notes

2.7.4.1 Stored Energy

The stored energy W_{st} corresponding to an ellipsoidal inclusion with the uniform eigenstrain $\underline{\underline{\varepsilon}}^{\mathrm{F}}$ in an infinite load-free matrix is easily obtained from the basic analysis. We have

$$\begin{aligned} W_{\mathrm{st}} &= \int \sigma_{ij}\, \varepsilon_{ij}^{\mathrm{e}}\ \mathrm{d}V = \int \sigma_{ij}\, \varepsilon_{ij}\ \mathrm{d}V - \int_{\mathrm{I}} \sigma_{ij}\, \varepsilon_{ij}^{\mathrm{F}}\ \mathrm{d}V \\ &= \int_S \sigma_{ij}\, u_i\, n_j\ \mathrm{d}S - V_{\mathrm{I}}\, \sigma_{ij}^{\mathrm{I}}\, \varepsilon_{ij}^{\mathrm{F}} \end{aligned} \tag{2.172}$$

and the integral over S at infinity vanishes since there is no applied load. Thus

$$W_{\text{st}} = -V_{\text{I}}\sigma^{\text{I}}_{ij}\varepsilon^{\text{F}}_{ij} = -V_{\text{I}}\underline{\underline{\varepsilon}}^{\text{F}} : \mathbf{C} : (\mathbf{S}^{\text{Esh}} - \mathbf{I}) : \underline{\underline{\varepsilon}}^{\text{F}} \tag{2.173}$$

For the sphere and the case of isotropy we find

$$W_{\text{st}} = V\left\{[\lambda(1-\alpha) + (2\mu/3)(\beta-\alpha)]\left(\varepsilon^{\text{F}}_{kk}\right)^2 + 2\mu(1-\beta)\varepsilon^{\text{F}}_{ij}\varepsilon^{\text{F}}_{ij}\right\} \tag{2.174}$$

The effect of a macroscopic load $(\underline{\underline{\Sigma}}, \underline{\underline{E}})$ at infinity can be found by using the general result obtained in Sect. 2.5.2.2 and the definition of the residual stresses $\underline{\underline{\sigma}}^{\text{r}}$.

The problem of elastic heterogeneities can be attacked by using, as before, the method of fictitious equivalent inclusions.

2.7.4.2 Inclusions and Dislocations

The inclusion problem could have been solved by using a different representation, that of a continuous distribution of dislocations. Giving a field $\underline{\underline{\varepsilon}}^{\text{F}}$ of stress-free strains is in fact equivalent to giving such a distribution. Although piece-wise continuous, $\underline{\underline{\varepsilon}}^{\text{F}}$ is not compatible: it corresponds to surface incompatibility, localised at the interface between the inclusion and the matrix; and this incompatibility is itself equivalent to a surface distribution of dislocations, defined completely by $\underline{\underline{\varepsilon}}^{\text{F}}$ and the geometry of the interface, localised at the boundary between the inclusion and the matrix. Thus the stress fields in the inclusion and the matrix can be calculated as the resultant of a given distribution of dislocations, localised at the interface. Using other terms, we could speak of such dislocations as "geometrically necessary dislocations"; the problem would remain the same, with the same underlying physics. – see Chap. 3 on Elastoplasticity.

2.8 Improved Bounds and Estimates for Elastic Inhomogeneous Materials

The solution of the basic problem of inclusions and the underlying Green techniques can be used in several ways to help solving the problem of concentration (see Sect. 1.4.4.1b) within the general methodology of homogenisation. This can be done to obtain improved bounds and estimates.

2.8.1 Methodology for Getting Sharper Bounds (Outline)

Sharper bounds than Voigt-Reuss ones can only result from taking into account additional information with respect to the sole volume fractions of the phases.

Whereas the latter can be considered as first-order data, *i.e.,* resulting from the knowledge of one-point correlation functions for the elastic moduli, improved bounds are expected to need the knowledge of higher order correlation functions. In Sect. 2.9, we give a brief outline of a systematic way to associate sharper and sharper bounds with the use of correlation functions of increasing order. Here, we only suggest a method for getting "second-order" bounds using the knowledge of second-order information on the spatial distribution of the phases. The basic idea consists in using as trial fields the solutions to the general problem of an infinite homogeneous elastic matrix subjected to an arbitrary distribution of fictitious eigenstrains (or "polarisation stresses"), thanks to the Green techniques. These fields are then introduced in the general extremum theorems and restricted to specific expressions which allow us to calculate the elastic energy when specific properties of the phase spatial distribution are assumed. Improved bounds then result from an optimisation process with respect to the remaining parameters which specify the distribution of the considered eigenstrains. This can be made as follows.

(i) Consider first the basic extremum theorems for an RVE with homogeneous boundary conditions derived from their general expression (2.128) and (2.131) and used in Sect. 2.5.3 for getting Voigt-Reuss bounds, namely for a given macroscopic strain $\underline{\underline{E}}$

$$< \underline{\underline{\sigma}}^* >: \underline{\underline{E}} - \frac{1}{2} < \underline{\underline{\sigma}}^* : \mathbf{s} : \underline{\underline{\sigma}}^* > \leq \frac{1}{2} \underline{\underline{E}} : \mathbf{C}^{\text{eff}} : \underline{\underline{E}} \leq \frac{1}{2} < \underline{\underline{\varepsilon}}' : \mathbf{c} : \underline{\underline{\varepsilon}}' > \tag{2.175}$$

(ii) The trial fields $\underline{\underline{\sigma}}^*(\underline{x})$ and $\underline{\underline{\varepsilon}}'(\underline{x})$ are now obtained as the solution of the problem of an infinite elastic body with arbitrary uniform moduli $\mathbf{C}^0$, which is subjected to an arbitrary distribution of fictitious eigenstrains $\underline{\underline{\varepsilon}}^{\text{F}}(\underline{x})$: it is easy to check that they are SA and KA for the heterogeneous RVE, respectively. Alternatively, a distribution of eigenstresses $\underline{\underline{p}}(\underline{x})$ can be considered; the solution is the same as before if $\underline{\underline{p}}(\underline{x}) = -\mathbf{C}^0 : \underline{\underline{\varepsilon}}^{\text{F}}(\underline{x})$. A third (equivalent) way to proceed is to apply a fictitious distribution of body forces $\underline{f}(\underline{x})$ deriving from the potential $\underline{\underline{\pi}}(\underline{x})$ (*i.e.,* such that $\underline{f}(\underline{x}) = \text{div}\underline{\underline{\pi}}(\underline{x})$), where $\underline{\underline{\pi}}(\underline{x})$ is a field of symmetric second-order tensors, homogeneous to stresses: it is named "*polarisation stress field*"; the solution $\underline{\underline{\varepsilon}}'(\underline{x})$ to this third problem is KA too, but $\underline{\underline{\sigma}}'(\underline{x}) = \mathbf{C}^0 : \underline{\underline{\varepsilon}}'(\underline{x})$ is not SA for the original problem; suffice then to consider instead of $\underline{\underline{\sigma}}'(\underline{x})$ the stress field $\underline{\underline{\sigma}}^*(\underline{x}) = \underline{\underline{\sigma}}'(\underline{x}) + \underline{\underline{\pi}}(\underline{x})$ which is SA. The solution is the same as before when $\underline{\underline{\pi}}(\underline{x}) = \underline{\underline{p}}(\underline{x}) = -\mathbf{C}^0 : \underline{\underline{\varepsilon}}^{\text{F}}(\underline{x})$. This solution can be obtained, at least formally (see Sect. 2.9), by generalisation of the resolution of the inclusion problem using Green techniques.

(iii) We now apply (2.175) to the foregoing trial fields. From Hill's lemma, we get

$$\begin{aligned} <\underline{\underline{\sigma}}^* : \underline{\underline{\varepsilon}}'> &= <\underline{\underline{\varepsilon}}' : \underline{\underline{\sigma}}^*> = <\underline{\underline{\varepsilon}}'> : <\underline{\underline{\sigma}}^*> \\ &= \underline{\underline{E}} : <\mathbf{C}^0 : \underline{\underline{\varepsilon}}' + \underline{\underline{p}}> = \underline{\underline{E}} : \mathbf{C}^0 : \underline{\underline{E}} + \underline{\underline{E}} : <\underline{\underline{p}}> \\ &= <\underline{\underline{\varepsilon}}' : \mathbf{C}^0 : \underline{\underline{\varepsilon}}'> + <\underline{\underline{\varepsilon}}' : \underline{\underline{p}}> \end{aligned} \tag{2.176}$$

so that the last term in (2.173) can be put in the more convenient form in view of extracting second-order correlation functions and getting rid of higher order ones, with $\delta\mathbf{c}^0(\underline{x}) = \mathbf{c}(\underline{x}) - \mathbf{C}^0$:

$$\begin{aligned} \frac{1}{2} <\underline{\underline{\varepsilon}}' : \mathbf{c} : \underline{\underline{\varepsilon}}'> &= \frac{1}{2} <\underline{\underline{\varepsilon}}' : \mathbf{C}^0 : \underline{\underline{\varepsilon}}'> + \frac{1}{2} <\underline{\underline{\varepsilon}}' : \delta\mathbf{c}^0 : \underline{\underline{\varepsilon}}'> \\ &= \frac{1}{2}\underline{\underline{E}} : \mathbf{C}^0 : \underline{\underline{E}} + \frac{1}{2}\underline{\underline{E}} : <\underline{\underline{p}}> - \frac{1}{2} <\underline{\underline{\varepsilon}}' : \underline{\underline{p}}> + \frac{1}{2} <\underline{\underline{\varepsilon}}' : \delta\mathbf{c}^0 : \underline{\underline{\varepsilon}}'> \end{aligned} \tag{2.177}$$

Additional elementary algebra helps transforming this expression into

$$\begin{aligned} &\frac{1}{2} <\underline{\underline{\varepsilon}}' : \mathbf{c} : \underline{\underline{\varepsilon}}'> = H^0(\underline{\underline{p}}) + \delta u^0 \\ &H^0(\underline{\underline{p}}) = \frac{1}{2}\underline{\underline{E}} : \mathbf{C}^0 : \underline{\underline{E}} + \frac{1}{2}\underline{\underline{E}} : <\underline{\underline{p}}> + \frac{1}{2} <\underline{\underline{\varepsilon}}' : \underline{\underline{p}}> - \frac{1}{2} <\underline{\underline{p}} : \mathbf{m}^0 : \underline{\underline{p}}> \\ &\delta u^0 = \frac{1}{2} <(\underline{\underline{\varepsilon}}' - \underline{\underline{p}} : \mathbf{m}^0) : \delta\mathbf{c}^0 : (\underline{\underline{\varepsilon}}' - \mathbf{m}^0 : \underline{\underline{p}})> \quad \text{with } \mathbf{m}^0 = (\delta\mathbf{c}^0)^{-1} \end{aligned} \tag{2.178}$$

The main advantage of this transformation lies in the fact that most of the contributions to the upper bound $(1/2) <\underline{\underline{\varepsilon}}' : \mathbf{c} : \underline{\underline{\varepsilon}}'>$ which involve high order correlation functions are now included in δu^0, whereas the remaining contribution, the so-called Hashin-Shtrikman[27] functional $H^0(\underline{\underline{p}})$, only depends on the polarisation stress field and is likely to be calculable with help of second-order information only.

[27] Zvi Hashin (1929–) is an Israeli physicist and engineer, Emeritus Professor at the Tel Aviv University. He derived the Hashin-Shtrikman bounds in 1963 in collaboration with his colleague Shmuel Shtrikman (1930–2003), a Belarusian-born physicist of the Weizmann Institute of Science, Rehovot, Israel, and of the Franklin Institute Laboratories, Philadelphia, USA, when he was a faculty member at the University of Pennsylvania, Philadelphia (USA).

If the same treatment is now applied to the term $< \underline{\underline{\sigma}}^* : \mathbf{s} : \underline{\underline{\sigma}}^* >$ instead of $< \underline{\underline{\varepsilon}}' : \mathbf{c} : \underline{\underline{\varepsilon}}' >$, it leads to

$$< \underline{\underline{\sigma}}^* >: \underline{\underline{E}} - \frac{1}{2} < \underline{\underline{\sigma}}^* : \mathbf{s} : \underline{\underline{\sigma}}^* >= H^0(\underline{\underline{p}}) - \delta v^0$$

$$\delta v^0 = \frac{1}{2} < (\underline{\underline{\sigma}}^* - \underline{\underline{p}} : \mathbf{S}^0 : \mathbf{l}^0) : \delta s^0 : (\underline{\underline{\sigma}}^* - \mathbf{l}^0 : \mathbf{S}^0 : \underline{\underline{p}}) > \quad \text{with } \mathbf{l}^0 = -(\delta \mathbf{s}^0)^{-1} \tag{2.179}$$

so that the bounding Eq. 2.175 reads simply

$$H^0(\underline{\underline{p}}) - \delta v^0 \leq \frac{1}{2} \underline{\underline{E}} : \mathbf{C}^{\text{eff}} : \underline{\underline{E}} \leq H^0(\underline{\underline{p}}) + \delta u^0 \tag{2.180}$$

From that we can state that, since the terms δu^0 and δv^0 can be hardly evaluated with only two-point information ($\underline{\underline{\varepsilon}}'$ and $\underline{\underline{\sigma}}^*$ already imply two-point correlation functions), they can be removed from the bounding Eq. 2.180 provided that they are negative for any $\underline{\underline{p}}$. This is easily obtained by choosing $\mathbf{C}^0$ and $\mathbf{S}^0$ large enough for $\delta \mathbf{c}^0$ and $\delta \mathbf{s}^0$ to be negative everywhere in the RVE, respectively. For such a choice of $\mathbf{C}^0$ or $\mathbf{S}^0$, say $\mathbf{C}^{\text{sup}}$ and $\mathbf{S}^{\text{inf}}$, the functional $H^0(\underline{\underline{p}})$, which is then equal to $H^{\text{sup}}(\underline{\underline{p}})$or $H^{\text{inf}}(\underline{\underline{p}})$, yields an upper or a lower bound for the effective moduli, respectively.

(iv) In the functional $H^0(\underline{\underline{p}})$, the contribution of the term $< \underline{\underline{\varepsilon}}' : \underline{\underline{p}} >$ is the most difficult to be computed since the strain tensor at point $\underline{x}$, $\underline{\underline{\varepsilon}}'(\underline{x})$, depends on the value of the polarisation stress field at any point $\underline{\underline{p}}(\underline{x}')$. This will need first specific enough polarisation fields and then some adequate properties of the spatial distribution of the phases to be known. The classical Hashin and Shtrikman bounds can be obtained by choosing $\underline{\underline{p}}(\underline{x})$ piecewise constant per phase – say $\underline{\underline{p}}(\underline{x}) = \sum_r \underline{\underline{p}}_r \varphi_r(\underline{x})$ where $\varphi_r(\underline{x})$ is the characteristic function of phase (r), by assuming the distribution of the phases to be *isotropic* and by optimising the values of $\underline{\underline{p}}_r$ as well as $\mathbf{C}^{\text{sup}}$ and $\mathbf{S}^{\text{inf}}$. The isotropy of the phase distribution can be defined by the property of the two-point correlation function $\varphi_{rs}(\underline{u})$ given by

$$\varphi_{rs}(\underline{u}) = \frac{1}{V_r} \int_V \varphi_r(\underline{x}) \varphi_s(\underline{x} + \underline{u}) \, \mathrm{d}V \tag{2.181}$$

not to depend on the direction of $\underline{u}$, *i.e.*, $\varphi_{rs}(\underline{u}) = \varphi_{rs}(|\underline{u}|)$, $\forall r, s$, with $\varphi_{rs}(\underline{0}) = \delta_{rs}$.

(v) The resulting optimal polarisation stress fields $\underline{\underline{p}}_r^{\text{opt}}$ are found to be

$$\underline{\underline{p}}_r^{\text{opt}} = \delta\mathbf{c}_r^0 : \mathbf{a}_r^0 :< \mathbf{a}^0 >^{-1} : \underline{\underline{E}} \quad \text{with} \quad \mathbf{a}_r^0 = (\mathbf{I} + \mathbf{P}_{\text{Sph}}^0 : \delta\mathbf{c}_r^0)^{-1} \tag{2.182}$$

where Hill's tensor defined by (2.153) refers to an infinite body with moduli $\mathbf{C}^0$ and to a spherical inclusion. According to (2.178), the corresponding value of the bounding functional H_{opt}^0 is

$$H_{\text{opt}}^0 = \frac{1}{2}\underline{\underline{E}} :< \mathbf{c} : \mathbf{a}^0 >:< \mathbf{a}^0 >^{-1} : \underline{\underline{E}} \tag{2.183}$$

This value has still to be optimised with respect to $\mathbf{C}^0$ by choosing $\mathbf{C}^{\text{sup}}$ (resp. $\mathbf{S}^{\text{inf}}$) as "low" (in the sense of the associated quadratic form) as possible while still making $\delta\mathbf{c}^{\text{sup}}$ (resp. $\delta\mathbf{s}^{\text{inf}}$) negative everywhere in the RVE. This discussion is easier when all the phase moduli are isotropic. For this case, the overall moduli $\mathbf{C}^{\text{eff}}$ are isotropic too and defined, for instance, by the shear and bulk moduli μ^{eff} and k^{eff}. Their Hashin-Shtrikman upper (μ^{HS^+} and k^{HS^+}) and lower bounds (μ^{HS^-} and k^{HS^-}) are easily found (see Volume III for an explicit derivation) to be, with obvious notation

$$\begin{cases} \mu^{\text{HS}^\pm} =< \mu\left(1 + \dfrac{\beta^\pm\delta\mu^\pm}{\mu^\pm}\right)^{-1} >< \left(1 + \dfrac{\beta^\pm\delta\mu^\pm}{\mu^\pm}\right)^{-1} >^{-1} \\ k^{\text{HS}^\pm} =< k(1 + \dfrac{\alpha^\pm\delta k^\pm}{k^\pm})^{-1} >< \left(1 + \dfrac{\alpha^\pm\delta k^\pm}{k^\pm}\right)^{-1} >^{-1} \end{cases} \tag{2.184}$$

with α and β the Eshelby coefficients (2.159) and under the assumption of isotropic phase distribution and isotropic phase moduli.

In these equations, $\mu^+ = \sup_r \{\mu_r\}, k^+ = \sup_r \{k_r\}$ and $\mu^- = \inf_r \{\mu_r\}, k^- = \inf_r \{k_r\}$.

Note that μ^+ (resp. μ^-) and k^+ (resp. k^-) do not necessarily refer to the same phase: one phase can be the "stiffest" one for μ and not for k and conversely. For a two-phase isotropic composite, it is easy to check the general result that Hashin-Shtrikman bounds always improve on Voigt and Reuss ones: they define a "Hashin-Shtrikman domain" which lies inside Hill's one (see Fig. 2.11).

Three additional comments can be given:

– formulae (2.182) and (2.183) suggest, through the occurrence of the term $\mathbf{a}_r^0 = (\mathbf{I} + \mathbf{P}_{\text{Sph}}^0 : \delta\mathbf{c}_r^0)^{-1}$, that some connexion may exist between Hashin-Shtrikman bounds and the solution of Eshelby's problem of a spherical heterogeneity in an infinite matrix – see (2.159) and (2.160). This point is clarified in the next section.

– Hashin and Shtrikman bounds can also be derived for other phase distributions than an isotropic one. This is particularly the case for "ellipsoidal distributions" which give rise to an orthotropic anisotropy of the elastic moduli.
– Sharper bounds than Hashin-Shtrikman ones can be derived, at least formally, from taking into account richer information than the symmetry of the anisotropy of the phase distribution. A systematic way to do so is presented briefly in Sect. 2.9.

We leave the reader to consult specialist works (*e.g.* Hashin and Shtrikman 1963; Kröner 1977; Willis 1977; Zaoui 2000) for proofs and detailed developments on this subject.

2.8.2 Improved Estimates

We can easily construct a whole set of estimates of the overall moduli $\mathbf{C}^{\text{est}}$ by assuming that the average mechanical state $(\underline{\underline{\sigma}}_r, \underline{\underline{\varepsilon}}_r)$ in any phase (r) of a multiphase material is the same as that of some ellipsoidal heterogeneity I_r with the same moduli $\mathbf{c}_r$, embedded in an infinite homogeneous matrix with moduli $\mathbf{C}^0$ and the uniform strain at infinity $\underline{\underline{E}}^0$. The shape and orientation of I_r can be chosen from microscopic observation of the (r) phase domains (fibres, particles, grains, etc.). The auxiliary variable $\underline{\underline{E}}^0$ is to be derived from the condition that the average of $\underline{\underline{\varepsilon}}_r$ for all r values is the prescribed macroscopic strain $\underline{\underline{E}}$ (see below), so that the only free variable is $\mathbf{C}^0$. These moduli can be chosen so as to express some morphological properties of the material under consideration; each choice leads to a specific estimate for $\mathbf{C}^{\text{eff}}$ that is, to a specific inclusion-based model of inhomogeneous material.

From (2.164) we know the relation between $\underline{\underline{\varepsilon}}_r$ and $\underline{\underline{E}}^0$ for an ellipsoidal inhomogeneity in an infinite matrix

$$\underline{\underline{\varepsilon}}_r = (\mathbf{I} + \mathbf{P}^0 : \delta\mathbf{c}_r^0)^{-1} : \underline{\underline{E}}^0 \quad \text{with } \mathbf{P}^0 = \mathbf{S}^{0\text{Esh}} : (\mathbf{C}^0)^{-1}, \quad \delta\mathbf{c}_r^0 = \mathbf{c}_r - \mathbf{C}^0 \tag{2.185}$$

the index 0 emphasising the fact that these tensors depend on the choice of $\mathbf{C}^0$. The auxiliary variable $\underline{\underline{E}}^0$ is determined from the condition on the average, $<\underline{\underline{\varepsilon}}>= \underline{\underline{E}}$; thus:

$$< (\mathbf{I} + \mathbf{P}^0 : \delta\mathbf{c}^0)^{-1} >: \underline{\underline{E}}^0 = \underline{\underline{E}} \quad \Rightarrow \quad \underline{\underline{E}}^0 =< (\mathbf{I} + \mathbf{P}^0 : \delta\mathbf{c}^0)^{-1} >^{-1} : \underline{\underline{E}} \tag{2.186}$$

Then from the relation $< \mathbf{c} : \underline{\underline{\varepsilon}}^{\text{est}} >= \mathbf{C}^{\text{est}} : \underline{\underline{E}}$ defining the estimate of the overall moduli $\mathbf{C}^{\text{est}}$, we get

$$\mathbf{C}^{\text{est}} =< \mathbf{c} : (\mathbf{I} + \mathbf{P}^0 : \delta\mathbf{c}^0)^{-1} >:< (\mathbf{I} + \mathbf{P}^0 : \delta\mathbf{c}^0)^{-1} >^{-1} \tag{2.187}$$

or after some elementary algebra, with $\delta \mathbf{c}_r^{\text{est}} = \mathbf{c}_r - \mathbf{C}^{\text{est}}$

$$< \delta \mathbf{c}^{\text{est}} : (\mathbf{I} + \mathbf{P}^0 : \delta \mathbf{c}^0)^{-1} >= 0 \tag{2.188}$$

A detailed analysis would have shown that when $\mathbf{C}^0$ tends to infinity $\mathbf{C}^{\text{est}}$ tends to $<\mathbf{c}>$, which is Voigt's upper bound: this can be understood intuitively by recognising that we overestimate the effective moduli by embedding each phase in a rigid body. When $\mathbf{C}^0$ tends to 0, $\mathbf{C}^{\text{est}}$ tends to $<\mathbf{c}^{-1}>^{-1}$, which is Reuss's lower bound; the corresponding explanation is that we underestimate $\mathbf{C}^{\text{eff}}$ by embedding each phase in an infinitely soft matrix, which allows the stress at infinity to be transmitted without change to the inclusion, as in Reuss's assumption of uniform stress.

In addition to these extreme choices of $\mathbf{C}^0$, two special cases are often considered, corresponding roughly to composite materials and polycrystals respectively.

NB: When the volume fraction $f_2 = f$ of the particles with moduli $\mathbf{c}_2$ is very low, $\underline{\underline{E}}^0$ can be assimilated to $\underline{\underline{E}}$ and the "reference medium" with moduli $\mathbf{C}^0$ to the matrix itself (*i.e.,* $\mathbf{C}^0 = \mathbf{c}^{\text{mat}}$). So, this "*dilute approximation*" consists in estimating $\underline{\underline{\varepsilon}}_2^{\text{est}}$ as

$$\underline{\underline{\varepsilon}}_2^{\text{est}} \approx \left[\mathbf{I} + \mathbf{P}^{\text{mat}} : (\mathbf{c}_2 - \mathbf{c}^{\text{mat}})\right]^{-1} : \underline{\underline{E}} \tag{2.189}$$

and then $\underline{\underline{\varepsilon}}_1^{\text{est}}$, from $(1-f)\underline{\underline{\varepsilon}}_1^{\text{est}} + f\,\underline{\underline{\varepsilon}}_2^{\text{est}} = \underline{\underline{E}}$, as

$$(1-f)\underline{\underline{\varepsilon}}_1^{\text{est}} = \left[\mathbf{I} - f\left[\mathbf{I} + \mathbf{P}^{\text{mat}} : (\mathbf{c}_2 - \mathbf{c}^{\text{mat}})\right]^{-1}\right] : \underline{\underline{E}} \tag{2.190}$$

By $\mathbf{C}^{\text{est}} : \underline{\underline{E}} =< \mathbf{c} : \underline{\underline{\varepsilon}}^{\text{est}} >$, we find

$$\mathbf{C}^{\text{EE}} = \mathbf{c}^{\text{mat}} + f(\mathbf{c}_2 - \mathbf{c}^{\text{mat}})\left[\mathbf{I} + \mathbf{P}_2^{\text{mat}} : (\mathbf{c}_2 - \mathbf{c}^{\text{mat}})\right]^{-1} \tag{2.191}$$

recalling that $\mathbf{P}_2{}^{\text{mat}} = \mathbf{S}^{\text{mat Esh}} : (\mathbf{c}^{\text{mat}})^{-1}$, and where the index EE refers to Eshelby and Einstein, since this model derives straightforwardly from Eshelby's solution of the inclusion problem and the dilute approximation had been already proposed by Einstein[28] for dilute viscous suspensions with rigid particles. Note that this dilute approximation must be restricted to very low particle volume fractions and could even violate Voigt-Reuss bounds for not low enough concentrations.

[28] Albert Einstein (1879–1955) who is quoted here is the same German-born theoretical physicist as the universally known father of modern physics.

2.8.2.1 Composite Materials: Mori-Tanaka Estimate and Hashin-Shtrikman Bounds

For a particle- or fibre-reinforced composite material we can choose as the surrounding material $\mathbf{C}^0$ the matrix of the composite, $\mathbf{C}^0 = \mathbf{c}^{\text{mat}}$, implying that the matrix is continuous, with the particles or fibres dispersed in it. From (2.187) the corresponding estimate is

$$\mathbf{C}^{\text{MT}} = < \mathbf{c} : (\mathbf{I} + \mathbf{P}^{\text{mat}} : \delta\mathbf{c}^{\text{mat}})^{-1} > : < (\mathbf{I} + \mathbf{P}^{\text{mat}} : \delta\mathbf{c}^{\text{mat}})^{-1} >^{-1} \tag{2.192}$$

where the index MT refers to Mori and Tanaka (1973), who have proposed such a model for composite materials, although in an unnecessarily complicated manner.

For two phases – the matrix $\mathbf{c}_1 = \mathbf{c}^{\text{mat}}$ and the inclusions $\mathbf{c}_2$, – with volume fractions $f_1 = 1 - f$ and $f_2 = f$ respectively – we find easily

$$\mathbf{C}^{\text{MT}} = \mathbf{c}_2 - (1-f)\left(\mathbf{c}_2 - \mathbf{c}^{\text{mat}}\right) : \{(1-f)\,\mathbf{I} + f\,[\mathbf{I} + \mathbf{P}_2^{\text{mat}} : \left(\mathbf{c}_2 - \mathbf{c}^{\text{mat}}\right)]^{-1}\}^{-1} \tag{2.193}$$

If the matrix and the particles, assumed spherical, are elastically isotropic, this equation leads to isotropic overall moduli, defined by μ^{MT} and k^{MT}, given by

$$\begin{cases} \mu^{\text{MT}} = \mu_1 \dfrac{(1-f)\mu^{\text{mat}} + f\mu_2 + (1-f)\beta^{\text{mat}}(\mu_2 - \mu^{\text{mat}})}{\mu^{\text{mat}} + (1-f)\beta^{\text{mat}}(\mu_2 - \mu^{\text{mat}})} \\ k^{\text{MT}} = k_1 \dfrac{(1-f)k^{\text{mat}} + f k_2 + (1-f)\alpha^{\text{mat}}(k_2 - k^{\text{mat}})}{k^{\text{mat}} + (1-f)\alpha^{\text{mat}}(k_2 - k^{\text{mat}})} \end{cases} \tag{2.194}$$

where α^{mat} and β^{mat} are the Eshelby coefficients (2.159) for the matrix. It can be checked that these estimates for μ^{eff} and k^{eff} coincide with their Hashin-Shtrikman lower bounds given by (2.184) when $\mu^- = \mu^{\text{mat}}$, $k^- = k^{\text{mat}}$.

This result fits with the intuitive physical interpretation of these bounds: with the inclusions being spherical because of the isotropic distribution, we over- or under-estimate the effective moduli depending on whether we embed each phase in a homogeneous infinite body with the highest or lowest, respectively, of the moduli of the constituent phases.

Note that the Mori-Tanaka model is not restricted to spherical particles and two-phase materials: it has been intensively applied to fibre-reinforced composites as well as to several kinds of particles or fibres. Nevertheless, for these cases, the connexion with the exact results of Hashin-Shtrikman bounds is looser and it may even happen that the estimates $\mathbf{C}^{\text{MT}}$ do not exhibit the required diagonal symmetry anymore, which means that these empirical estimates must be used with special care.

2.8.2.2 Disordered Polycrystals: The Self-consistent Scheme

For multiphase materials for which no phase plays any prominent morphological role, contrary to composites with a continuous matrix, the Mori-Tanaka model is irrelevant. This is especially the case for well-disordered polycrystals: every phase then consists of grains all with the same lattice orientation and each grain family is surrounded, on average, by almost all the other families, so that the surrounding medium $\mathbf{C}^0$ could be chosen as the effective medium itself. Let $\mathbf{C}^{\mathrm{SC}}$ be the corresponding estimate, according to the so-called self-consistent scheme ($\mathbf{C}^0 = \mathbf{C}^{\mathrm{est}} = \mathbf{C}^{\mathrm{SC}}$). From (2.187) and (2.188) we have

$$\mathbf{C}^{\mathrm{SC}} = < \mathbf{c} : (\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : \delta\mathbf{c}^{\mathrm{SC}})^{-1} > : < (\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : \delta\mathbf{c}^{\mathrm{SC}})^{-1} >^{-1} \tag{2.195}$$

or

$$< \delta\mathbf{c}^{\mathrm{SC}} : (\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : \delta\mathbf{c}^{\mathrm{SC}})^{-1} > = 0 \tag{2.196}$$

NB

1. In general, this will be an implicit equation for $\mathbf{C}^{\mathrm{SC}}$, since $\mathbf{P}^{\mathrm{SC}}$ and $\delta\mathbf{c}^{\mathrm{SC}}$ both depend on $\mathbf{C}^{\mathrm{SC}}$.
2. Instead of $\mathbf{P}^{\mathrm{SC}}$ we can use Hill's "constraint tensor" $\mathbf{C}^*$, defined by (2.171); the relation between the two tensors is simply $\mathbf{P} = (\mathbf{C}^0 + \mathbf{C}^*)^{-1}$.
3. Here again, note that problems (lack of diagonal symmetry for $\mathbf{C}^{\mathrm{SC}}$) may arise when several shapes or orientations of ellipsoidal inclusions are considered. When all the phases are associated with parallel ellipsoids, $\mathbf{P}_r^{\mathrm{SC}}$ in (2.195) or (2.196) does not depend on r so that, after multiplication by the constant tensor $\mathbf{P}_r^{\mathrm{SC}} = \mathbf{P}^{\mathrm{SC}}$, we find $< (\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : \delta\mathbf{c}^{\mathrm{SC}})^{-1} > = \mathbf{I}$, which means from (2.186)

$$\underline{\underline{E}}^0 = < (\mathbf{I} + \mathbf{P}^0 : \delta\mathbf{c}^0)^{-1} >^{-1} : \underline{\underline{E}} = \underline{\underline{E}} \tag{2.197}$$

and then simply

$$\mathbf{C}^{\mathrm{SC}} = < \mathbf{c} : (\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : \delta\mathbf{c}^{\mathrm{SC}})^{-1} >, \quad \text{when } \mathbf{P}_r^{\mathrm{SC}} = \mathbf{P}^{\mathrm{SC}} \ \forall r \tag{2.198}$$

2.8.2.3 Hashin-Shtrikman Estimates

In addition to the Mori-Tanaka and self-consistent models, this approach yields as many models as choices for $\mathbf{C}^0$. Besides Voigt and Reuss bounds which, as shown above, are recovered for extreme values for $\mathbf{C}^0$, we get a continuous variety of models, the so-called "Hashin-Shtrikman estimates", by letting $\mathbf{C}^0$ take any value lying between (in the sense of the associated quadratic forms) those which correspond to the Hashin-Shtrikman bounds.

This is easier to understand by restricting the discussion to an *isotropic two-phase material, with isotropic incompressible phases and shear moduli* $\mu_1 < \mu_2$. It is easily found that the Hashin-Shtrikman estimates for μ^{eff} vary with the parameter μ^0 (with k^0 infinite, $\mu_1 \leq \mu^0 \leq \mu_2$ and $\beta^0 = 2/5$) according to the equation:

$$\mu^{\text{est}} = \frac{2\mu_1\mu_2 + 3\mu^0 < \mu >}{2\mu_1\mu_2 < 1/\mu > +3\mu^0} \tag{2.199}$$

As particular cases, we find

$\mu^0 = 0 \rightarrow \mu^{\text{R}} = 1/ < 1/\mu >$ (Reuss bound)

$\mu^0 = \mu_1 \rightarrow \mu^{\text{MT}} = \mu^{\text{HS}^-} = \dfrac{2\mu_2 + 3\mu^{\text{V}}}{2\mu_2 + 3\mu^{\text{R}}}\mu^{\text{R}}$ (MT estimate and HS lower bound)

$\mu^0 = \mu^{\text{est}} = \mu^{\text{SC}} \rightarrow \mu^{\text{SC}}$:> 0 root of

$$3X^2 + \left(\frac{2\mu_1\mu_2}{\mu^{\text{R}}} - 3\mu^{\text{V}}\right) X - 2\mu_1\mu_2 = 0 \text{ (SC estimate)}$$

$\mu^0 = \mu_2 \rightarrow \mu^{\text{HS}^+} = \dfrac{2\mu_1 + 3\mu^{\text{V}}}{2\mu_1 + 3\mu^{\text{R}}}\mu^{\text{R}}$ (HS upper bound)

$\mu^0 \rightarrow \infty \rightarrow \mu^{\text{V}} =< \mu >$ (Voigt bound)

2.9 Systematic Theory of the Elasticity of Random Media (Outline)

The above treatment of the Voigt/Reuss and Hashin-Shtrikman bounds and of the self-consistent model could be set, in linear elasticity, in the framework of a general theory of randomly heterogeneous media. Such a theory has been developed, especially by Kröner[29], over the past 40 years and for a full account we refer the reader to the specialist literature (Kröner 1977, 1978). Here we give only the underlying principles and the main results achieved.

2.9.1 General Equation for Heterogeneous Elastic Media

The problem of heterogeneous inclusions arises as a special case of a general problem of heterogeneous media, which can be solved, at least formally, by the

[29]Ekkehart Kröner (1919–2000) was a German physicist and a Professor at the University of Stuttgart (FRG).

Green's method already introduced (Sect. 2.7.1). We consider such a medium V with variable local elastic moduli $\mathbf{c}(\underline{x})$ and boundary conditions at the surface ∂V, for example displacements $\underline{u}^{\mathrm{g}}$. We introduce a fictitious reference medium with the same geometry and the same boundary conditions, but with uniform moduli $\mathbf{C}^0$.

Putting $\mathbf{c} = \mathbf{C}^0 + \delta\mathbf{c}^0$ we have

$$\underline{\underline{\sigma}} = \mathbf{c} : \underline{\underline{\varepsilon}} = (\mathbf{C}^0 + \delta\mathbf{c}^0) : \underline{\underline{\varepsilon}} = (\mathbf{C}^0 + \delta\mathbf{c}^0) : \mathrm{grad}\underline{u} \tag{2.200}$$

In the absence of body forces the equilibrium equation in rectangular cartesian co-ordinates can be written

$$\sigma_{ij,j} = [(C^0_{ijkl} + \delta c^0_{ijkl})u_{l,k}]_{,j} = C^0_{ijkl}u_{l,kj} + (\delta c^0_{ijkl}u_{l,k})_{,j} = 0 \tag{2.201}$$

We can regard the equilibrium equations as relating to the homogeneous reference medium to which are applied not only the displacements $\underline{u}^{\mathrm{g}}$ at the boundary ∂V but also the (fictitious) volume force density f_i defined by

$$f_i = \left(\delta c^0_{ijkl}u_{l,k}\right)_{,j} \tag{2.202}$$

Given the Green's tensor $\underline{\underline{G}}^0$ for the reference medium, we can find the displacement corresponding to unit point force under the conditions of zero displacement at the surface; this is the tensor solution of (2.147) with $\underline{\underline{G}}^0$ and $\mathbf{C}^0$ instead of $\underline{\underline{G}}$ and $\mathbf{C}$ and with $\underline{\underline{G}}^0$ zero on ∂V; the solution to (2.201) for the displacements u_i can be written

$$u_i(\underline{x}) = u^0_i(\underline{x}) + \int_V G^0_{ij}\left(\delta c^0_{jklm}u_{m,l}\right)_{,k'}\mathrm{d}V' \tag{2.203}$$

where $u^0_i(\underline{x})$ is the response of the reference medium to the boundary conditions and we use once again the notation $\mathrm{A}_{,k'} = \partial\mathrm{A}/\partial x'_k$ to distinguish derivatives with respect to $\underline{x}$ or $\underline{x}'$.

Integrating by parts and using the condition that $\underline{\underline{G}}^0$ vanishes on the boundary

$$\begin{aligned} u_i &= u^0_i + \int_V \left[\left(G^0_{ij}\delta c^0_{jklm}u_{m,l'}\right)_{,k'} - G^0_{ij,k'}\delta c^0_{jklm}u_{m,l'}\right]\mathrm{d}V' \\ &= u^0_i + \int_{\partial V} G^0_{ij}\delta c^0_{jklm}u_{m,l'}n'_k\,\mathrm{d}S' - \int_V G^0_{ij,k'}\delta c^0_{jklm}u_{m,l'}\,\mathrm{d}V' \\ &= u^0_i - \int_V G^0_{ij,k'}\delta c^0_{jklm}u_{m,l'}\,\mathrm{d}V' \end{aligned} \tag{2.204}$$

Differentiating $u_i(\underline{x})$ with respect to x_n gives (omitting some mathematical details)

$$u_{i,n} = u^0_{i,n} - \int_V G^0_{ij,k'n}\delta c^0_{jklm}u_{m,l'}\,\mathrm{d}V' \tag{2.205}$$

and symmetrisation with respect to (i,n) gives

$$\varepsilon_{in} = \varepsilon^0_{in} - \int_V \Gamma^0_{injk} \delta c^0_{jklm} u_{m,l'} \, \mathrm{d}V' \tag{2.206}$$

where, taking into account the symmetry of δc^0_{jklm} with respect to j and k,

$$\Gamma^0_{injk} = G^0_{ij,k'n_{(in)(jk)}} \tag{2.207}$$

Using a condensed notation with the symbol $\otimes$: meaning a convolution product combined with the product of a fourth-order tensor and a second-order tensor we have the integral equation

$$\underline{\underline{\varepsilon}}(\underline{x}) + \left[\tilde{\mathbf{\Gamma}}^0(\underline{x}, \underline{x}') : \delta \mathbf{c}^0(\underline{x}') \right] \otimes : \underline{\underline{\varepsilon}}(\underline{x}') = \underline{\underline{\varepsilon}}^0(\underline{x}) \tag{2.208}$$

or in a more abbreviated form which still shows the integral-equation type of relation between the operators

$$\underline{\underline{\varepsilon}} + \left(\tilde{\mathbf{\Gamma}}^0 : \delta \mathbf{c}^0 \right) \otimes : \underline{\underline{\varepsilon}} = \underline{\underline{\varepsilon}}^0 \tag{2.209}$$

This is the general equation we wanted; it involves the previously-defined fundamental Green's *operator* $\tilde{\mathbf{\Gamma}}^0$, the so-called *modified Green's operator*, here for deformations. There is a corresponding equation for the stresses, involving a corresponding operator $\tilde{\mathbf{\Delta}}^0$, for stresses: the following relations are easily deduced from (2.201):

$$\begin{cases} \underline{\underline{\sigma}} + \tilde{\mathbf{\Delta}}^0 : \delta \mathbf{s}^0 \otimes : \underline{\underline{\sigma}} = \underline{\underline{\sigma}}^0 \\ \mathbf{S}^0 : \underline{\underline{\sigma}}^0 = \underline{\underline{\varepsilon}}^0 \\ \tilde{\mathbf{\Delta}}^0 = \tilde{\mathbf{C}}^0 - \mathbf{C}^0 : \tilde{\mathbf{\Gamma}}^0 : \mathbf{C}^0 \end{cases} \tag{2.210}$$

where $\tilde{\mathbf{C}}^0 = \mathbf{C}^0 \delta(\underline{x}, \underline{x}')$.

2.9.2 *Properties of the Modified Green's Operator*

It can be shown that the operator $\tilde{\mathbf{\Gamma}}^0$ can be separated into two parts, a singular part $\tilde{\mathbf{E}}^0$ called *local*, and a regular part, $\tilde{\mathbf{F}}^0$ called, *long-range*, such that

$$\tilde{\mathbf{\Gamma}}^0 = \tilde{\mathbf{E}}^0 + \tilde{\mathbf{F}}^0, \quad \tilde{\mathbf{E}}^0(\underline{x}) = \mathbf{E}^0(\underline{x}) \delta(\underline{x}, \underline{x}') \tag{2.211}$$

where $\tilde{\mathbf{F}}^0(\underline{x}, \underline{x}')$ decreases like $1/|\underline{x}-\underline{x}'|$ and $\mathbf{E}^0$ is a simple fourth-order tensor. In an infinite medium the integral of $\tilde{\mathbf{F}}^0$ over the whole volume is zero and that of $\tilde{\mathbf{\Gamma}}^0$ reduces to $\mathbf{E}^0$. Further, the integral of $\tilde{\mathbf{F}}^0$ over the volume of an ellipsoid is uniform (and zero for a sphere) if $\underline{x}$ is in the interior. These are remarkable properties; they account for the simplicity of the solution of Eshelby's problem of an ellipsoidal inclusion in an infinite medium, and in particular for the uniformity of the solution within the inclusion.

Unfortunately, it is not possible to give an explicit analytic expression for $\tilde{\mathbf{\Gamma}}^0$ for an arbitrary anisotropic medium. However, it can be expressed simply in the case of an *infinite isotropic medium*; and in what follows we shall, for the most part, limit ourselves to this case – which could be extended to that of transverse isotropy. According to (2.154), we can also write

$$G^0_{ij}(\underline{x}, \underline{x}') = G^0_{ij}(|\underline{x} - \underline{x}'|) = G^0_{ij}(\rho) = \frac{(3k^0 + 7\mu^0)\delta_{ij} + (3k^0 + \mu^0)e_i e_j}{8\pi\mu^0(3k^0 + 4\mu^0)\rho} \tag{2.212}$$

with $\underline{e} = \frac{\underline{x}-\underline{x}'}{\rho}$. From that, we find

$$\begin{aligned}
\Gamma^0_{ijkl} &= E^0_{ijkl}\delta(\rho) + F^0_{ijkl}(\underline{x}, \underline{x}') \\
E^0_{ijkl} &= \frac{-(3k^0 + \mu^0)\delta_{ij}\delta_{kl} + 9(k^0 + 2\mu^0)I_{ijkl}}{15\mu^0(3k^0 + 4\mu^0)} \\
F^0_{ijkl} &= \frac{-1}{8\pi\mu^0(3k^0 + 4\mu^0)\rho^3} \\
&\quad \times \Big[(3k^0 + \mu^0)(\delta_{ij}\delta_{kl} - 3e_i e_j \delta_{kl} - 3e_k e_l \delta_{ij}) - 6\mu^0 I_{ijkl} + 15((3k^0 + \mu^0) \\
&\qquad e_i e_j e_k e_l - 3/2(3k^0 + 2\mu^0)(e_j e_l \delta_{ik} + e_i e_l \delta_{jk} + e_j e_k \delta_{il} + e_i e_k \delta_{jl})\Big] \\
F^0_{iijj} &= F^0_{ijij} = 0
\end{aligned} \tag{2.213}$$

From these expressions, together with the properties just described, it is easy to find the solution within *a spherical heterogeneity* I *of moduli* $\mathbf{c}$ *embedded in an infinite homogeneous matrix of moduli* $\mathbf{C}$ *and uniform deformation* $\underline{\underline{\varepsilon}}^0$ *at infinity* (which, in the absence of the heterogeneity, would become established throughout the matrix medium). If we put $\mathbf{C}^0 = \mathbf{C}$ then $\delta\mathbf{c} = 0$ everywhere except in the inclusion and the integral equation (2.208) has a uniform solution $\underline{\underline{\varepsilon}}^{\mathrm{I}}$ given by

$$(\mathbf{I} + \mathbf{E} : \delta\mathbf{c}) : \underline{\underline{\varepsilon}}^{\mathrm{I}} = \underline{\underline{\varepsilon}}^0 \tag{2.214}$$

It is left as an exercise for the reader to show that if $\mathbf{c}$ and $\mathbf{C}$ are isotropic this reproduces Eshelby's classic result. In the case of an *ellipsoidal heterogeneity* H the result involves also the long-range part $\tilde{\mathbf{F}}$:

$$\left(\mathbf{I} + \int_{\mathrm{H}} \boldsymbol{\Gamma}(\underline{x}, \underline{x}')\mathrm{d}V' : \delta c\right) : \underline{\underline{\varepsilon}}^{\mathrm{H}} = \underline{\underline{\varepsilon}}^{0} \tag{2.215}$$

With Eshelby's tensor defined as by (2.152) and (2.153), *i.e.*,

$$\mathbf{S}^{\mathrm{Esh}} = \left(\int_{\mathrm{H}} \boldsymbol{\Gamma}(\underline{x}, \underline{x}')dV'\right) : \mathbf{C} \tag{2.216}$$

we thus have again the result already found in Sect. 2.7.3.4, taking account of the fact that the integral is independent of $\underline{x}$ when this is inside H.

2.9.3 Equation for the Effective Moduli

The effective moduli of the heterogeneous medium can be found from the solution of the general Eq. 2.209. Formally

$$\underline{\underline{\varepsilon}} = \left(\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Gamma}}^{0} : \delta\mathbf{c}^{0}\right)^{-1} \otimes : \underline{\underline{\varepsilon}}^{0}, \quad \text{where } \tilde{\mathbf{I}} = \mathbf{I}\delta(\underline{x}, \underline{x}') \tag{2.217}$$

Eliminating $\underline{\underline{\varepsilon}}^{0}$ from this equation by averaging gives the strain concentration law:

$$\begin{aligned} &< \underline{\underline{\varepsilon}} > = < \left(\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Gamma}}^{0} : \delta\mathbf{c}^{0}\right)^{-1} > \otimes : \underline{\underline{\varepsilon}}^{0} \\ &\Rightarrow \quad \underline{\underline{\varepsilon}} = (\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Gamma}}^{0} : \delta\mathbf{c}^{0})^{-1} \otimes : \left[< (\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Gamma}}^{0} : \delta\mathbf{c}^{0})^{-1} >\right]^{-1} : < \underline{\underline{\varepsilon}} > \end{aligned} \tag{2.218}$$

Multiplying by $\mathbf{c}$ and taking the average we have the effective moduli $\mathbf{C}^{\mathrm{eff}}$:

$$\mathbf{C}^{\mathrm{eff}} = < \mathbf{c} : (\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Gamma}}^{0} : \delta\mathbf{c}^{0})^{-1} > \otimes : \left[< (\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Gamma}}^{0} : \delta\mathbf{c}^{0})^{-1} >\right]^{-1} \tag{2.219}$$

Similarly, from (2.210) we would find

$$\mathbf{S}^{\mathrm{eff}} = < \mathbf{s} : (\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Delta}}^{0} : \delta\mathbf{s}^{0})^{-1} > \otimes : \left[< (\tilde{\mathbf{I}} + \tilde{\boldsymbol{\Delta}}^{0} : \delta\mathbf{s}^{0})^{-1} >\right]^{-1} \tag{2.220}$$

This can be simplified by making a particular choice for the reference medium; thus with $\mathbf{C}^{0} = \mathbf{C}^{\mathrm{eff}}$ (2.219) becomes after some algebra

$$< \delta\mathbf{c}^{\text{eff}} : (\tilde{\mathbf{I}} + \tilde{\mathbf{\Gamma}}^{\text{eff}} : \delta\mathbf{c}^{\text{eff}})^{-1} >= 0 \tag{2.221}$$

and similarly from (2.220)

$$< \delta\mathbf{s}^{\text{eff}} : (\tilde{\mathbf{I}} + \tilde{\mathbf{\Delta}}^{\text{eff}} : \delta\mathbf{s}^{\text{eff}})^{-1} >= 0 \tag{2.222}$$

The apparent simplicity of Eqs. 2.221 and 2.222 conceals considerable complexity: mathematical complexity, because they are in fact integral equations, whose solution requires the difficult convolution inversion of the operators $\tilde{\mathbf{\Gamma}}$ and $\tilde{\mathbf{\Delta}}$, the analytic form for which, as we have said, is seldom known; and physical complexity, because their solution requires full information concerning the distribution $\mathbf{c}(\underline{x})$, and this is seldom available. As we know, there are two ways in which we can try to overcome these difficulties:

– by looking for bounds for $\mathbf{C}^{\text{eff}}$ and $\mathbf{S}^{\text{eff}}$ that will take account of our incomplete knowledge of the distributions $\mathbf{c}(\underline{x})$ and $\mathbf{s}(\underline{x})$
– by making simplifying assumptions that will enable us to propose estimates by finding solutions for particular spatial distributions of the elastic heterogeneities.

We end this discussion by giving, without proof, a few of the more significant results that have been obtained by these approaches.

2.9.4 *Bounds and Estimates for the Effective Moduli*

Formally, the inverse that appears in (2.219) can be developed as a series:

$$(\tilde{\mathbf{I}} + \tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0)^{-1} = \tilde{\mathbf{I}} - \tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0 + (\tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0) \otimes : (\tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0) - \cdots \tag{2.223}$$

With the assumption that this converges – which may not be true in the case of very heterogeneous (such as porous) materials – this gives

$$\begin{aligned}\mathbf{C}^{\text{eff}} = \mathbf{C}^0 + < \delta\mathbf{c}^0 > - < \mathbf{c}' : \tilde{\mathbf{\Gamma}}^0 : \mathbf{c}' > + < \mathbf{c}' : \tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0 : \tilde{\mathbf{\Gamma}}^0 : \mathbf{c}' > \cdots \\ \cdots - < (\mathbf{c}' : \tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0)' : \tilde{\mathbf{\Gamma}}^0 : (\delta\mathbf{c}^0 : \tilde{\mathbf{\Gamma}}^0 : \mathbf{c}')' > \cdots \\ + < (\mathbf{c}' : \tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0)' : \tilde{\mathbf{\Gamma}}^0 : \delta\mathbf{c}^0 : \tilde{\mathbf{\Gamma}}^0 : (\delta\mathbf{c}^0 : \tilde{\mathbf{\Gamma}}^0 : \mathbf{c}')' > - \cdots\end{aligned} \tag{2.224}$$

where $\mathbf{c}' = \mathbf{c} - < \mathbf{c} >$, and generally $\mathbf{A}' = \mathbf{A} - < \mathbf{A} >$.

Equation 2.224 involves the successive correlation functions of increasing order of the moduli: $<\mathbf{c}>$, $<\mathbf{c}:\mathbf{c}>$, $<\mathbf{c}:\mathbf{c}:\mathbf{c}>$ and so on. If we truncate it at order n and use only correlations of order less than or equal to n, it can be shown that this gives an upper bound for $\mathbf{C}^{\text{eff}}$ that decreases with increasing n. Similarly, from (2.220)

for $\mathbf{S}^{\text{eff}}$ we get a sequence of increasing lower bounds for $\mathbf{C}^{\text{eff}}$. Note that for $n = 1$ we reproduce the Voigt/Reuss bounds, which use only first-order correlations – and therefore only the volume fractions of the different constituents.

With the additional assumption of macroscopic isotropy it is found (Kröner 1978) that if the medium satisfies the condition of *graded disorder* defined by

$$< \mathbf{c}' : (\tilde{\mathbf{F}}^0 : \mathbf{c}')^p >= 0, \quad p = 1, 2, \ldots, n-1 \tag{2.225}$$

then (2.224) gives as n order upper bounds the moduli $\mathbf{C}^{(n)}$ defined by

$$\mathbf{C}^{(n)} =< \mathbf{c} : (\mathbf{I} + \mathbf{E}^{(n-2)} : \delta\mathbf{c}^{(n-2)})^{-1} >:< (\mathbf{I} + \mathbf{E}^{(n-2)} : \delta\mathbf{c}^{(n-2)})^{-1} >^{-1} \tag{2.226}$$

and as corresponding n order lower bounds the moduli $\mathbf{C}^{(-n)}$, obtained similarly by replacing $(n-2)$ by $(-n-2)$. Here $\mathbf{E}^{(n-2)}$ denotes the local part of the modified Green's operator $\tilde{\mathbf{\Gamma}}^{(n-2)}$ for a medium of moduli $\mathbf{C}^{(n-2)}$.

Study of the expression for $\mathbf{C}^{(n)}$ shows that this involves terms that come from the solution of a problem of a spherical heterogeneity with moduli $\mathbf{c}$ in a matrix of moduli $\mathbf{C}^{(n-2)}$. Thus for $n = 2$, $\mathbf{C}^{(2)}$ is the Hashin-Shtrikman upper bound, which can be calculated from the solution to problems of spherical inclusions in an infinite matrix of moduli $\mathbf{C}^{(+0)}$ (the "stiffest" phase). $\mathbf{C}^{(-2)}$ is the corresponding lower bound, obtained with matrix of moduli $\mathbf{C}^{(-0)}$, the "softest" phase. If we started from the moduli of Voigt ($\mathbf{C}^{(1)}$) or of Reuss ($\mathbf{C}^{(-1)}$) bounds we should obtain bounds of rank 3, sharper than those of Hashin and Shtrikman and valid for a medium of disorder of rank 3, and so on. As $n \to \infty$ the sequence of estimates converges to $\mathbf{C}^{\pm\infty}$ given by

$$\begin{aligned} \mathbf{C}^{\pm\infty} &=< \mathbf{c} : (\mathbf{I} + \mathbf{E}^{\pm\infty} : \delta\mathbf{c}^{\pm\infty})^{-1} >:< (\mathbf{I} + \mathbf{E}^{\pm\infty} : \delta\mathbf{c}^{\pm\infty})^{-1} >^{-1} \\ \text{or } &< \delta\mathbf{c}^{\pm\infty} : (\mathbf{I} + \mathbf{E}^{\pm\infty} : \delta\mathbf{c}^{\pm\infty})^{-1} >= 0 \end{aligned} \tag{2.227}$$

which is Eq. 2.219 with $\tilde{\mathbf{\Gamma}}^{\text{eff}}$ replaced by its local part. Taking this together with (2.214) and (2.195), since $\mathbf{P}_{\text{Sph}} = \mathbf{E}$, it will be seen that this is strictly the self-consistent model with isotropic distribution of the phases, which thus takes on the significance of a model of a *perfectly disordered heterogeneous material.*

Thus the different bounds and models we have studied can be brought together within a general framework which enables them to be compared with one another, and their fields of applicability better defined. Note nevertheless that this approach, based on the use of point-correlation functions, is not quite appropriate to the study of composite materials for which one phase at least, the matrix, is continuous: such a morphological characteristic is hardly described by low-order point-correlation functions. For these cases, improvements on the classical Mori-Tanaka or Hashin-Shtrikman models can be achieved with treatments dealing with shell-chore-type basic morphological units instead of points, such as the Hashin "composite sphere

(or cylinder) assemblage" (Hashin 1962; Hashin and Rosen 1964), the "three-phase model" (Christensen and Lo 1979), the "representative morphological patterns" (Bornert et al.1996) or "unit cell" and periodic homogenisation approaches (Suquet 1987). The reader is invited to consult this literature for more information.

References

Bornert M, Stolz C, Zaoui A (1996) Morphologically representative pattern-based bounding in elasticity. J Mech Phys Solids 44:307–331

Christensen RM, Lo KH (1979) Solutions for effective shear properties in three phase sphere and cylinder models. J Mech Phys Solids 27:315–330

Eshelby JD (1957) The determination of the elastic field of an ellipsoidal inclusion and related problems. Proc R Soc Lond Ser A 241:376–396

Eshelby JD (1961) Elastic inclusions and inhomogeneities in Progress in solid mechanics II. In: Sneddon I.N., Hill R. (ed), vol 3. North-Holland, Amsterdam, 87–140

Ferry JD (1980) Viscoelastic properties of polymers. Wiley, New York

Flory PJ (1953) Principles of polymer chemistry. Cornell University Press, Ithaca

Flory PJ (1985) Molecular theory of rubber elasticity. Polymer J 17:1–12

Flory PJ, Rehner J Jr (1943) Statistical mechanics of cross-linked polymer networks. II. Swelling. J Chem Phys 11:521–526

Grumbell SM, Mullins L, Rivlin RS (1953) Departures of the elastic behaviour of rubbers in simple extension from the kinetic theory. Trans Faraday Soc 49:1495

Hashin Z (1962) The elastic moduli of heterogeneous materials. J Appl Mech 29:143–150

Hashin Z, Rosen BW (1964) The elastic moduli of fiber-reinforced materials. J Appl Mech 31:223–232

Hashin Z, Shtrikman S (1963) A variational approach to the theory of the elastic behavior of multiphase materials. J Mech Phys Solids 11:127–140

Higgs PG, Gaylord RJ (1990) Slip-links, hoops and tubes - Tests of entanglement models of rubber elasticity. Polymer 31:70

James HM, Guth E (1947) Theory of the increase in rigidity of rubber during cure. J Chem Phys 15:669

Kröner E (1977) Bounds for effective elastic moduli of disordered materials. J Mech Phys Solids 25:137–155

Kröner E (1978) Self-consistent scheme and graded disorder in polycrystal elasticity. J Phys F 8:2261–2267

Mandel J (1972) Plasticité classique et viscoplasticité CISM Lecture notes 97 Springer Wien

Meyer KH, Ferri C (1935) Sur l'élasticité du caoutchouc. Helv. Chim. Acta 18:570–589

Mori T, Tanaka K (1973) Average stress in the matrix and average elastic energy of materials with misfitting inclusions. Acta Metall 21:571–574

Mura T (1987) Micromechanics of defects in solids. Kluwer Academic Publishers, The Hague

Nemat-Nasser S, Hori M (1999) Micromechanics: overall properties of heterogeneous materials. 2nd rev ed, Elsevier Science B.V, Amsterdam

Salençon J (2001) Handbook of continuum mechanics General concepts and thermoelasticity. Springer, Berlin/Heidelberg/New York

Suquet P (1987) Elements of homogenization for inelastic solid mechanics in Homogenization techniques for composite media. In: Sanchez-Palencia E, Zaoui A (eds). Lecture notes in physics, vol 272, Springer, New York, 193–278

Treloar LRG (1943) The elasticity of a network of long chain molecules (I and II). Trans Faraday Soc 39:36–41

Treloar LRG (1975) The physics of rubber elasticity. Oxford University Press, Oxford
Willis JR (1977) Bounds and self-consistent estimates for the overall properties of anisotropic composites. J Mech Phys Solids 25:185–202
Zaoui A (2000) Matériaux hétérogènes et composites Editions de l'Ecole Polytechnique Palaiseau

Chapter 3
Elastoplasticity

Abstract Phenomenological descriptions are given of the elastoplastic behaviour in uniaxial loading. Monotonic and cyclic loadings are then considered. It is shown how this can be generalised in three dimensions. Plastic deformation requires the displacement of dislocations, which are described first in a continuous medium and then in crystallographic structures. The forces on dislocations and their stress and strain fields and energy are introduced. Mutual interactions between dislocations and interaction with interfaces are studied. Twinning is described. Then hardening due to interactions with the crystallographic lattice, dislocation network, grain boundaries, foreign atoms either isolated or in the form of precipitates or inclusions is described. Hardening due to twinning, martensitic transformation and transformation plasticity are described. The case of some steels is given as examples. Fibres reinforcement is introduced. The macroscopic formulation of plastic behaviour is based on single crystal plasticity and this leads to plasticity criteria considering isotropic and kinematic hardening. The limit analysis of structures is studied. The plasticity of heterogeneous materials is introduced.

3.1 Introduction

This chapter is devoted to the analysis of a type of behaviour, nonlinear by definition, shown by a great variety of materials in response to a great variety of types of loading. The main restrictions here will be the assumptions that viscosity can be neglected and that the loading does not damage the material – no cracks or voids, for example, are produced. The absence of viscosity does not necessarily mean that the temperatures are low, for if the temperature is low enough for the diffusion mechanisms to be regarded as inhibited then there will in general be a low-temperature viscoplastic regime with other phenomena to be taken into account (see Chap. 4). It is best to take the view that the validity of the elastoplasticity approximation, as treated below, considered as a limiting case of elastoviscoplasticity, holds over

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1_3, © Springer Science+Business Media B.V. 2012

a domain which, under suitable temperature conditions, adjoins both that of loads varying infinitely slowly, giving an asymptotically stable response, and that of very rapidly varying loads in which viscosity "has no time to act". Consequently, we shall treat plasticity as independent of time and strain rate – *i.e., rate-independent plasticity.* If a parameter t, called "time" for convenience, appears in a constitutive equation this will not be physical time but a purely kinematic time, enabling us to locate a sequence of events on an increasing scale; these events will usually be related to changes to the loading and will represent the load path, without regard to the rate at which any changes are made. So far as the assumption of no damage is concerned, this is an essential restriction but nevertheless a provisional one which in the end we shall lift when considering the coupling between plasticity and damage (see Volume II). Finally, to avoid unnecessary complications, we shall consider the elastic behaviour as linear; this will enable us to attribute all nonlinear aspects unambiguously to plasticity.

A final restriction, and not the least, is that we shall limit our treatment to small transformations. Finite transformations in elastoplasticity, especially when the elastic deformations themselves cease to remain small, give rise to complex problems that are important but difficult both to formulate and to analyse; these lie at the heart of much current research, and we leave the reader to consult specialist works on the subject.

The plan of the chapter is this:

1. A short introduction, summarising the most significant phenomenological aspects of elastoplasticity and the main questions raised by its description and its study.
2. A treatment of the most important physical aspects of elastoplastic behaviour, especially for crystalline materials, which up to the present have provided the material for the deepest studies of the microscopic mechanisms of plasticity.
3. A transition section on the problems raised by upscaling treatments from microscopic to macroscopic plasticity
4. A final section formulating elastoplastic behaviour within the framework of macroscopic rheology, with an indication of methods for carrying out calculations for elastoplastic structures and of the "heterogeneous materials" approach in plasticity.

3.2 General: Phenomenological Aspects

The study of the response of a sample, which, starting from an initial state that can be taken as reference, is subjected to a variety of tension/compression loadings and for which viscosity and damage can be neglected, will reveal the main phenomenological aspects of elastoplasticity. On the basis of such observations we can formulate the main problems that a local three-dimensional representation will have to solve.

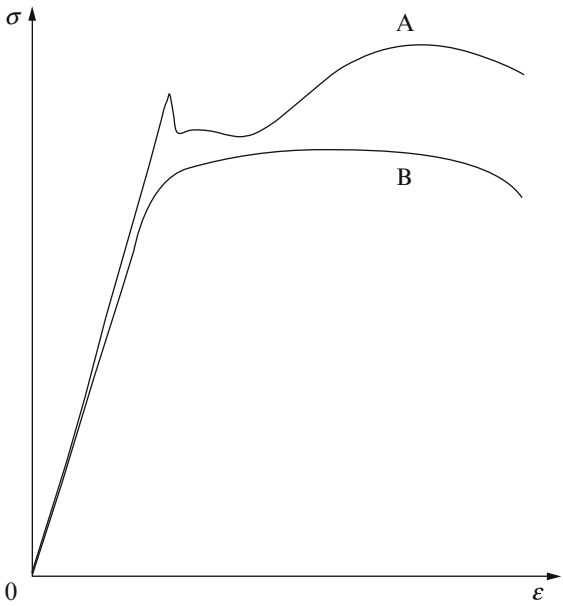


Fig. 3.1 Stress-strain curves

3.2.1 One-Dimensional Response

3.2.1.1 Yield Strength

The overall response of a sample to a monotonic load can be represented by a (σ,ε) diagram as in Fig. 3.1 for tension, recording the true values of stress and strain. This will show the yield strength; whether the curve has a discontinuity, as does A in the diagram, or remains smooth, as does B, we have to define in a conventional manner the stress $\sigma_{0,}$ or R_e, at which plasticity begins to appear. Since the departure from linear response is difficult to establish in a reproducible manner it is usual to record the axial stress R_p corresponding to an irreversible (*i.e.*, permanent) axial deformation of some stated amount, which is currently taken as 0.2%. The load has to be removed to show the permanent deformation ε^p, and the total deformation ε can then be expressed as the sum of a reversible elastic part ε^e and an irreversible or plastic part ε^p, namely $\varepsilon = \varepsilon^e + \varepsilon^p$.

Starting from the same initial state, compression would give a corresponding compressive yield stress σ_0', so defining an interval of elasticity $[\sigma_0', \sigma_0]$.

3.2.1.2 Work-Hardening

If the tensile test is continued up to a moderate axial plastic deformation ε_1^p with corresponding stress σ_1, and the stress then reduced to zero, this unloading will take

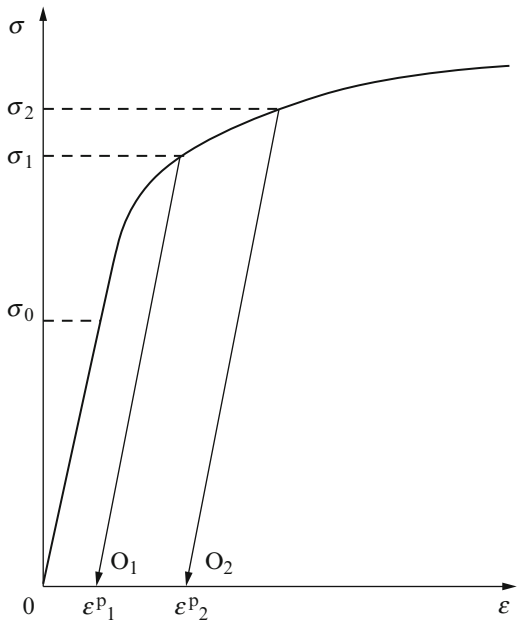


Fig. 3.2 Load-unload sequences

place elastically with a Young's modulus which can be regarded as remaining the same as the original modulus; the effect is to return to the point O_1 $(0, \varepsilon_1^p)$ of Fig. 3.2. The state represented by O_1 is stable, and a second experimenter setting out to find the tensile yield stress in this state will record not σ_0 but σ_1. If he had worked with compression he would have found a value σ_1' and thus an elastic interval $[\sigma_1', \sigma_1]$; and similarly $[\sigma_2', \sigma_2]$ if he had started from state O_2, etc.

This is expressed by the term *work-hardening*, meaning, for uniaxial stressing, the effect of plastic flow on the elastic interval, changing both its width and the positions of its boundaries. The perfect plasticity scheme neglects this change; this is what one can do for example, in the region of the plateau in curve A of Fig. 3.1. In general, it is helpful to be able to specify this evolution in the case of a work-hardenable material, that is, to give the limits of the elastic interval for each state of the sample.

How are we to define this *state* and to determine the elastic intervals for each one of the states that arises? Figure 3.2 suggests representing the state O_1 by the plastic strain ε_1^p. In general, hardening increases in the direction of the flow ($\sigma_1 > \sigma_0$) and decreases in the opposite direction ($|\sigma_1'| < |\sigma_0'|$); this is the *Bauschinger effect*, often observed in metallic materials. However, this parametrisation of work-hardening by plastic deformation is not very satisfactory, for

- we may not know the value of the parameter when we perform the test on a sample whose previous history we do not know;
- more seriously, we may find that different widths of the interval correspond to the same value of the plastic strain.

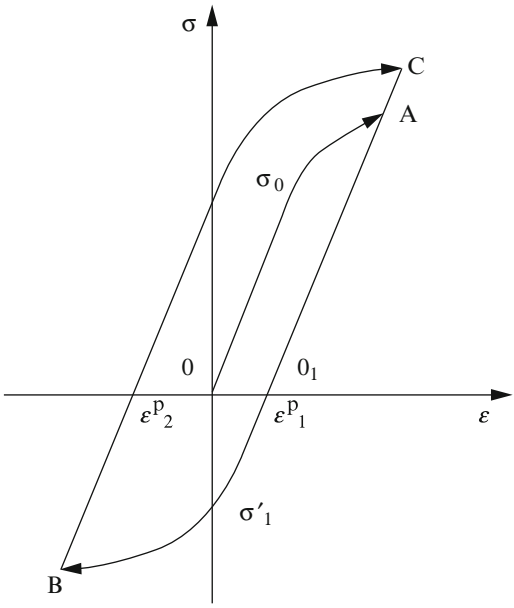


Fig. 3.3 Cyclic loading

Thus in Fig. 3.3, the path OAO_1BCO_1 returns us to the "state O_1" defined by ε_1^p, with different (tensile) elastic limits for the first passage (σ_A) and the second (σ_C).

So, there is no simple answer to our question, and in fact it represents one of the major difficulties of plasticity. Here we might consider parametrisation by a cumulative plastic deformation p, the sum of the absolute values of the variations of the plastic strain: $p_A = \varepsilon_1^p$, $p_C = \varepsilon_1^p + 2(\varepsilon_1^p - \varepsilon_2^p)$; but this can be vitiated by more complex histories of loadings, such as cyclic loadings which cause a great variety of responses.

3.2.1.3 Cyclic Work-Hardening

Among the different materials, and also, for a given material, according to its range of loading, there is a great variety of responses to cyclic loadings, whether these are applied as stresses between limits σ_{min} and σ_{max} (with $\Delta\sigma = \sigma_{max} - \sigma_{min}$) or strains between limits ε_{min} and ε_{max} (with $\Delta\varepsilon = \varepsilon_{max} - \varepsilon_{min}$). The response to an imposed stress may become stable, settling down to a steady limit cycle – the case of *shakedown*, or, if the cycle is reduced to a line segment, *adaptation*; if not, the deformation increases at each cycle and produces what is called the *ratchet* effect. For imposed strains the first would lead to the average stress becoming constant at the value $\sigma = (\sigma_{max} + \sigma_{min})/2$, the second to this decreasing steadily.

If the cycle is alternating (*i.e.,* symmetric: $\sigma_{min} = -\sigma_{max}$ or $\varepsilon_{min} = -\varepsilon_{max}$), and a stabilised response is reached, we may see, by reference to the response to a monotonic loading, either a softening or a hardening cyclic effect. In the first case

(softening), if the cycle is of imposed stresses, the strain amplitude $\Delta\varepsilon$ will increase and if it is of imposed strains the stress amplitude $\Delta\sigma$ will decrease, both until the stable limit cycle is reached; with the opposite effects in the second case (hardening) (see also Volume II). If the stable response cycle is symmetric about the origin we can link the semi-amplitudes of stress and of plastic strain and obtain, for a range of levels of loading, a graph of *stable work-hardening cycles*; and comparison with the corresponding curve for monotonic loading will give a direct indication of the presence of cyclic softening or cyclic hardening.

These are just a few indications of the variety of responses that can be found within a single class of behaviour; a variety that any model aiming at generality, even if simply phenomenological, must be able to represent without needing to be completely reformulated for each new case.

3.2.1.4 Laws of Plastic Flow and Work-Hardening

Describing plastic flow is, in principle, a completely different question from that of work-hardening since plastic strain, or any quantity directly related to this, has not *a priori* to be taken as a work-hardening variable. Thus we have to develop the two sets of laws in parallel: those for the work-hardening parameters, from which the way in which the plastic thresholds (in fact, the limits of the elastic interval) change in the course of the loading, and those for the plastic flow, from which the amount of plastic strain can be found at each step.

The two questions are closely related:

- on the one hand, it is usual to take the plastic strain ε^{p} or the cumulative strain p derived from this as one of the parameters of work-hardening, with the result that the law of plastic flow is often assimilated into one of the laws concerning the evolution of the work-hardening parameters;
- on the other, and even if the first does not apply, the separate laws of flow and work-hardening, taken globally, have features in common which distinguish both from those of elastic or viscoelastic behaviour.

We can state two of these common features now:

- In both cases the laws are *incremental*, meaning that they do not give a direct relation between the current values of the parameters concerned (those of work-hardening or plastic deformation) at a given step and the stresses at that step, but only their increments from the state at that step. It is clear (from Fig. 3.3, for example) that the same value of one of these parameters can correspond to an unlimited number of levels of the stresses; thus one cannot expect, as one can for the case of elasticity, closed-form equations relating stresses to strains.
- Further, both sets of laws are *multi-branched*, requiring different expressions according to the increment considered. At the least one has to consider separately the cases of *plastic loading*, in which both sets of parameters vary, and of *elastic unloading* if the conditions up to then were those of plastic flow, or *elastic response* if one is already in the elastic regime, when neither varies.

We have been discussing certain features of elastoplastic behaviour in the context of one-dimensional response; we now take up the study in more detail, in both local and three-dimensional contexts, with the aim of developing a complete formulation of the laws of elastoplastic behaviour.

3.2.2 Three-Dimensional Behaviour

Instead of studying the overall response of a sample to a tension/compression cycle we now set out to describe the elastoplastic behaviour of a material volume element subjected to a stress state $\underline{\underline{\sigma}}$. We no longer have to deal with only a pair of global parameters (σ,ε); the load is now represented by six parameters, the components of the symmetric stress tensor $\underline{\underline{\sigma}}$, so we have to generalise and re-formulate the questions studied above.

3.2.2.1 Plastic Criterion, Load Function

The volume element will have some initial state that we can regard as given. The concept of yield strength (for tension or compression) must be replaced by that of a *criterion for plasticity*: what stress states, corresponding to all possible modes of loading, can induce the plastic regime? The question is answered by finding a scalar-valued function $f\left(\underline{\underline{\sigma}}\right)$ of the stress tensor which, by convention, is negative when and only when the response to the load is purely elastic, and vanishes when the plastic regime is entered. This, called the *load function* or *yield function*, is used to construct a criterion for plasticity, based on the sign of its value; we shall see later that positive values have no physical significance.

The concept of elastic interval $[\sigma_0', \sigma_0]$ is replaced by that of *elastic domain*, which can conveniently be represented geometrically in the stress space. This is a 6-dimensional space in which, with some suitable reference system, a point represents a particular stress tensor. The elastic domain is the interior of the surface $f(\underline{\underline{\sigma}}) = 0$, called the *load* (or *flow*) *surface*. Starting from any interior point, any of an unlimited number of paths in the space will lead to the plastic regime when, but only when, it reaches the surface.

3.2.2.2 Work-Hardening

We have seen how the yield strength for a work-hardenable material can be altered by plastic flow, whether tensile or compressive; we must now investigate how the load surface, and therefore the load (yield) function, can change during plastic flow, the point representing the load remaining on the surface. For this we have to make this function depend on other variables in addition to the stress components – in fact, on the work-hardening parameters, which can be scalars, vectors or tensors

according to the case, and which we denote by Y_m, $1 \leq m \leq \mathrm{M}$. These act as internal parameters or internal variables in a formulation in terms of hidden variables, which could be developed in the framework of the thermodynamics of irreversible processes.

Let $\mathrm{f}(\underline{\underline{\sigma}}, Y_m)$ be the load function, taking account of work-hardening; it can be used to define the plastic criterion: as long as $\mathrm{f}(\underline{\underline{\sigma}}, Y_m) < 0$, with the values of the Y_m corresponding to the current state of work-hardening, there is no plastic response. The latter can appear only when $\mathrm{f} = 0$, and when this occurs the Y_m change as the plastic flow develops in such a way that f remains zero. It is thus not possible for f to become positive, for, starting from negative, a positive value would have had to be preceded by zero and this would have been maintained while the variables $\underline{\underline{\sigma}}$ and Y_m changed in the course of the flow.

The changes to the load surface with plastic flow can be complex. In general there will not be just a simple shift of the surface as a whole, following the load point, but a complete transformation in which any change of shape is possible. Further, the plastic threshold does not change only in a direction that is either the same as or opposite to that of the load, as in simple tension/compression, but can change in the direction of any or all of the load components. The shape changes of the load surface are determined by the way in which f depends on the parameters Y_m, and thus by the laws governing the variation of these parameters as the flow proceeds.

Starting from a given shape of the load surface $\mathrm{f}(\underline{\underline{\sigma}}) = 0$, various simple evolutionary schemes for reference models of work-hardening can be devised; the following are those most commonly used at present.

(a) *Isotropic*. This assumes a simple dilatation of the load surface with respect to the origin; only one parameter Y is needed ($\mathrm{M} = 1$) and we have

$$\mathrm{f}(\underline{\underline{\sigma}}, Y) = \mathrm{f}(\underline{\underline{\sigma}}) - \mathrm{h}(Y) \tag{3.1}$$

where, for positive work-hardening, Y is positive and h is an increasing function.

(b) *Linear kinematic*. The change is a simple translation, without rotation or deformation. A tensorial parameter $\underline{\underline{X}}$, say, of the same order as $\underline{\underline{\sigma}}$, is needed and the load function is

$$\mathrm{f}(\underline{\underline{\sigma}}, \underline{\underline{X}}) = \mathrm{f}(\underline{\underline{\sigma}} - \underline{\underline{X}}) \tag{3.2}$$

After plastic flow the point $\underline{\underline{X}}$ will be in the same position with respect to the load surface as the origin was with respect to the initial surface. If the loading is uniaxial, as in simple tension/compression, this implies a strong Bauschinger[1] effect (equal amplitudes of hardening in one direction and softening in the opposite), which is not in accord with isotropic work-hardening (equal hardening in both directions). But since the compressive plastic response would be obtained by translation of the

[1] Johann Bauschinger (1834–1893) was a German mathematician. He taught mechanics at the Technical University of München.

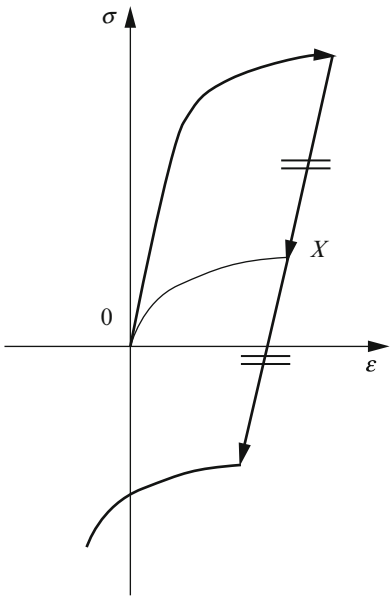


Fig. 3.4 Kinematic work-hardening

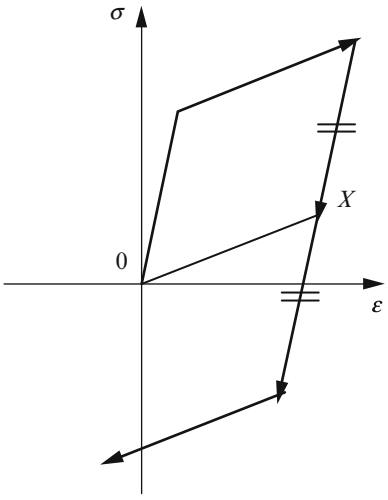


Fig. 3.5 Linear kinematic work-hardening

tensile response, we should find forms of the cycle that bear little resemblance to what is found in practice if we did not assume the work-hardening to be linear (*cf.* Fig. 3.4); this is why this simple model of kinematic work-hardening is used with the linear assumption (Fig. 3.5).

Starting with these two extremes, we can imagine various extensions, such as

- nonlinear kinematic
- a combination of isotropic and kinematic (simultaneous dilatation and translation of the load surface)

Comparing what it predicts for the response to a cyclic loading with what is observed in practice can test the validity of any model. We return to this in Sect. 3.5.2.4.

3.2.2.3 Laws of Plastic Flow and Work-Hardening

Having chosen the work-hardening parameters and the incorporation of these into the load function, we have now to determine the laws that describe their evolution, as well as those for plastic flow. We continue to treat the total deformation of the material as the sum of an elastic and a plastic term (or at least treat the increments so, writing $\mathrm{d}\underline{\underline{\varepsilon}} = \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{e}} + \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}$); taking account of what we have said concerning the 1-dimensional case, in the plastic regime we have equations of this form:

$$\begin{aligned} \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} &= \mathbf{A}(\underline{\underline{\sigma}}, Y_m) : \mathrm{d}\underline{\underline{\sigma}} \quad \text{or} \quad \mathrm{d}\varepsilon_{ij}^{p} = A_{ijkl}(\sigma_{pq}, Y_m)\mathrm{d}\sigma_{kl} \\ \mathrm{d}Y_m &= \mathbf{K}(\underline{\underline{\sigma}}, Y_m) : \mathrm{d}\underline{\underline{\sigma}} \quad \text{or} \quad \mathrm{d}Y_m = K_{mkl}(\sigma_{pq}, Y_n)\mathrm{d}\sigma_{kl} \end{aligned} \tag{3.3}$$

and in the elastic regime $\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = 0$ and $\mathrm{d}Y_m = 0$.

We have already seen that a necessary condition for plastic flow is $\mathrm{f}(\underline{\underline{\sigma}}, Y_m) = 0$; but this is not sufficient, for if at a point on the load surface an increment $\mathrm{d}\underline{\underline{\sigma}}$ is imposed which returns the point to the interior, no plastic flow develops. If the surface is regular and has an outward-pointing normal that is uniquely defined, parallel to and in the same direction as $\partial \mathrm{f}/\partial\underline{\underline{\sigma}}$, this can occur if $\partial \mathrm{f}/\partial\underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\sigma}} < 0$. Thus an increment $\mathrm{d}\underline{\underline{\sigma}}$ at a point $\underline{\underline{\sigma}}$ on the load surface will lead to plastic flow if and only if

$$\mathrm{f}(\underline{\underline{\sigma}}, Y_m) = 0 \quad \text{and} \quad \partial \mathrm{f}/\partial\underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\sigma}} > 0 \tag{3.4}$$

There are two distinct situations in which plastic flow does not occur:

$$\begin{aligned} \mathrm{f}(\underline{\underline{\sigma}}, Y_m) < 0 \quad \forall\, \mathrm{d}\underline{\underline{\sigma}} \quad \text{(elastic regime)} \\ \mathrm{f}(\underline{\underline{\sigma}}, Y_m) = 0 \quad \text{and} \quad \partial \mathrm{f}/\partial\underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\sigma}} < 0 \quad \text{(elastic unloading)} \end{aligned}$$

The case $\partial \mathrm{f}/\partial\underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\sigma}} = 0$ corresponds to a neutral loading, with no new plastic flow for a work-hardenable material.

The differential $\mathrm{d}_Y \mathrm{f} = \partial \mathrm{f}/\partial\underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\sigma}}$ that appears in these expressions is simply a partial differential of f with the parameters Y_m held constant. During plastic flow the total differential is zero since f keeps its zero value, and therefore in these circumstances

$$\mathrm{df} = \partial \mathrm{f}/\partial\underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\sigma}} + \partial \mathrm{f}/\partial Y_m : \mathrm{d}Y_m = 0 \tag{3.5}$$

which provides a "consistency condition" for the laws that describe the variations of the work-hardening parameters during plastic flow; this condition can be written as follows:

$$\left(\frac{\partial \mathrm{f}}{\partial \sigma_{kl}} + \frac{\partial \mathrm{f}}{\partial Y_m} K_{mkl}\right) \mathrm{d}\sigma_{kl} = 0, \quad \forall \mathrm{d}\sigma_{kl}, \quad \frac{\partial \mathrm{f}}{\partial \sigma_{kl}} \mathrm{d}\sigma_{kl} \geq 0, \quad \mathrm{f} = 0 \tag{3.6}$$

For perfect plasticity (no hardening) the law of plastic flow is simply

$$\frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} : \mathrm{d}\underline{\underline{\sigma}} = 0 \tag{3.7}$$

Several other conclusions can be drawn from these equations; we shall return to them in Sect. 3.5.2, considering particular forms of the laws. For the moment we just state that for a complete description of plasticity three categories of information must be available together:

- the load function $\mathrm{f}(\underline{\underline{\sigma}}, Y_m)$
- the flow functions $\mathbf{A}(\underline{\underline{\sigma}}, Y_m)$
- the work-hardening functions $\mathbf{K}(\underline{\underline{\sigma}}, Y_m)$.

These are not completely independent, but none of them can be assimilated into any of the others. One of the major concerns in the coming study of the physical mechanisms of plastic flow is to minimise the number of independent items of information needed for a complete specification of elastoplastic behaviour, giving additional relations between these three functions.

3.3 Physical Mechanisms of Plasticity

3.3.1 The Problem

In this chapter we shall see that plastic deformation in a crystalline material results from the movement of defects in the lattice – that is, of *dislocations* and also of *twins*. From a study of the various obstacles, which they have to surmount, it is possible to estimate the stress that has to be applied in order to generate a plastic deformation. Such a study will lead to explanations of the differences in hardness between covalent materials and metals, of the nature of work-hardening and of the influence of grain size and that of foreign atoms in solid solutions or in the form of precipitates or inclusions. We shall also study mechanical *twinning*, a common deformation mechanism, and *transformation plasticity*, which is very important in connexion with residual stresses and the particular phenomena of *shape memory*. All these are matters of great practical importance; they are fundamental for the development of materials with high yield strength, for specifying the behaviour of

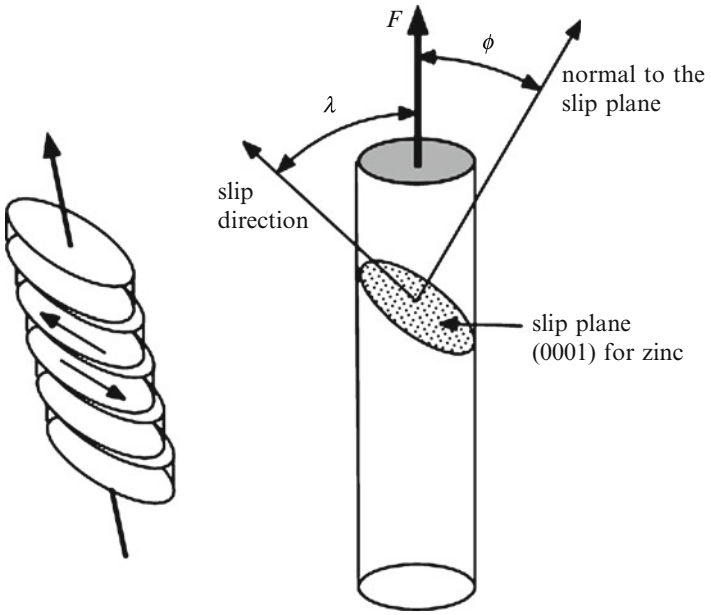


Fig. 3.6 Deformation of a hexagonal close packed single crystal by slip along the basal plane

materials undergoing forming by plastic deformation and for understanding how some properties evolve with time.

First of all, we need to relate what can be observed when deforming single crystals and compare their behaviour with theoretical predictions. This will show why it is necessary to consider the intervention of defects.

3.3.2 Deformation of a Single Crystal. Discrepancy Between Experiment and Theory

3.3.2.1 Schmid Law

We first begin with not too complicated a behaviour. Observation shows that the plastic deformation of a single crystal of the hexagonal metal zinc is due to slip along the basal (0001) planes (refer to Sect. 3.3.5.2). Figure 3.6 shows a tensile test piece in which these planes are inclined with respect to the tensile axis. It deforms like a pack of playing cards.

More generally, crystalline solids deform plastically at low temperatures by slip along the densest crystallographic planes. Zinc is a particularly simple case because a single slip plane dominates. The situation is more complicated for face-centred cubic materials (Fig. 3.7) (refer to Sect. 3.3.4) because here there are 12 equivalent

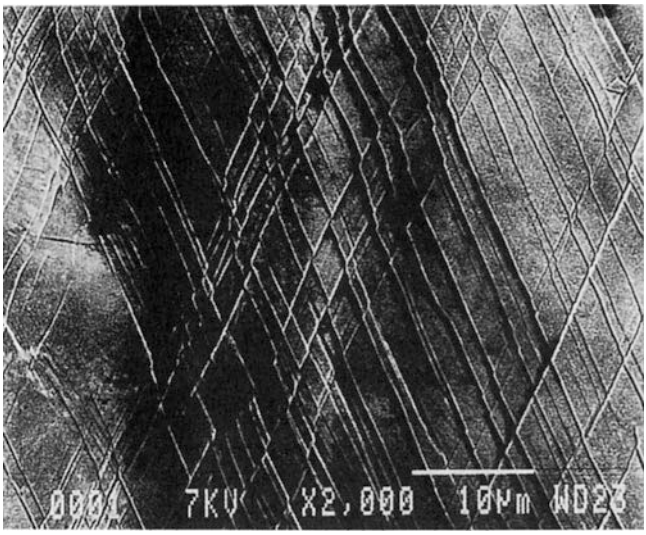


Fig. 3.7 Slip lines in a single crystal of copper oriented for double glide

slip systems – that is, systems characterised by one plane and one direction – and more complicated still for body-centred cubic materials, for which the slip in a direction is associated with a number of planes, all including a <111> direction (refer to Sect. 3.3.5.1), gives rise to what is called *pencil glide*: the deformation resembles that of a stack of hexagonal shaped pencils. Table 3.1 identifies these planes for a number of materials, and gives the directions in which the slip occurs.

We are interested in the critical conditions necessary for slip to occur. It is clear that to produce slip along a particular plane there must be a shear force acting on that plane of at least some critical value; this is the Schmid-Boas law[2] (Schmid and Boas 1935):

> Slip begins when the resolved part of the shear stress over a crystallographic slip plane in the slip direction reaches a critical value

To return to the case of zinc, it follows that the tensile yield strength σ_p for a single crystal for which the tensile axis makes an angle ϕ with the normal to the slip plane and an angle λ to the direction of slip (*cf.* Fig. 3.6) satisfies

$$\sigma_p \cos\phi \cos\lambda = \tau_c \tag{3.8}$$

[2]Erich Schmid (1896–1983) was an Austrian physicist. Walter Boas (1904–1982) was a German-Australian physicist.

Table 3.1 Slip systems and yield strengths for various single crystals

Element	Structure	Slip system at ambient temperature		Critical shear stress τ_c MPa	Theoretical shear stress τ_0 MPa
–		Plane	Direction	–	
Cu	FCC	{111}	<110>	1.0	51,500
Ag				0.6	32,000
Au				0.92	11,300
Al					
Ni				5.8	16,600
Pb					
Fe	BCC	{110},	<111>	–	80,800
V		{112},			
Nb		{123},			
		Planes in zone			
Ta		with <111>			
...					
Mg	Hex.	{0001}	$<11\bar{2}0>$	0.83	3,000
Zn				0.94	6,200
Cd				0.58	4,600
Ti, Zr		{1120}			
NaCl	Cubic	{110}	$<1\bar{1}0>$	0.75	2,840
LiF					
MgO					

Owing to the number of slip systems in BCC, the determination of the critical shear stress is difficult. The published results are scattered

Figure 3.8 gives some experimental results that support this relation. The quantity

$$m_S = \cos\phi \cos\lambda$$

is called the *Schmid factor*.

For an FCC single crystal of copper, according to its orientation, either only one slip plane can be active (single slip) or two, or even more, for example if the tensile axis coincides with a <100> direction (Fig. 3.9).

Simple geometrical considerations show that the orientation of a single crystal is changed by slip along the crystallographic planes. The normal to the slip plane tends to become perpendicular to the tensile axis. In simple slip this rotation reduces the shear component and thus produces a consolidation of a purely geometrical nature.

The yield strength for various single crystals is listed in Table 3.1.

3.3.2.2 Theoretical Critical Shear Stress

Consider now the calculation of the theoretical *critical shear stress* τ_0 necessary to result in plastic slip. It is clear from the periodic structure of a single crystal that

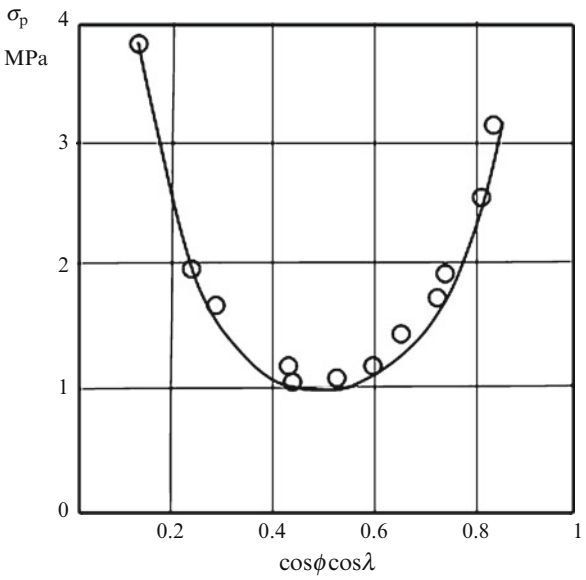


Fig. 3.8 Variation of the yield strength of a single crystal of magnesium with the Schmid's factor. The curve is the Schmid's law

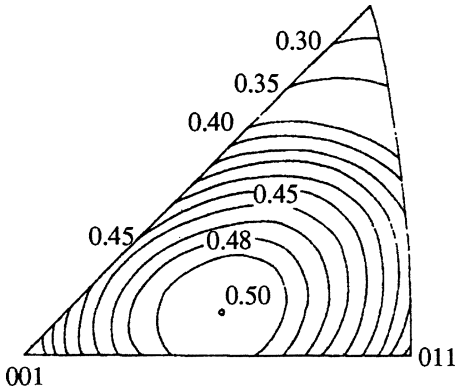


Fig. 3.9 Contours for the Schmid factor for a face-centred cubic material

the energy of the crystal is a periodic function of the slip distance (Fig. 3.10). In the absence of stresses the atoms position themselves so as to minimise this energy; and if the applied shear is small they will return to this position when it is removed: this is elastic behaviour. But if the shear exceeds a certain value an energy barrier will be crossed and the plane will slip one inter-atomic distance in returning to a minimum-energy position; if the shear is then removed a residual slip will remain and the behaviour will be plastic. To calculate the yield strength note that the shear is the derivative of the energy with respect to slip; this too is periodic, with period equal to the inter-atomic distance b, and the first term in its Fourier series development can be written:

$$\tau = \tau_0 \sin(2\pi x/b) \tag{3.9}$$

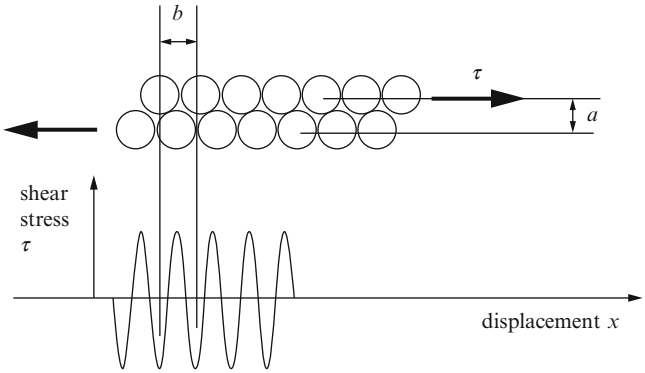


Fig. 3.10 Theoretical shear stress model for slip

where τ_0 is the amplitude, equal to the theoretical yield strength, and x is the displacement of one part of the crystal with respect to the other.

This can be written in terms of shear $\gamma = x/a$, where a is the distance between the slip planes. If the slip distance is small the behaviour is elastic, and as $\gamma = \tau/\mu$, where μ is the shear modulus, it follows that:

$$\tau_0 = \mu b/2\pi a \tag{3.10}$$

which means that τ_0 is of the order of $\mu/10$.

The yield strengths for single crystals given by this method are several orders of magnitude greater than those measured: Fig. 3.10 and Table 3.1 give the values for a number of materials, including some (amorphous materials and certain polymers) to which the method of calculation clearly would not apply (Fig. 3.11).

3.3.2.3 Introduction of Dislocations

This complete disagreement between theory and experiment led Taylor[3] in the 1930s to suggest that slip was the result of movement of defects in the lattice, that is, of dislocations, as described mathematically by Volterra[4] in 1905. These defects, introduced on purely theoretical grounds, were observed when electron microscopy was developed (Fig. 3.12 as an example). It has since been shown that for a single crystal with no dislocations, such as a trichite or *whisker*, the yield stress is indeed

[3]Geoffrey Ingram Taylor (1886–1975) was an English physicist. Egon Orowan (1902–1989) and Michael Polanyi (1891–1976), Hungarian-British physicists, proved the necessity of the intervention of dislocations to explain plastic deformation at about the same time.

[4]Vito Volterra (1860–1940) was an Italian mathematician.

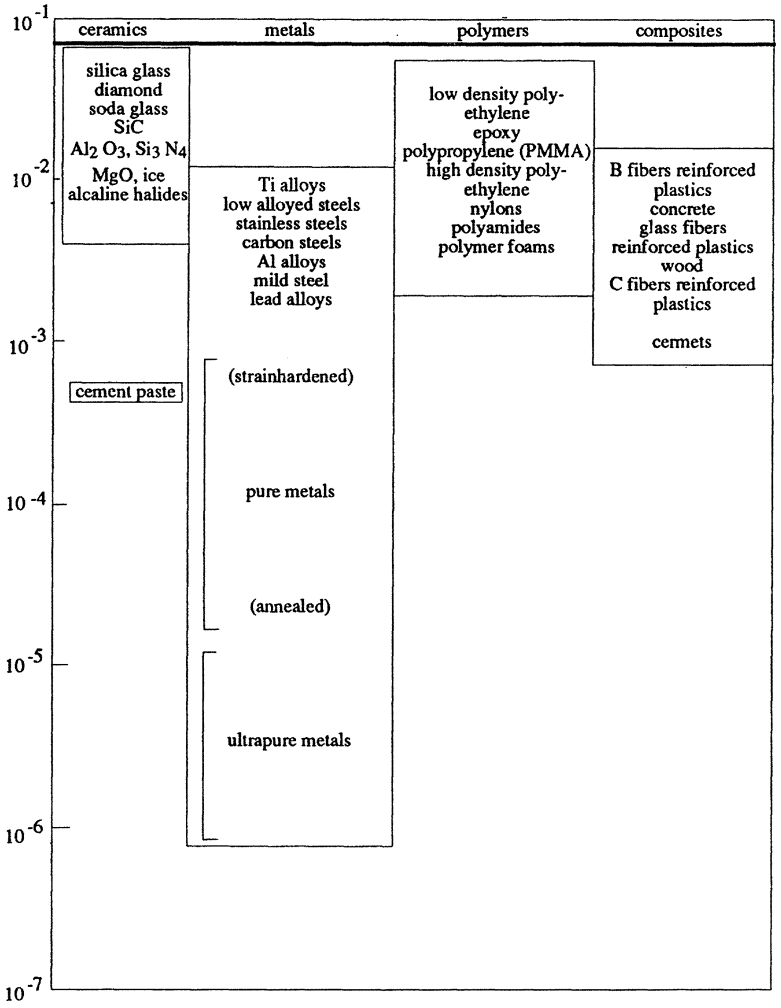


Fig. 3.11 Ratio of yield strength to Young modulus for various materials

of the order of $\mu/10$ (see Table 3.2). Whiskers are filamentary single crystals a few microns in diameter, first obtained by growth at a surface under particular conditions of oxidation (Fig. 3.13).

Let us imagine, with Taylor, that there is a defect in the lattice at the level of the slip plane, as in Fig. 3.14. The movement of this, called a dislocation, will finally produce a step of height b on the surface, and thus the traverse of the whole single crystal by such a defect will generate a slip of amplitude b. The energy demand is only that needed to break the inter-atomic links in a single row of the lattice, whereas the previous calculation assumed that all the links in the whole slip plane had to be broken at the same time. Thus the energy, and consequently the stress, are much less

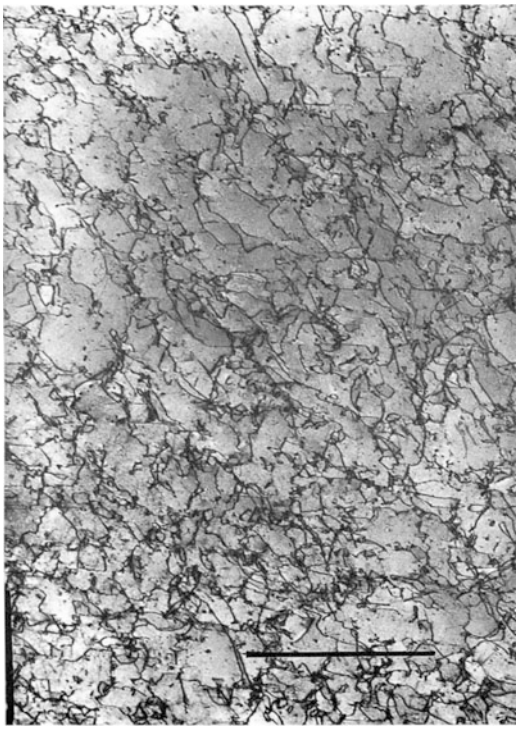


Fig. 3.12 Electron microscopy view of dislocations in an austenitic stainless steel. (The bar is one micron long)

Table 3.2 Mechanical properties of whiskers

–	Graphite	Al_2O_3	Fe	SiC	Si	AlN	Cu
R_m (GPa)	19.6	15.4	12.6	20–40	7	7	2
E (GPa)	686	532	196	700	182	150	192

than the calculated values. Mott[5] gave a striking picture of the effect by remarking that a rug laid over a carpet is easily moved by pushing a wrinkle along (Fig. 3.15).

A way to visualise dislocations is to look at a bubbles raft such as the one shown in Fig. 3.16 in which there are often such a defect. It can be seen by looking at the oblique bubbles alignments from the lower right of the picture. Natural objects, such as corncobs, zebra stripes, sunflowers display edge dislocations. Caterpillars and worms use dislocations to crawl.

The general situation is that plastic deformation is intimately related to the movement of dislocations; we now study in detail the geometrical properties of dislocations and the consequences of their displacements.

[5]Sir Nevill Mott (1905–1996) was a British physicist who won the Nobel Prize.

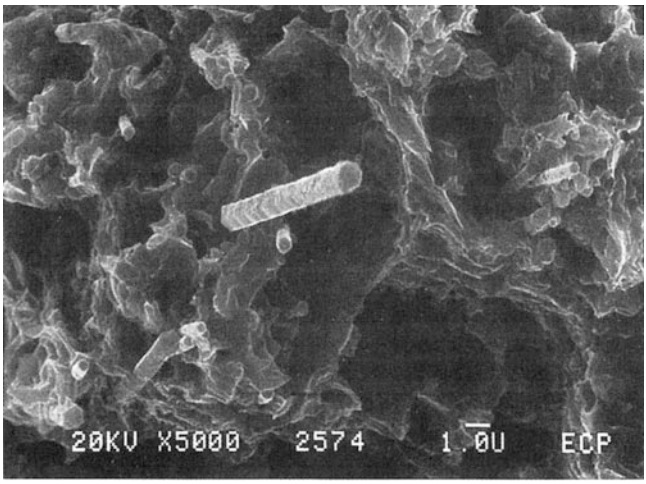


Fig. 3.13 Whiskers of silicon carbide sticking out of the fracture surface of an aluminium matrix composite

3.3.3 Dislocations: Definition, Geometrical Properties

3.3.3.1 The Volterra Dislocation

The mathematician Volterra studied the residual stresses in a solid resulting from a dislocation. The diagram of Fig. 3.17 represents the way he defined it.

Imagine, in a continuous solid, a cut area A defined by a closed boundary curve l, and the two surfaces of which are moved relative to each other by a vector $\underline{b}$, the same all over the area, called the Burgers vector: this will require the application of some force system.[6] Suppose now that the discontinuity is removed either by filling the resulting empty volume with the same material or, if material had had to be removed in making the displacement, by sticking the surfaces of the cut together; this will allow the applied forces to be removed, leaving an internal force system. It can be shown that this internal force system is independent of the area A and depends only on the curve l and the vector $\underline{b}$. The curve l is a dislocation characterised by its Burgers vector $\underline{b}$.

It follows from this definition that a dislocation is either a closed path or a path that ends either at the surface of the solid or at another dislocation (that is, at a *node*) (Fig. 3.18); and in the latter case there is the relation $\underline{b}_1 + \underline{b}_2 + \underline{b}_3 = 0$, just as for the currents in the branches of an electrical network. This last relation implies *a precise*

[6] Johannes (Jan) Martinus Burgers (1895–1981) was a Dutch physicist.

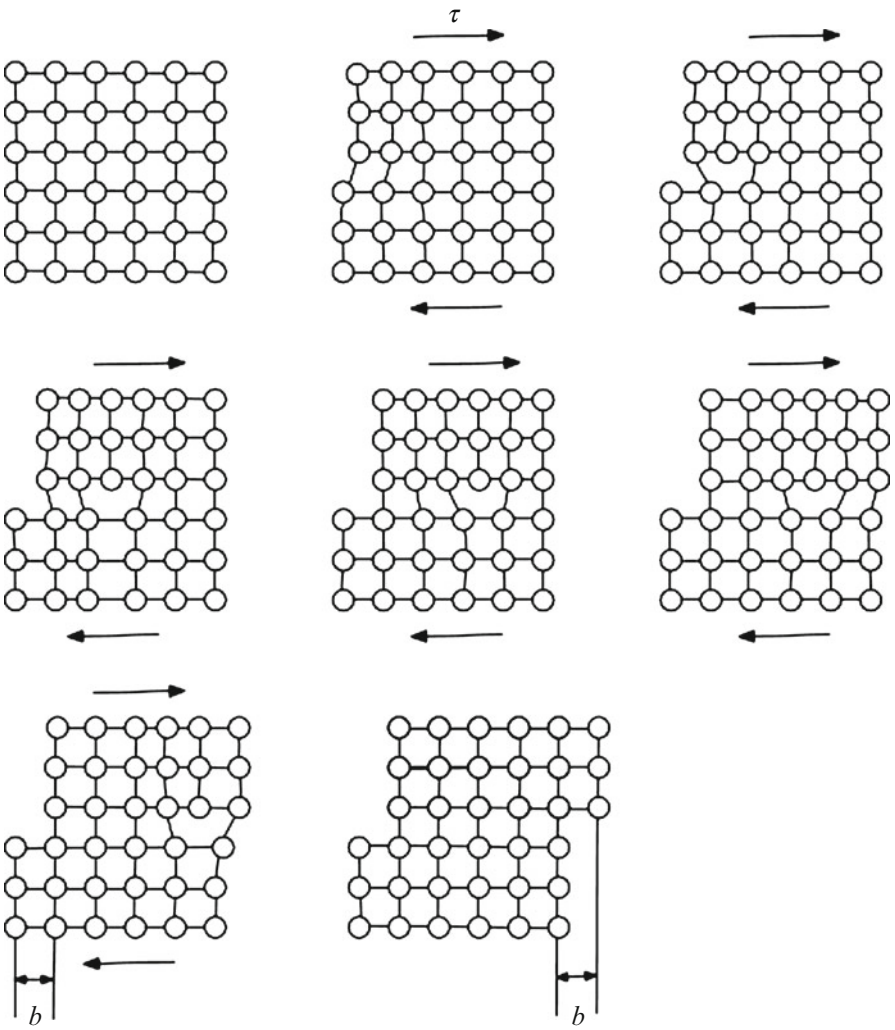


Fig. 3.14 Slip resulting from the movement of a dislocation

definition of the orientation of the dislocation and of its Burgers vector. When the direction in which the path l is followed has been chosen, the Burgers vector is the vector joining the starting to the finishing points of what was a direct closed circuit AB, surrounding the line l, that has been made open by the cut (Fig. 3.19).

Dislocations are joined together by nodes and they form a three-dimensional net, the Frank's network (Fig. 3.20).[7]

[7]Frederick Charles Frank (1911–1998) was a British physicist.

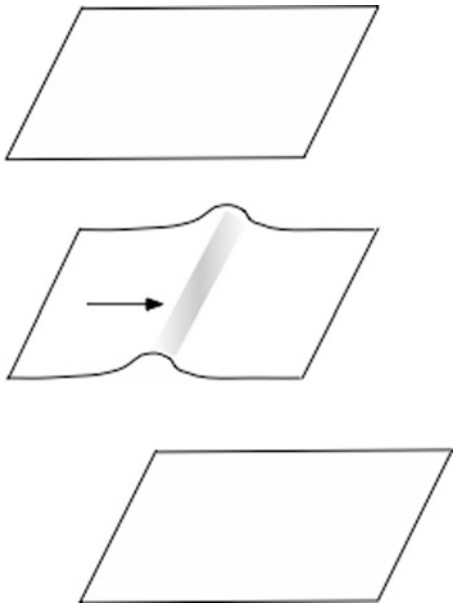


Fig. 3.15 Mott's rug analogue

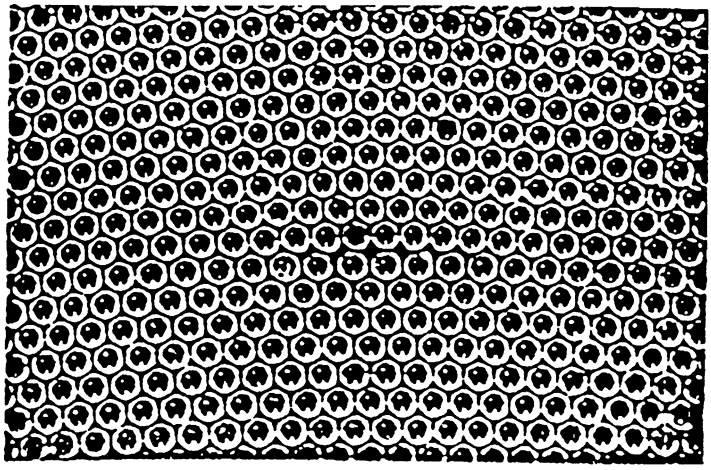


Fig. 3.16 Dislocation in a lattice of bubbles

3.3.3.2 Screw and Edge Dislocations

A *screw dislocation* is one for which the Burgers vector is parallel to the line l; it is so called because a circuit round the dislocation is transformed into a helix (Fig. 3.21).

An *edge dislocation* is such that the Burgers vector is normal to the line; it is so called because it can be created by inserting an "edge" of the material into the

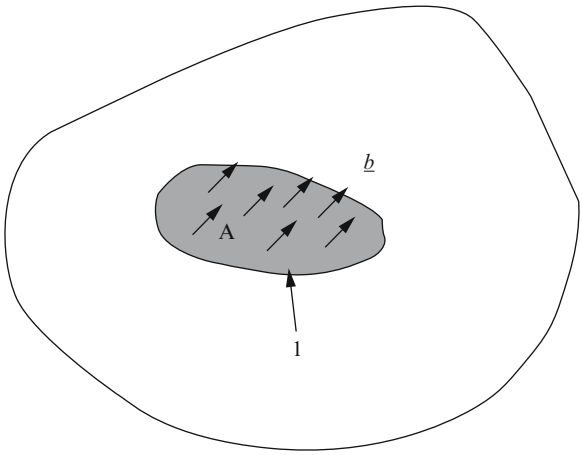


Fig. 3.17 Mathematical definition of a dislocation

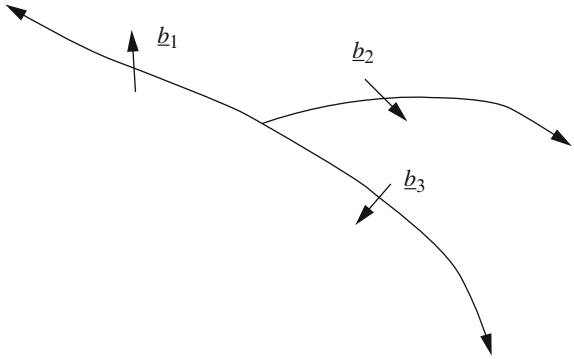


Fig. 3.18 Node of dislocations

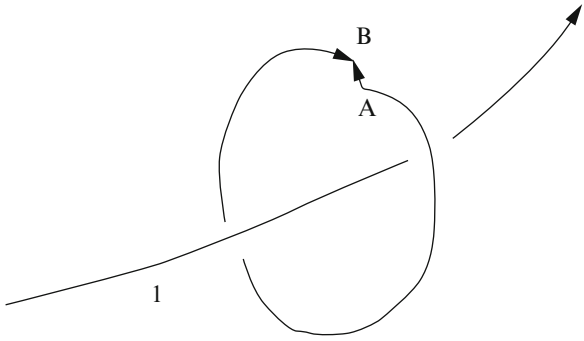


Fig. 3.19 Definition of the Burgers vector $\underline{b} = \underline{\mathrm{AB}}$

solid (Fig. 3.22). But as we indicated above, the same dislocation can be produced, according to the orientation of the cut, by taking out an edge or by glide of the sides. On figures, an edge dislocation is represented by an upside down T, whose stem is in the direction of the inserted edge.

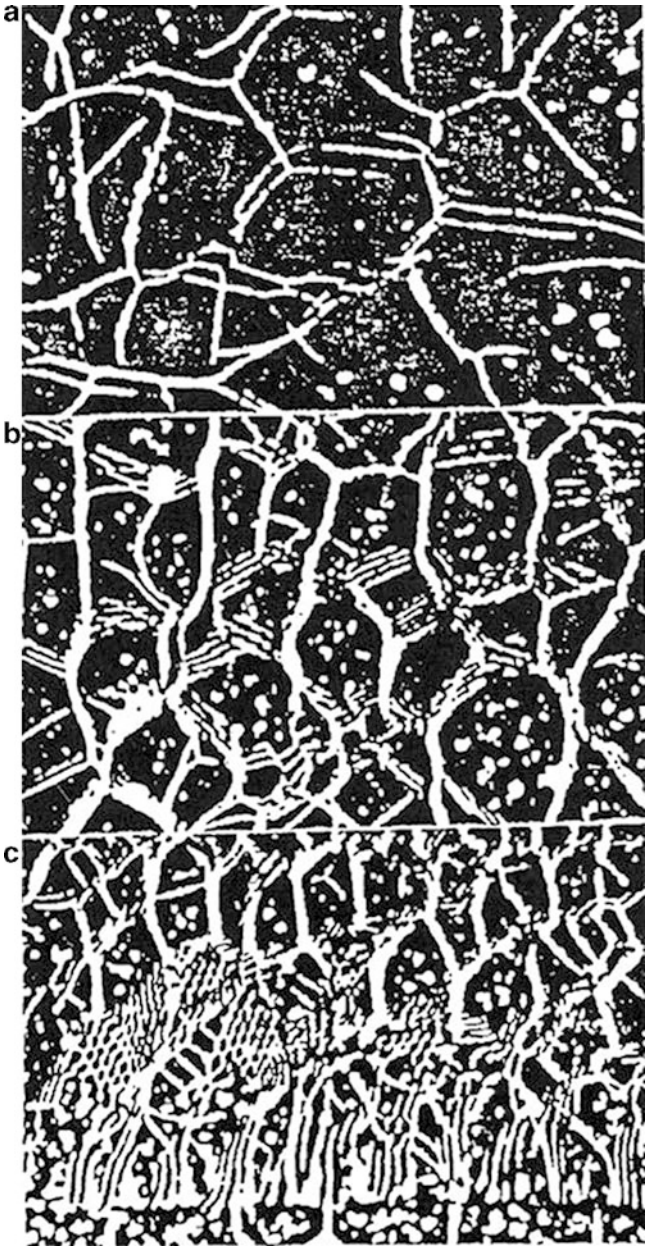


Fig. 3.20 Frank's network of dislocations in NaCl. The three pictures concern three different dislocation densities

Any dislocation can be decomposed into a screw part and an edge part, since the creation processes are purely elastic. In general, a dislocation will consist of purely "edge" parts and purely "screw" parts separated by parts of mixed nature (Fig. 3.23).

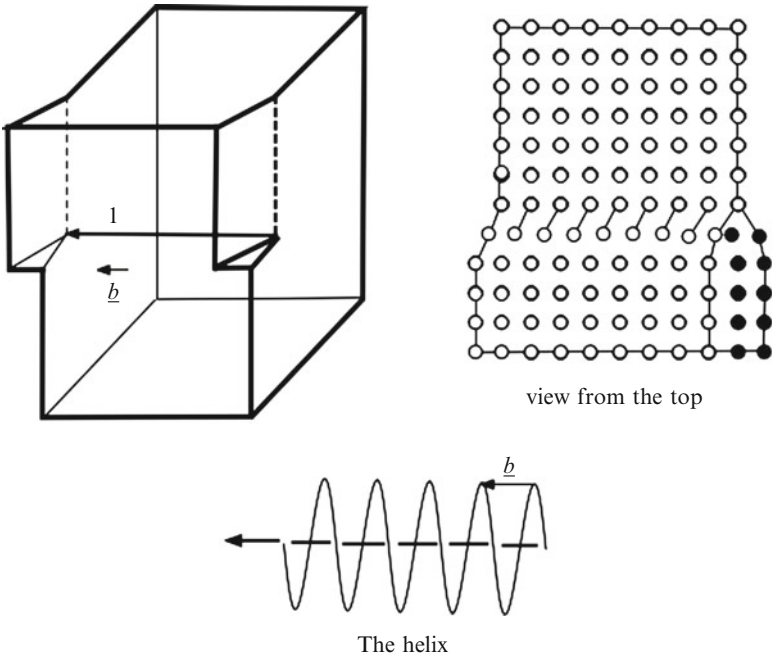


Fig. 3.21 Screw dislocation

3.3.3.3 Displacement of Dislocations

In general, any displacement of the dislocation line l requires addition or removal of material. This movement is called a *climb*, with no implied relation to any reference system, and is non-conservative. But a displacement in the plane defined by part of the line l and the Burgers vector $\underline{b}$ is a *slip* and is conservative (Fig. 3.24). A screw dislocation has an infinite number of slip planes.

A dislocation loop in general has screw and edge parts. The screw portions can slip (or glide) in any case, whereas edge portions can do so only when the Burgers vector is in the plane of the loop. However, a loop whose Burgers vector is perpendicular to its plane can glide along a cylinder, the axis of which is perpendicular to the plane of the loop (Fig. 3.25). Indentation of the surface of a solid creates such loops. They can also result from the condensation of excess vacancies or interstitials (refer to Sect. 3.3.4.2). Such dislocations are called prismatic and the corresponding glide a *prismatic glide*.

3.3.3.4 Dislocation Density

We define the *dislocation density* ρ_D as the total length of the dislocations per unit volume:

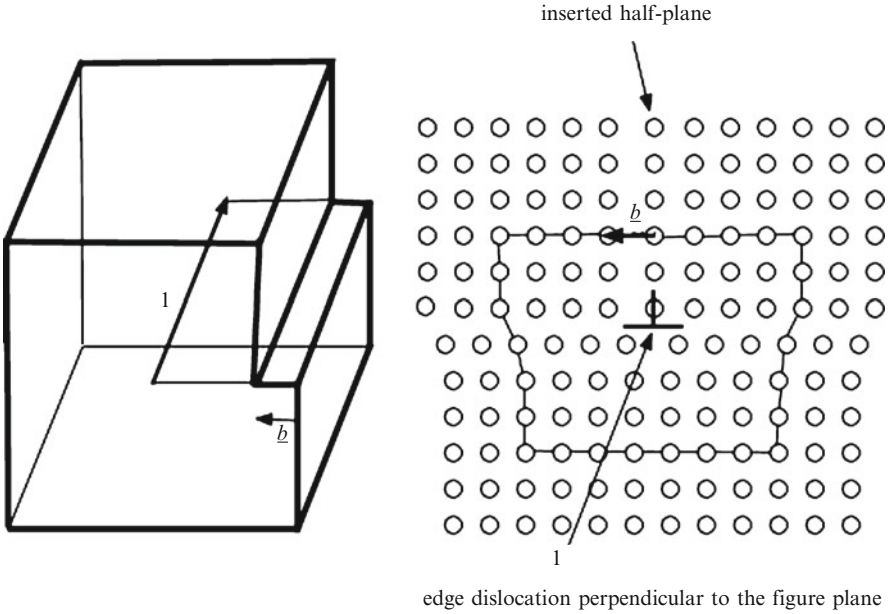


Fig. 3.22 Edge dislocation

$$\rho_D = NL/V \tag{3.11}$$

if there are N dislocations of length L (Table 3.3).

This also equals $1/L_F^2$, the number of dislocations that cross a unit area, L_F being the mean distance between dislocations or the size of the Frank's network.

> The dislocation density is the total length of dislocations per unit volume or the number of dislocations, which cross a unit area. It is expressed in m^{-2}.

When a fraction only of the dislocations can move, we can also define the *mobile dislocation density* ρ_M.

3.3.3.5 Geometrically Necessary Dislocations

Forming a solid into a particular shape by plastic deformation will introduce a certain distribution of dislocations, called *geometrically necessary dislocations.* An example is a single-crystal bar, bent to have a curvature $1/R$ (Fig. 3.26). To find

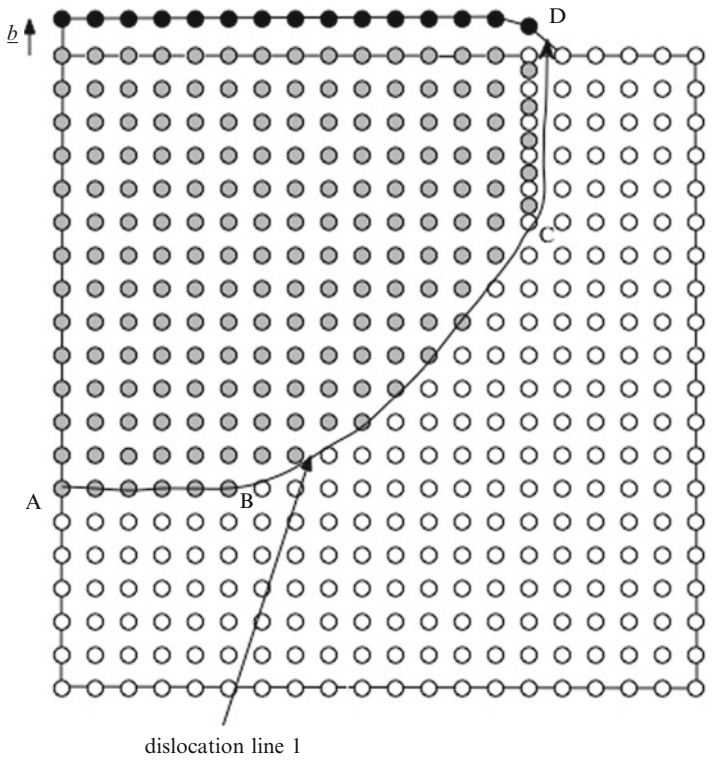


Fig. 3.23 Dislocation with edge part AB, screw part CD and mixed part BC. The grey atoms are those of the lower cut surface, which was displaced a distance $\underline{b}$

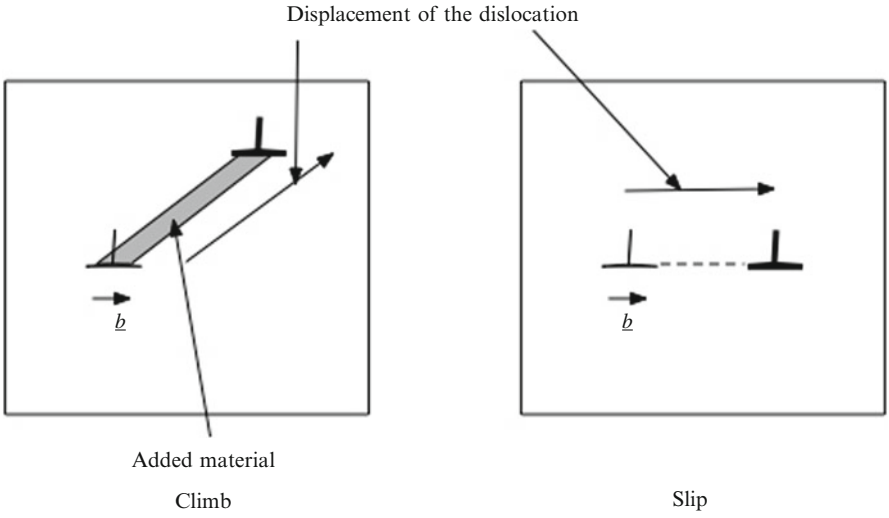


Fig. 3.24 Climb and slip of a dislocation. The dislocation line is perpendicular to the plane of the figure

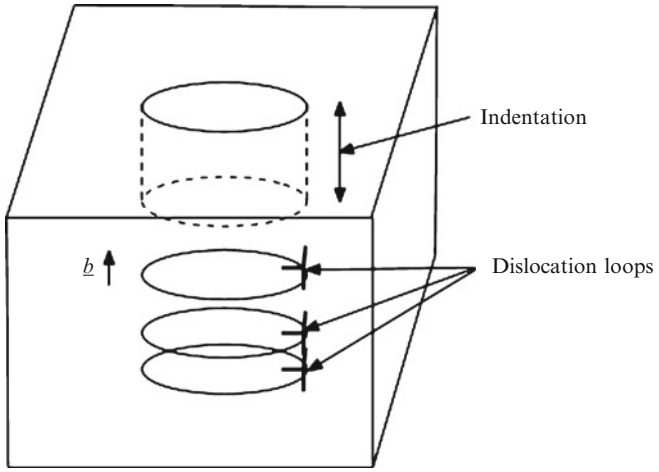


Fig. 3.25 Dislocation loops created by indentation of the surface of the solid

Table 3.3 Approximate dislocation densities

State	Dislocation density ρ_D (m^{-2})
Single crystal solidified with precautions	10^5–10^6
Annealed single crystal	10^9–10^{10}
Annealed polycrystal	10^{10}–10^{11}
Work-hardened polycrystal	10^{13}–10^{16}

the density ρ_G of these dislocations, let ABCD be a Burgers circuit in the crystal before deformation, becoming ABCD' after bending. The closure fault DD' is equal to the total Burgers vector, which is nb if n is the number of geometrically necessary dislocations. The density is this number n divided by the area of the circuit:

$$\rho_G = \frac{n}{AB.BC} = \frac{AD' - AD}{b.AB.BC} = \frac{(R + AB)\,\theta - R\theta}{b.AB.R\theta} = \frac{1}{Rb} \tag{3.12}$$

Here are a few other examples. The first one (Fig. 3.25) is provided by the driving of an indenter of diameter ϕ a distance u into a material. The geometrically necessary dislocations can be grouped into loops of edge dislocations of diameter ϕ; their number n is such that $nb = u$ and the density is

$$\rho_G = \pi\phi u/bV \tag{3.13}$$

where V is the volume of the crystal.

To find the density of such dislocations due to rigid inclusions of diameter D in a plastic matrix we consider a cube of side L, the average distance between the inclusions, subjected to a shear strain $\gamma = u/L$. Figure 3.27 shows that this requires

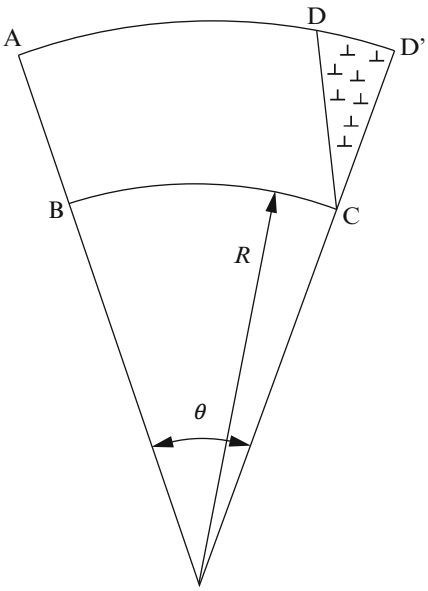


Fig. 3.26 Geometrically necessary dislocations for accommodation of plastic bending deformation

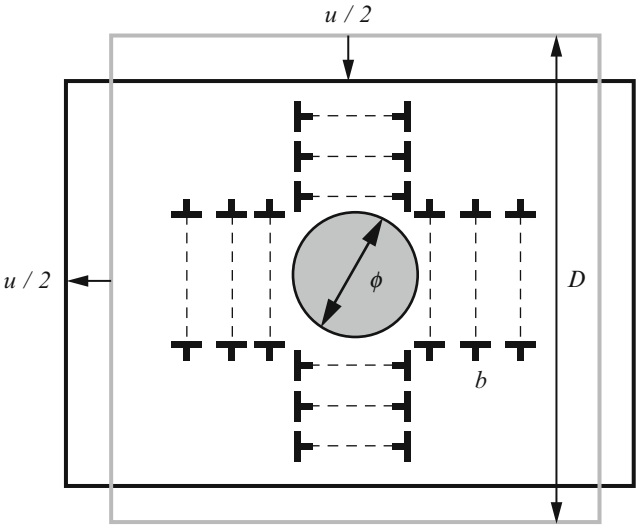


Fig. 3.27 Geometrically necessary dislocations around a rigid inclusion

a number n of loops of dislocations such that $nb = 4(u/2)$, giving a density

$$\rho_G = \frac{2u}{b}\pi D \Big/ L^3 = 2\pi\gamma \frac{D}{bL^2} \tag{3.14}$$

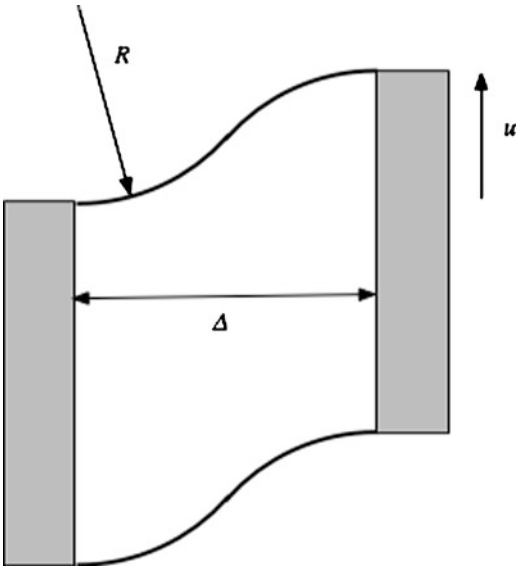


Fig. 3.28 Geometrically necessary dislocations generated by shearing two rigid plates

Since the volume fraction of inclusions is $f_V = \frac{\pi}{6}\frac{D^3}{L^3}$, this can be written:

$$\rho_G = (288\pi)^{1/3}\frac{\gamma}{bD}f_v^{2/3} \simeq 10\frac{\gamma}{bD}f_v^{2/3} \tag{3.15}$$

Finally for this question we consider the deformation resulting from a shear strain γ in a material clamped between two rigid plates distant of Δ displaced by u (Fig. 3.28). It will acquire a curvature $1/R$ such that:

$$R = \frac{\Delta^2 + u^2}{4u} = \frac{1+\gamma^2}{4\gamma}\Delta \simeq \frac{\Delta}{4\gamma} \tag{3.16}$$

(since γ is small) and substituting this in (3.12) we have:

$$\rho_G = \frac{1}{Rb} = \frac{4\gamma}{\Delta b} \tag{3.17}$$

3.3.4 Dislocations in Face Centred Cubic (FCC) Crystals

Dislocations were introduced essentially to explain plastic deformation of crystalline materials. After the preceding description of elementary properties of dislocations in a plain medium, we need to understand their particularities when they are within a crystal. An important one is that the Burgers vectors are as small as possible so as to reduce the elastic strain energy of the dislocations, as we shall see in

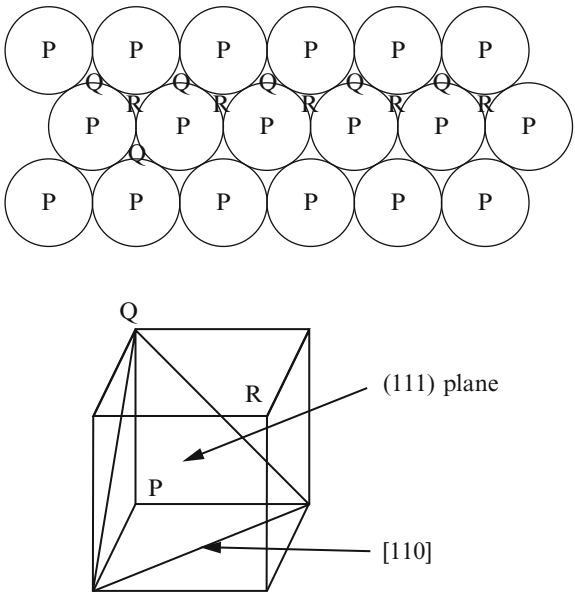


Fig. 3.29 Diagram of a close-packed FCC (111) plane

Sect. 3.3.7. We treat first the case of the FCC structure, which is easier to understand and concerns several alloys of practical importance; aluminium alloys, austenitic stainless steel and superalloys of nickel, copper alloys, lead, precious metals (gold and silver).

3.3.4.1 Perfect and Imperfect Dislocations, Stacking Faults

Because of the periodic structure of a crystal, the Burgers vector cannot be arbitrary: it must be such that the periodicity of the lattice is re-established after the displacement, and therefore must be a sum of multiples of the lattice vectors.

Dislocations having such vectors are called *perfect*. However, there exist *imperfect* dislocations, and these introduce *stacking faults*. The study of these latter is simplest for the FCC case, a compact packing of atoms, which we can represent by spheres; a {111} plane of the structure (Fig. 3.29) consists of a compact packing of these located at the vertices of equilateral triangles.

In the diagram of Fig. 3.29 the points P are the locations of these atoms; Q those of the atoms in the next {111} upper plane and R those in the plane next adjacent; the Q atoms are at the vertical of centroids of the first P triangles; the R atoms at the vertical of centroids of the Q triangles. The FCC structures can be represented

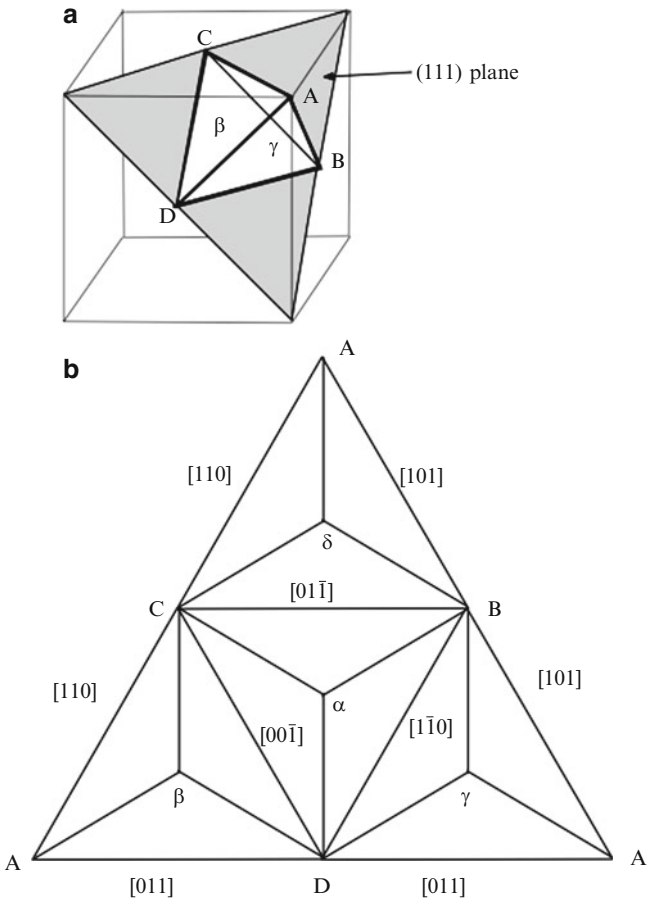


Fig. 3.30 Thompson tetrahedron: (**a**) three-dimensional view; (**b**) plane view after cutting and opening the faces. Faces *a*, *b*, *c*, *d* are the (111), $(\bar{1}11)$, $(11\bar{1})$ and $(1\bar{1}1)$ planes respectively. The segments such as αB, which join the vertices to the centres of the faces are <112> directions

by a PQRPQRPQR stacking, symbolically ΔΔΔΔΔΔ. Dislocations with Burgers vector PP, that is, ½ <110>, are *perfect*. Their vectors form the edges of a regular tetrahedron, called the Thompson tetrahedron (Fig. 3.30) (Thompson 1953). (It can be a good exercise to fabricate a Thompson's tetrahedron with cardboard).

The faces of the Thompson tetrahedron ABCD, whose centres are labelled α, β, γ, δ, are the {111} slip planes a, b, c, d. A dislocation with Burgers vector Aγ, for example, is *imperfect*: it moves the atoms at Q to the positions R and changes the stacking to PQRPRPQR, or ΔΔΔ∇ΔΔΔ, thus introducing a *stacking fault*. A perfect dislocation can be broken into a pair of imperfect dislocations, for example

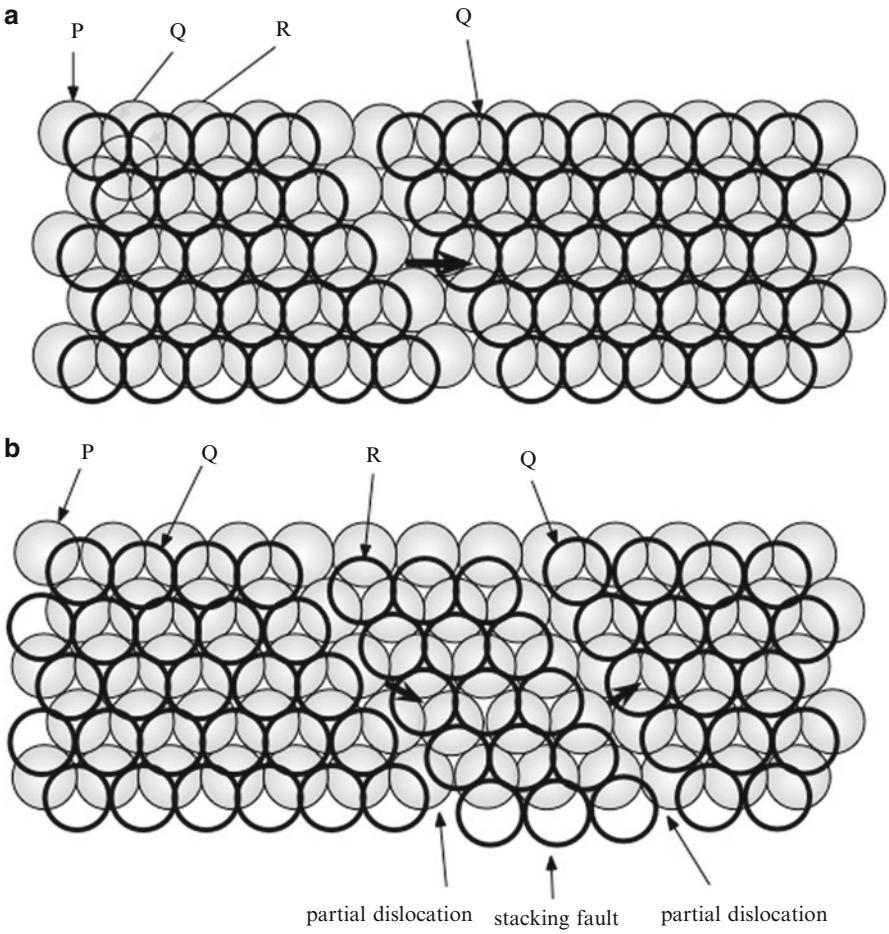


Fig. 3.31 (**a**) Perfect dislocation in a FCC structure. (**b**) Stacking fault bordered by two Shockley imperfect dislocations. (Heavy arrows are the Burgers vectors)

AB into Aδ+δB – equivalently $1/2\,[101] = 1/6\,[211] + 1/6\left[1\bar{1}2\right]$; there will be a stacking fault between the two imperfect dislocations. These are called Shockley dislocations[8] (Fig. 3.31).

3.3.4.2 Cross-Slip, Glissile and Sessile Dislocations

Every perfect dislocation belongs to two of the {111} slip planes: thus BC can slip in either ABC or DBC. A slip that starts in one plane and moves into another is

[8]William B. Shockley (1910–1989) was an American physicist who won the Nobel Prize in 1956 for the invention of the transistor.

called a *cross-slip*. An imperfect dislocation, on the other hand, can slip in only one plane, for example Bα in BCD, and cannot cross-slip.

Two Shockley dislocations can combine to form a dislocation with a vector such as αβ, for example. Such a dislocation cannot slip because its Burgers vector is not contained in any of the slip planes: it is called a *sessile* dislocation, in contrast to the Shockley dislocations, which are *glissile*. Thus an interaction of this kind, called the *Lomer-Cottrell lock*, can block the slip. Being at the intersection of two slip planes, carpet-like, such a dislocation is called a *stair-rod*. Later we shall see the importance of this for work-hardening (Exercise in Volume III).

A crystal will always contain vacancies. If as a result of quenching or irradiation these are supersaturated with respect to the thermodynamic equilibrium, they will tend to become eliminated. An effective method of elimination is for them to combine into pairs, triplets, etc. and finally into cavities appearing as discs over a dense plane. If these exceed a certain critical size the crystal energy will decrease as a result of a breakdown of the structure, eliminating part of a P plane for instance; this will create a stacking fault of the type PQRQRPQR, or ΔΔΔ∇ΔΔΔ, surrounded by a sessile *Frank dislocation* loop (Frank and Read 1950) with Burgers vector of type Aα, or (1/3) <111> (Fig. 3.32). Such a stacking fault is called an *intrinsic stacking fault.*

In the same way, interstitial atoms displaced by irradiation can condense to form disks, which are bordered also by a Frank dislocation loop. Those disks constitute an *extrinsic stacking fault* PQRPQPRPQR or ΔΔΔΔ∇∇ΔΔΔ (Fig. 3.33).

3.3.5 Dislocations in Other Crystalline Structures

3.3.5.1 Body Centred Cubic Structure (BCC)

Dislocations in the BCC structure concern the important case of ferritic steels, and also of transition metals (V, Nb, Ta, Cr, Mo, W) and of the high temperature phase of titanium and zirconium, which are found stabilised in some alloys.

In this structure there has been some uncertainty about the slip systems because the Burgers vector being ½ <111>, many possible slip planes contain this direction, particularly the {110}, {123} and {112} planes (Fig. 3.34). Another dislocation can be a <100> dislocation, the result of the reaction of the type: $1/2[111] + 1/2[1\bar{1}\bar{1}] = [100]$.

Figure 3.35 shows the (112) plane of the BCC structure. Black circles N°1 represent the atoms in this plane. The circles numbered 2, 3, 4, 5 and 6 are the projections of the atoms in the successive (112) planes. The stacking is 123456 and can be represented by ΔΔΔΔΔΔ. A twin corresponds to the stacking in the reverse order: 165432 or ∇∇∇∇∇∇. The sequence ΔΔΔ∇ΔΔΔ represents a stacking fault shifting atoms in position 4 to position 3 by a translation $\frac{1}{6}a[11\bar{1}]$. The stacking fault will be suppressed by a further shift of $\frac{1}{3}a[11\bar{1}]$ corresponding to a stacking fault ΔΔΔ∇∇ΔΔΔ.

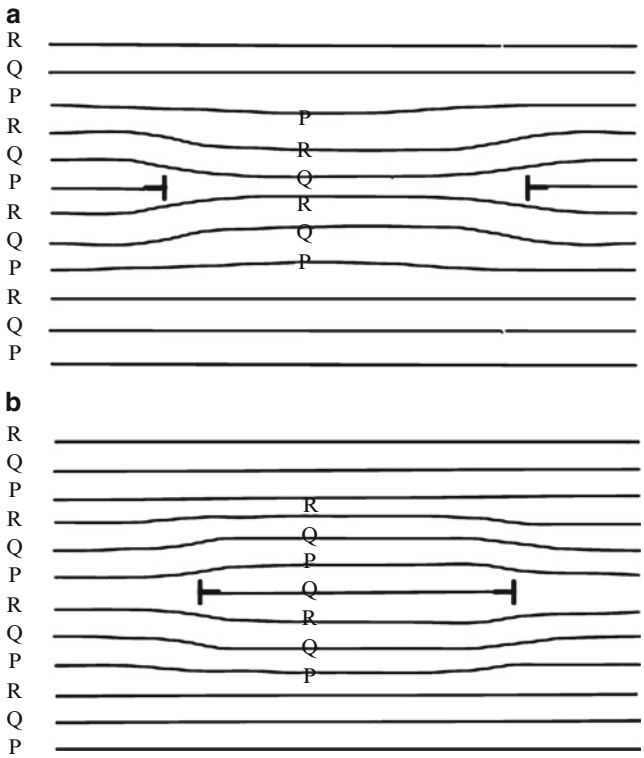


Fig. 3.32 (**a**) Frank dislocation loop created by condensation of vacancies; (**b**) Frank dislocation loop created by condensation of interstitial atoms

The (1/2)a[111] screw dislocation plays an important role in the plastic deformation of BCC metals. This screw dislocation (1/2)a[111] could split into three imperfect dislocations (1/6)a[111] in the planes $(\bar{2}11)$, $(1\bar{2}1)$ and $(11\bar{2})$, bordered by ribbons of stacking faults (Fig. 3.36). These dislocations, having the same Burgers vector, would be pushed by the same force. According to the direction of this force, one of them would move away from the axis [111], while the two others would meet along this axis, recombine and join the first; or the contrary would happen, in which case the dislocations would be locked (grey arrows in Fig. 3.36). However, this dissociation does not take place. But, this mechanism was modified in considering that the core of the dislocation would be slightly extended in the three {112} planes as in Fig. 3.36. Calculations showed that this could be a likely possibility. For some time, it has been the basis of dislocation behaviour in BCC metals. However, more recent calculations have shown that the screw dislocation core is not de-generated by extension in the {112} planes (Fredericksen and Jacobsen 2003; Ventelon and Willaime 2007; Domain and Monnet 2005; Mendelev et al .2003; Li et al .2004). Furthermore, it has been demonstrated, by careful observations in the electron microscope, that *screw dislocations do not glide on* {112} *planes, as the model of*

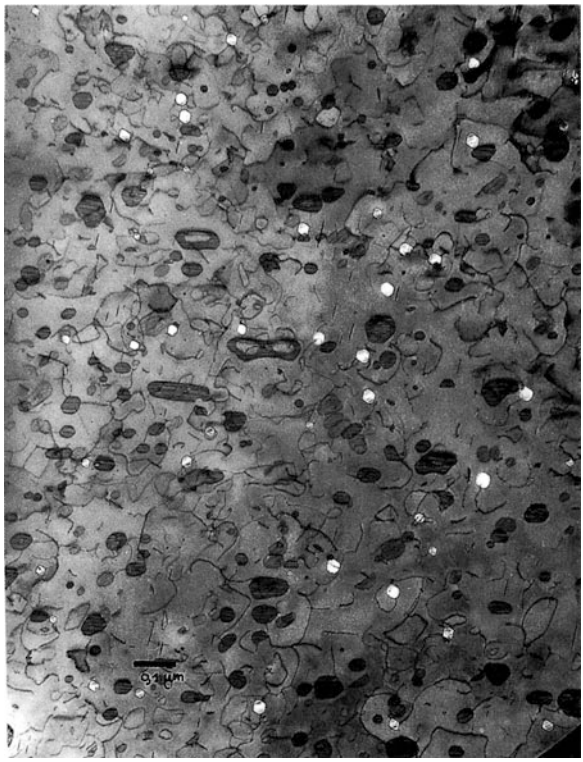


Fig. 3.33 Frank dislocation loops in irradiated austenitic stainless steel. Note that a number of the stacking faults in the centre of the loops were eliminated by introduction of smaller loops (The stacking faults appear darker)

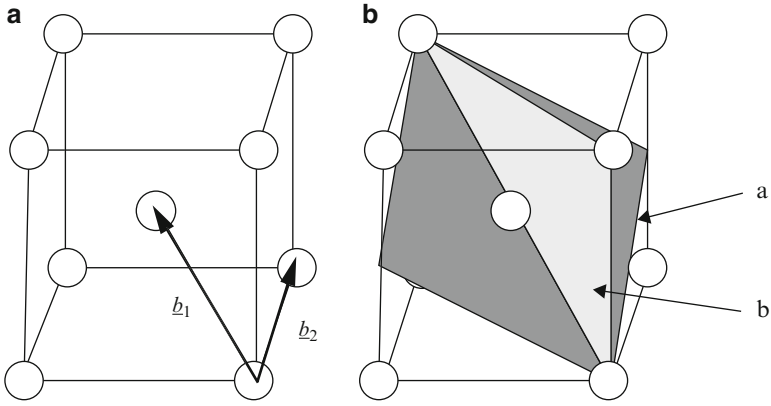


Fig. 3.34 (**a**) Burgers vectors in a BCC structure: $\underline{b}_1 = [111]$; $\underline{b}_2 = [001]$. (**b**) Slip planes: a $= (11\bar{2})$; b $= (110)$

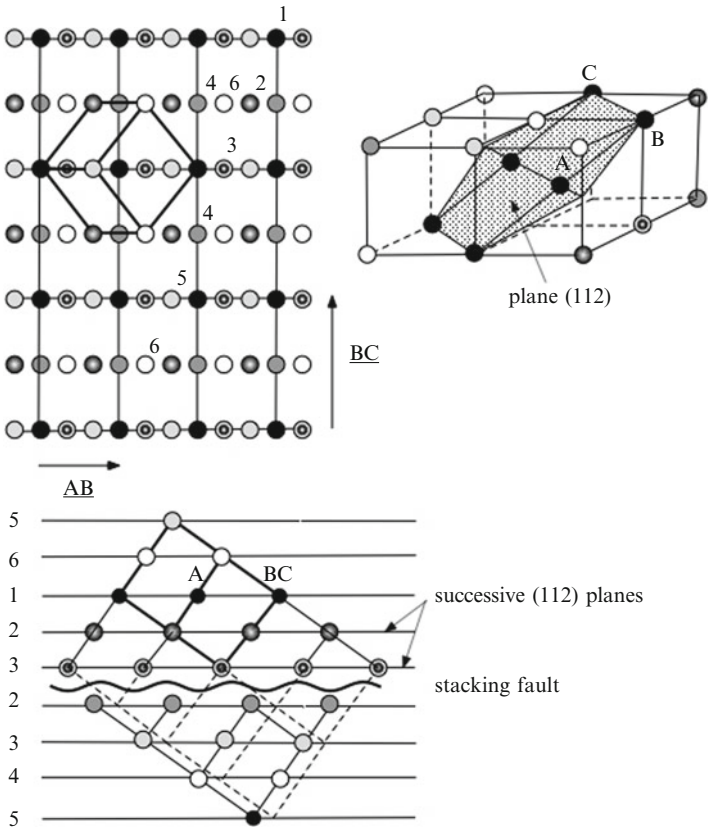


Fig. 3.35 (112) plane of the BCC structure displaying the successive planes (112) numbered 1, 2, 3, 4, 5 and 6. AB is $(1/2)\,a\left[11\bar{1}\right]$ and BC is $a\left[1\bar{1}0\right]$

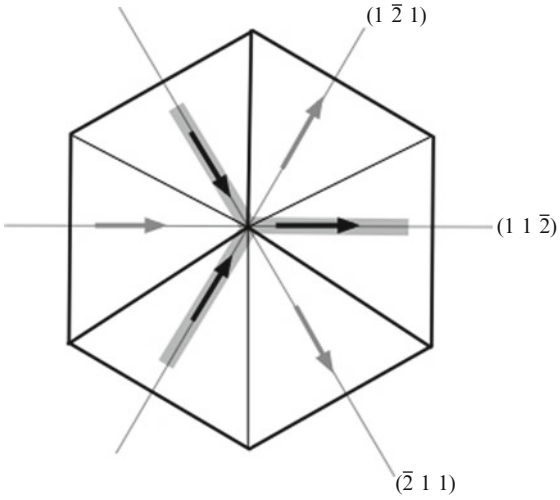


Fig. 3.36 Possible splitting of a (1/2)*a*[111] dislocation, perpendicular to the plane of the figure, into three (1/6)*a*[111] dislocations in three {112} planes, creating ribbons of stacking fault. Pushed by the same force the three dislocations would recombine (*black arrows*) and glide would be possible, while, if they were in the planes corresponding to the *grey arrows* they would be locked

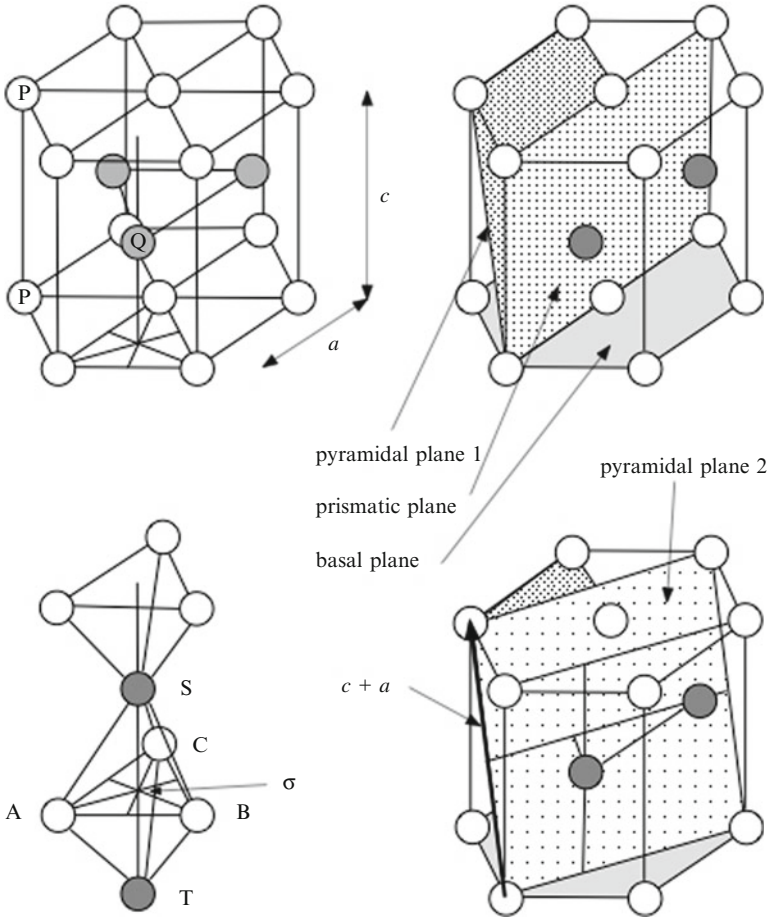


Fig. 3.37 Slip planes in a hexagonal close packed structure (Pyramidal plane $1 = (1\bar{1}01)$; pyramidal plane $2 = (11\bar{2}2)$)

the extended dislocation core would imply, but instead on {110} *planes* (Caillard 2010a, b). This will bear consequences on strain-hardening of BCC metals.

3.3.5.2 Hexagonal Close Packed Structure (HCP)

The alloys of industrial importance concerned are those of zinc, magnesium, cadmium, zirconium and titanium. Beryllium is also of interest. A hexagonal close packed structure differs from the FCC one by the position of the second neighbouring atoms only. It is also a pile of close packed planes, the basal (0001) plane, but the successive layers follow the sequence PQPQPQPQ or $\Delta\nabla\Delta\nabla\Delta\nabla\Delta$. Unlike FCC piling, the distance between successive layers is not constant in HCP (Fig. 3.37): the *c/a* ratio differs according to the various elements (see Table 3.4) (in

Table 3.4 Main slip planes and Burgers vectors in the HCP structure

Basal plane ABC		Prismatic plane AσS		
(0001)		$\{1\bar{1}00\}$		
AB	Aσ	AB	TS	σS
$1/3 < 11\bar{2}0 >$	$1/3 < 1100 >$	$1/3 < 11\bar{2}0 >$	[0001]	1/2 [0001]
a	$a/\sqrt{3}$	a	c	$c/2$
–				
Pyramidal plane ABS		Pyramidal plane		
$\{10\bar{1}1\}$		$\{11\bar{2}2\}$		
AB	AS	$c+a$	AS	
$1/3 < 11\bar{2}0 >$	$1/6 < \bar{2}203 >$	$1/3<11\bar{2}3>$	$1/6 < \bar{2}203 >$	
a	$\sqrt{c^2/4 + a^2/3}$	$\sqrt{c^2 + a^2}$	$\sqrt{c^2/4 + a^2/3}$	

crystallographic nomenclatures a designates the distance between closest atoms in the basal plane, whereas so far we designated by a the distance between slip planes). It follows that dislocations with a Burgers vector joining two atoms in the basal plane are perfect and can be decomposed in two Shockley imperfect dislocations. This is the case in Zn, Be, Mg, Cd for which the basal plane is the preferred slip plane. However, the situation is different for Ti and Zr for which the preferred slip planes are the prismatic ones (Fig. 3.37). Notice that an intrinsic stacking fault in a FCC structure is like a layer of hexagonal material, and that in the same way, a stacking fault along the basal plane in a HCP structure is like a layer of FCC material.

The main slip planes in the HCP structure are represented in Fig. 3.37. Table 3.4 enumerates the slip planes and the Burgers vectors, which they include.

It is important to note that slip in directions out of the basal plane requires the displacement of dislocations with rather long Burgers vectors, c or $c + a$, which have a large self-energy. When we discuss later the Taylor model, we shall see that five independent slip systems are needed to accommodate any prescribed deformation. Now, if no slip is possible out of the basal plane, the number of independent slip system is only three. This accounts for the brittleness of beryllium and zinc. We shall see that twinning compensates this lack of slip possibilities.

3.3.6 Force on Dislocations and Deformation Produced by Their Displacement

3.3.6.1 Glide Force on a Dislocation

We now consider the effect on a dislocation of external forces, starting with a very simple example. Suppose a cube of side L is acted on by a shear force τ per unit area, applied over a face parallel to the slip plane of a dislocation whose Burgers vector is $\underline{b}$ (Fig. 3.14). If the slip traverses the whole of the cube the upper part is

displaced a distance b with respect to the lower part and the work done W is:

$$W = \tau L^2 b \tag{3.18}$$

If instead the dislocation is shifted only a distance δx we can assume that the corresponding displacement of the cube is $\delta x/L$ and the work done is

$$\delta W = \tau L^2 b(\delta x/L) \tag{3.19}$$

This is as if a configuration force f per unit length was applied on the dislocation line. Its work would be:

$$\delta W = f L \delta x \tag{3.20}$$

So, equating these two expressions, we have:

$$f = \tau b \tag{3.21}$$

The glide force on a unit length of dislocation is equal to the product of the shear stress on the glide plane by the Burgers vector.

3.3.6.2 Deformation Produced by Dislocation Glide

The shear strain $\delta\gamma$ is such that:

$$\delta\gamma = (\delta x/L)\,(b/L) = L\,b\,\delta x/L^3 = b\,\delta A/V \tag{3.22}$$

where δA is the area swept by the dislocation and V the volume of the cube.

The shear strain produced by a dislocation glide is equal to the product per unit volume of the Burgers vector by the area swept by the dislocation.

If the density ρ_{D} of dislocations of Burgers vector of norm b is moved an average distance l_{D} we have:

$$\gamma = N\,b\,A/V = \rho_{\mathrm{D}}\,b\,l_{\mathrm{D}} \tag{3.23}$$

In reality only some fraction will be mobile; if the density of these is ρ_{M} and their average velocity is v, the slip rate is:

$$\dot{\gamma} = \rho_{\mathrm{M}} b v \tag{3.24}$$

The shear strain rate is equal to the product of the mobile dislocations density by their average velocity and by the norm of the Burgers vector. This law is sometimes called the Orowan law.

For geometrically necessary dislocations, by analogy with (3.23), we define a characteristic length l_G, the *mean free path*, by $l_G = \gamma / \rho_G b$. Thus, this mean free path is equal to $\phi \Big/ 9.67 f_V^{2/3}$ for rigid inclusions and to $D/4$ for rigid plates, as can be calculated from Eqs. 3.15 and 3.17. These quantities are presently largely discussed in the theories of plasticity of polycrystalline materials *with internal lengths.*

3.3.6.3 Force on a Dislocation; General Case

We now treat these questions in a more general manner.

The internal stress field of the dislocation introduces an elastic deformation energy E_D into the solid; the externally-applied forces $\underline{T}$ generate their own deformation energy E_T independently of E_D, so the total stored elastic energy is $E = E_D + E_T$.

If the forces $\underline{T}$ vary then the internal dislocation stresses, zero on the surface $\partial\Omega$, do no work, so any change in the elastic energy is equal to the work done by the external forces and does not depend on the presence of the dislocation. This, therefore, cannot be detected by measuring any elastic constants: the dislocation must be made to move for that to be possible.

Consider now the work done in creating a dislocation. Before the cut of area A was made (Fig. 3.17) the deformation energy was E_T. Forces $\underline{T}_D$ are applied to move this area the vector distance $\underline{b}$. The work done is that done by these forces, equal to the dislocation energy E_D, plus the work done by the external forces in being displaced by $\underline{u}_D$, plus that done by the stresses $\underline{\underline{\sigma}}$ acting over the plane of A (deriving from the externally-applied actions and possibly from other sources of internal stresses.)

To calculate this energy, let $\underline{n}$ be the outward (*i.e.*, towards the exterior of the solid) normal at the boundary of the cut where the Burgers circuit originates (Fig. 3.38). With the convention concerning the direction of this vector (Sect. 3.3.3.1), the work done by the stresses is $-\int_A b_i \sigma_{ij} n_j \, \mathrm{d}A$.

The stored energy is:

$$E = E_T + E_D + \int_{\partial\Omega} \underline{T}.\underline{u}_D \mathrm{d}S - \int_A b_i \sigma_{ij} n_j \, \mathrm{d}A \tag{3.25}$$

and since $E = E_D + E_T$

$$\int_{\partial\Omega} \underline{T}.\underline{u}_D \mathrm{d}S = \int_A b_i \sigma_{ij} n_j \, \mathrm{d}A \tag{3.26}$$

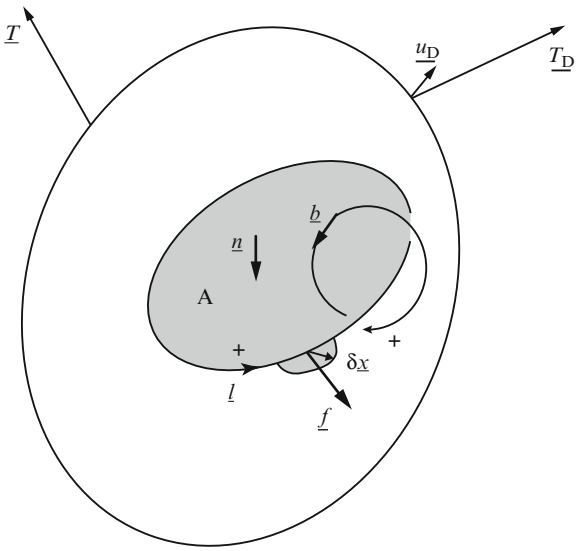


Fig. 3.38 Force exerted on a dislocation

If now the dislocation is moved a distance $\delta x(s)$, where s is the curvilinear coordinate along the length of the dislocation line l, the work done by the external actions is:

$$\delta W = \int_{\partial\Omega} \underline{T}.\delta\underline{u}_{\mathrm{D}}\mathrm{d}S = \delta \int_{A} \underline{b}.\underline{\underline{\sigma}}.\underline{n}\mathrm{d}A = \int_{l} \underline{b}.\underline{\underline{\sigma}}.\mathrm{d}\underline{l} \wedge \delta\underline{x} \tag{3.27}$$

with $\underline{l}$ the unit vector along the dislocation line.

We can consider this as the work done by a configuration force $\underline{f}$ per unit length along the dislocation line; then:

$$\delta W = \int_{l} \underline{f}.\delta\underline{x}\mathrm{d}l \tag{3.28}$$

and equating the two expressions we have:

$$\underline{f} = \left(\underline{b}.\underline{\underline{\sigma}}\right) \wedge \underline{l} \tag{3.29}$$

This is called the *Peach – Koehler formula* (Koehler 1941; Peach and Koehler 1950) .[9]

The scalar product $\underline{f}.\underline{l}$ vanishes identically, showing that the force f is normal to the dislocation line. In the case of slip its component in the slip plane is $\tau\, b$, where, as we showed above, τ is the resolved part of the shear in the slip plane parallel to the Burgers vector.

[9] J.S. Koehler was professor at the University of Illinois and M.Peach was his PhD student.

3.3.6.4 Deformation Produced by the Displacement of a Dislocation; General Case

We now find the deformation produced in the solid by the movement of a dislocation whose Burgers vector is $\underline{b}$ in sweeping an area δA in a plane whose normal is $\underline{n}$ (Fig. 3.38) (Exercise in Volume III and Sect. 3.5.1.2).

Suppose the solid subjected to a homogeneous stress such that the σ_{IJ} component, with I and J fixed, is different from zero. The above expression (3.27) for the work done can then be written:

$$\int_{\partial\Omega} (n_I \sigma_{IJ} \delta u_J + n_J \sigma_{IJ} \delta u_I) \mathrm{d}S = (b_I \sigma_{IJ} n_J + b_J \sigma_{IJ} n_I)\, \delta A \tag{3.30}$$

Simplifying by σ_{IJ}, using the result for any value *i,j* of I,J and applying Green's theorem, we obtain:

$$\int_{\partial\Omega} \left(n_i \delta u_i + n_j \delta u_j\right) \mathrm{d}S = \int_{\Omega} \left(\delta u_{i,j} + \delta u_{j,i}\right) \mathrm{d}V \tag{3.31}$$

This provides an expression for the average strain

$$\delta\overline{\varepsilon_{ij}} = \frac{1}{2V} \int_{\Omega} \left(\delta u_{i,j} + \delta u_{j,i}\right) \mathrm{d}V = \left(b_i n_j + b_j n_i\right) \frac{\delta A}{2V} \tag{3.32a}$$

which can also be written:

$$\delta\underline{\underline{\bar{\varepsilon}}} = (\underline{b} \otimes \underline{n})^S \frac{\delta A}{V} \tag{3.32b}$$

where $\otimes$ denotes a tensorial product and S the symmetrical part of the expression between braces. In particular, this formula gives the expression (3.22) for the shear strain produced by dislocation glide. Dislocation climb results in extension.

3.3.7 Stress, Strain Fields and Deformation Energy Associated with a Dislocation

The Volterra process (Sect. 3.3.3.1) shows that creating a dislocation in a solid needs energy and introduces residual stresses; we now show how these stresses and also the stored elastic energy can be calculated. We shall find that in order to minimise this energy the dislocations tend to keep to straight lines and behave like stretched elastic bands. The calculation will show also that a dislocation does not change the volume of the crystal, and thus supports the assumption usually made that volume remains constant in plastic deformation. We first discuss screw dislocations, which are easier to treat than edge dislocations.

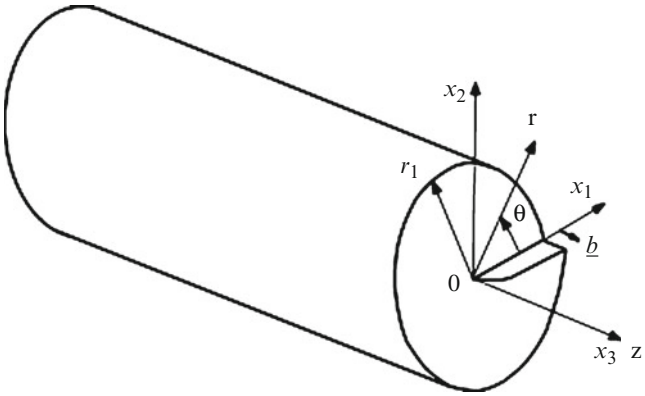


Fig. 3.39 Displacement field for a screw dislocation

3.3.7.1 Strains and Stresses Around a Screw Dislocation

Suppose the dislocation lies in the direction of Ox_3 (Oz in cylindrical polars); the resulting displacement will be in the same direction (Fig. 3.39):

$$u_z = b\theta/2\pi \tag{3.33}$$

The strain field is easily found by differentiation:

$$\varepsilon_{\theta z} = \frac{b}{4\pi r} \tag{3.34}$$

all other components being equal to zero.

The corresponding stresses, obtained from the Hooke's relations, are:

$$\sigma_{\theta z} = \frac{\mu b}{2\pi r} \tag{3.35a}$$

or

$$\sigma_{13} = -\frac{\mu b}{2\pi}\frac{x_2}{x_1^2 + x_2^2}$$
$$\sigma_{23} = \frac{\mu b}{2\pi}\frac{x_1}{x_1^2 + x_2^2} \tag{3.35b}$$

and all other components are zero.

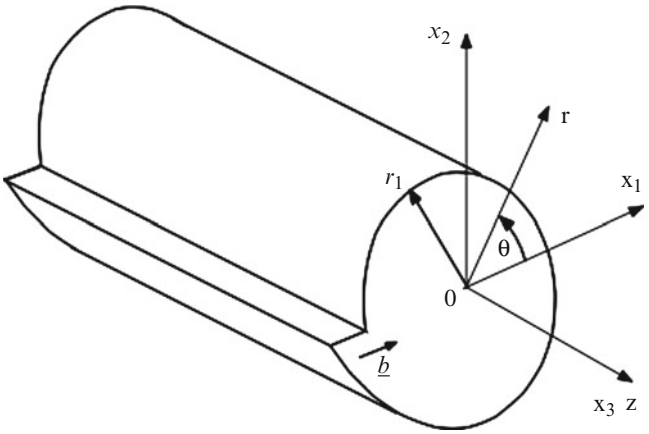


Fig. 3.40 Displacement field for an edge dislocation

The stress system is thus pure shear, with no dilatation. The stress decreases with distance from the dislocation like $1/r$; it cannot of course go to infinity when r goes to zero, and there is what is called a core zone, of radius $r_0 = \alpha b$, in which the laws of linear elasticity cease to hold because the strains are too great. Calculations have shown the value of α to be about 3 for ionic crystals and between 0.5 and 2 for metals.

The stresses $\sigma_{\theta z}$ act on the two ends of the cylinder and form a torque C:

$$C = \int_{r_0}^{r_1} 2\pi r \left(r\sigma_{\theta z}\right) \mathrm{d}r = \frac{\mu b}{2}\left(r_1^2 - r_0^2\right) \tag{3.36}$$

This can be balanced by superposing a stress field corresponding to a displacement of torsion $u_\theta = \frac{rz}{R}$, R being a constant found by equating the torque exerted by this stress field to the one given by the preceding Eq. 3.36. Hence:

$$\sigma_{\theta z} = -\frac{\mu b}{\pi\left(r_1^2 + r_0^2\right)} \tag{3.37}$$

This stress field is negligible in relation to the main field given by Eq. 3.35 above.

3.3.7.2 Strains and Stresses Around an Edge Dislocation

Calculation of the strain and stress field around an edge dislocation is not so straightforward (Nabarro 1967) (Fig. 3.40).

The situation is that of plane strain. A usual method to solve the problem is then to introduce an Airy stress function χ. The stresses are found by differentiation:

$$\sigma_{11} = \frac{\partial^2 \chi}{\partial x_2^2}$$
$$\sigma_{22} = \frac{\partial^2 \chi}{\partial x_1^2}$$
$$\sigma_{12} = -\frac{\partial^2 \chi}{\partial x_1 \partial x_2} \tag{3.38a}$$

or in polars

$$\sigma_{rr} = \frac{1}{r}\frac{\partial \chi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \chi}{\partial \theta^2}$$
$$\sigma_{\theta\theta} = \frac{\partial^2 \chi}{\partial r^2}$$
$$\sigma_{r\theta} = -\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \chi}{\partial \theta}\right) \tag{3.38b}$$

For choosing the right function χ, equations of compatibility require that $\nabla^4 \chi = 0$ and it can be written:

$$\chi = 2x_2 \mathrm{P} + \mathrm{R}'' \tag{3.39}$$

P and R″ being harmonic functions. It is found that

$$\chi = 2Ax_2 \log r \tag{3.40}$$

provides the solution for an edge dislocation with $\underline{b} = (b, 0, 0)$ and $A = \frac{-\mu b}{4\pi(1-\nu)}$, ν being the Poisson ratio. It follows that:

$$u_1 = \frac{b}{4\pi(1-\nu)}\frac{x_1 x_2}{x_1^2 + x_2^2} - \frac{b}{2\pi}\arctan\frac{x_1}{x_2}$$
$$u_2 = \frac{b}{4\pi(1-\nu)}\frac{x_2^2}{x_1^2 + x_2^2} - \frac{(1-2\nu)\,b}{8\pi(1-\nu)}\log\frac{x_1^2 + x_2^2}{b^2} \tag{3.41}$$

$$\varepsilon_{11} = -\frac{b}{4\pi(1-\nu)}\frac{x_2\left(x_1^2 - x_2^2\right)}{\left(x_1^2 + x_2^2\right)^2} - \frac{b}{2\pi}\frac{x_2}{x_1^2 + x_2^2}$$
$$\varepsilon_{22} = \frac{b}{4\pi(1-\nu)}\frac{x_2\left(3x_1^2 + x_2^2\right)}{\left(x_1^2 + x_2^2\right)^2} - \frac{b}{2\pi}\frac{x_2}{x_1^2 + x_2^2}$$
$$\varepsilon_{12} = \frac{b}{4\pi(1-\nu)}\frac{x_1\left(x_1^2 - x_2^2\right)}{\left(x_1^2 + x_2^2\right)^2} \tag{3.42a}$$

$$\varepsilon_{rr} = \varepsilon_{\theta\theta} = -\frac{b\,(1-2\nu)}{4\pi\,(1-\nu)}\frac{\sin\theta}{r}$$

$$\varepsilon_{r\theta} = \frac{b}{4\pi\,(1-\nu)}\frac{\cos\theta}{r} \tag{3.42b}$$

$$\sigma_{11} = -\frac{\mu b}{2\pi\,(1-\nu)}\frac{x_2\left(3x_1^2+x_2^2\right)}{\left(x_1^2+x_2^2\right)}$$

$$\sigma_{22} = \frac{\mu b}{2\pi\,(1-\nu)}\frac{x_2\left(x_1^2-x_2^2\right)}{\left(x_1^2+x_2^2\right)}$$

$$\sigma_{12} = \frac{\mu b}{2\pi\,(1-\nu)}\frac{x_1\left(x_1^2-x_2^2\right)}{\left(x_1^2+x_2^2\right)} \tag{3.43a}$$

$$\sigma_{rr} = \sigma_{\theta\theta} = -\frac{\mu b}{2\pi\,(1-\nu)}\frac{\sin\theta}{r}$$

$$\sigma_{r\theta} = \frac{\mu b}{2\pi\,(1-\nu)}\frac{\cos\theta}{r} \tag{3.43b}$$

These solutions inspire several remarks. The displacement u_2 diverges for large distances from the dislocation. This is an unresolved difficulty.

As with the screw dislocation, the stresses fall off like $1/r$ and this solution is no longer valid for distances less than a value r_0. The material is in compression for $0 < \theta < \pi$ and in tension for $\pi < \theta < 2\pi$; and the shear stress is maximum on the plane $\theta = k\pi$, k integer.

The stress σ_{rr} is exerted over the external surface of the cylinder, of radius r_1; this can be balanced by applying stresses which vary like $r/{r_1}^2$ and which, as before, are negligible compared with those just calculated.

As an illustration, the displacement field around an edge dislocation, visualised by geometric phase method in high resolution transmission electron microscopy in diamond and in boron nitride, is shown in Fig. 3.41.

3.3.7.3 Change of Density of a Solid Due to the Presence of Dislocations

We have seen that a screw dislocation generates a system of deformations in which there is no dilatation. It is not the same for an edge dislocation, but it is easily shown that the reduction in volume over $0 < \theta < \pi$ is exactly balanced by the increase over $\pi < \theta < 2\pi$. That the overall change must be zero is obvious from the fact that the mean value of an internal stress field is zero (Exercise in Volume III)

However, in the dislocation core the deformation is too great for the laws of linear elasticity to hold there, and it has been shown that this produces a local dilatation

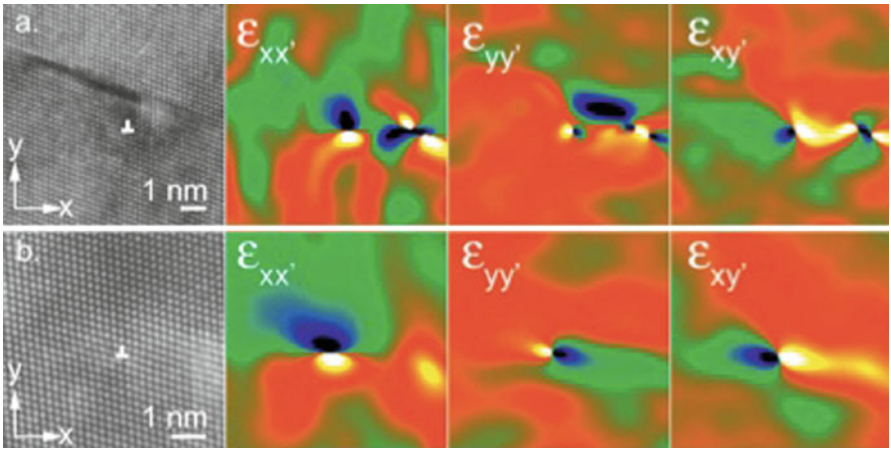


Fig. 3.41 Displacement fields around an edge dislocation in diamond and in cBN obtained by the "geometric phase method" in HREM (Willems et al. 2005)

approximately equivalent to a tube of vacancies. There are consequential variations in density, but these are very small, even for high concentrations of dislocations: typical values of $\Delta V/V = \rho_D b^2$ are 6.25×10^{-8} for $\rho_D = 10^{12}$ m^{-2}, 6.25×10^{-3} for $\rho_D = 10^{17}$ m^{-2}.

The presence of dislocations hardly modify the density of a solid.

An important consequence is that plastic deformation, which derives from the movement and, as we shall see, the multiplication of dislocations, is accompanied by only a negligible change of volume. There is therefore a sound foundation for the assumption usually made in the mechanics of coherent solids, that *plastic deformation takes place at constant volume*.

Nevertheless, the core dilatation can interact with the hydrostatic stress, in such a way that the force on the dislocation is not completely independent of it, especially for screw dislocations in BCC materials.

3.3.7.4 Contrast and Visibility of Dislocations in Electron Microscopy

A dislocation alters both the distances between the crystallographic planes and their orientation. If a crystal is mounted in such a way that the Bragg equation $2d \sin\theta = n\lambda$ holds for a particular plane, a dislocation will destroy the reflection conditions locally and the effect will be seen as a dark line in the field of view (see for instance Fig. 3.12). But in certain orientations this will not occur, in particular

for a screw dislocation in the reflecting plane and for an edge dislocation normal to this, since in these cases the displacements produced by the dislocation are in the reflecting plane and affect neither d nor θ.

The situation can be summarised by saying that the condition for visibility is

$$\underline{g} \cdot \underline{b} \neq 0 \tag{3.44}$$

where $\underline{g}$ is the vector of the diffracting plane in the reciprocal lattice. This has been obtained by a very approximate argument, but is confirmed by a deeper theoretical treatment (see *e.g.* Hirsch et al. 1965).

3.3.7.5 Energy of a Dislocation; Line Tension

The elastic energy E_D of a dislocation is equal to the work needed to separate the boundaries of the area A (Fig. 3.17) by the vector $\underline{b}$; that is:

$$E_D = \frac{1}{2} \int_A b_i \sigma_{ij} n_j \, dA \tag{3.45}$$

$\underline{n}$ being the normal to A.

Replacing σ_{ij} by $\sigma_{\theta z}(\theta = 0)$ we have, per unit of length,

$$E_D = \frac{1}{2} \int_{r_0}^{r_1} M \frac{b^2}{r} dr = \frac{M}{2} b^2 \log \frac{r_1}{r_0} \tag{3.46}$$

where $M = \mu/2\pi$ for a screw dislocation and $\mu/2\pi(1-\nu)$ for an edge dislocation (refer also to Sect. 3.3.7.7).

Thus the dislocation energy is considerable for a large volume of solid. However, it is very difficult to produce crystals with a low density of dislocations: typical values of ρ_D for annealed metals are of the order of 10^{12} m^{-2}, with a corresponding mean distance between dislocations $1/\sqrt{\rho_D}$ of the order of 10^{-6} m. Their stresses will cancel each other at distances of this order, since on average there will be as many dislocations of one sign as of the other, so we can take this as the value for r_1, and with r_0 of the order of 10^{-9} m the values of the logarithm (base e) in (3.46) is close to 7.

It is generally agreed to take the dislocation energy as $0.5\mu b^2$ per unit length.

For copper, for example, with approximate values $\mu = 4{\cdot}10^{10}$ Nm^{-2}, $b = 2.5{\cdot}10^{-10}$ m this is $1.25{\cdot}10^{9}$ Jm^{-1}, or about 2 eV per atom along the dislocation.

It follows that increasing the length of a dislocation – by bending it, say – will increase the total energy; there is therefore a configuring force along the line, corresponding to dE_D/dl, that is, $t = 0.5\,\mu b^2$: this is the *line tension* of the dislocation. Dislocations tend to be as straight as possible. However, in Fig. 3.12,

most dislocations appear more or less curved. This is because they meet the inhomogeneous stress field of neighbouring dislocations.

The *energy of dislocation loops* was calculated. For a circular loop of radius R_c with its Burgers vector in the plane of the loop the energy is (Nabarro 1967):

$$E_{\text{loop}} = \frac{\mu b^2 R_c}{2}\left[\frac{1}{2}\left(1+\frac{1}{1-\nu}\right)\right]\log\frac{8R_c}{e^2 r_0} \tag{3.47}$$

For a prismatic loop with Burgers vector normal to its plane the mean energy is:

$$E_{\text{loop}} = \frac{\mu b^2 R_c}{2(1-\nu)}\log\frac{8R_c}{e^2 r_0} \tag{3.48}$$

If $\nu = 1/3$ the line tension of an edge dislocation is 3/2 times that of a screw dislocation. Along a dislocation loop the edge components will tend to be shorter than the screw components with the result that the shape becomes roughly elliptical.

The elastic energy of dislocations is greater than their entropy of configuration. If ρ_D dislocations cross the unit area, considering that there are $1/b^2$ possible sites, the entropy per unit volume is $kT\rho_D\left[\left(1-\rho_D b^2\right)\log\left(1-\rho_D b^2\right)-\log\left(\rho_D b^2\right)\right]$. It can be checked that this is smaller than $(1/2)\rho_D\mu b^2$ even for very large dislocation densities. It follows that dislocations increase the free energy and are not stable defects in a crystal. However, it is very difficult to eliminate dislocations, since they form self-stabilising networks.

3.3.7.6 Multiplication of Dislocations

The density of dislocations initially present in a crystal is too small to allow a plastic deformation of any great magnitude. Consider for example a cubical grain of side 100 μm containing 10^{12} dislocations/m^2 with Burgers vectors $2.5{\cdot}10^{-12}$ m; on the assumption that they sweep the entire cross-section of the grain, the maximum plastic deformation will be, by (3.21), $\gamma = \rho_D b l_D = 10^8 \times 2.5 \times 10^{-8} \times 10^{-2} = 2.5\%$.

Suppose that, for one reason or another, a dislocation is anchored at points A, B at distance l apart. The effect of the shear stress acting over the slip plane will be to cause this segment to curve, and it will take up a curvature such that an element ds will be in equilibrium under the shear force $f = \tau b$ and the line tension t, assumed constant independent of the type of the dislocation (Fig. 3.42). The radius of curvature R is such that

$$2t\frac{\mathrm{d}s}{2R} = f\,\mathrm{d}s \tag{3.49}$$

i.e.,

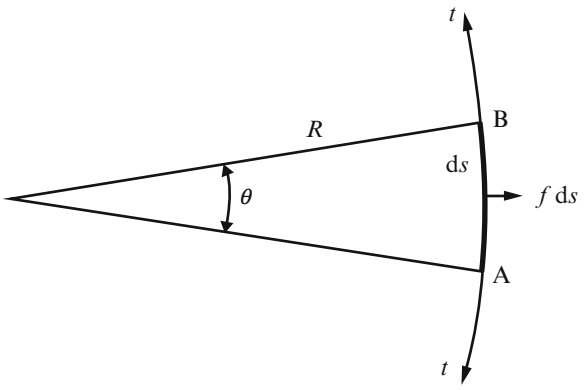


Fig. 3.42 Equilibrium of an element ds of dislocation under the action of a force per unit length $f = \tau b$

$$R = t/f = 0.5\,\mu b^2/\tau b = 0.5\,\mu b/\tau = t/\tau b \tag{3.50}$$

If τ increases R will decrease until it has fallen to $l/2$ when the arc becomes unstable and wraps around the points A and B. The parts of opposite sign then recombine (as shown in Fig. 3.43) forming a loop, which spreads into the slip plane, leaving behind a new segment AB to repeat the process. Putting $R = l/2$ in (3.50) we see that the activation stress for this process is (Exercise in Volume III)

$$\tau_{\mathrm{FR}} = \frac{\mu b}{l} \tag{3.51}$$

This mechanism of dislocation multiplication is called a *Frank-Read source*. It is considered as the main one and it plays an important role in plastic deformation.

Figure 3.44 shows a process similar to the Frank-Read source and illustrates the way it works. It is the photography of steps produced on the surface of aluminium by the uniform growth sideways of the initial step, which existed at the emergence of two screw dislocations.

Another source of dislocations could be grain boundaries.

3.3.7.7 Peierls's Model of the Core of a Dislocation

The core energy along a dislocation is minimum when the dislocation follows a dense row of atoms. To move from one such row to the next the energy has to be increased sufficiently to break some inter-atomic bonds. We say that the dislocation is in a *Peierls trough*[10] (Fig. 3.45).

The depth of these troughs is a function of the bond energy, and is especially great for covalent materials. In these the dislocations tend to run in very straight

[10] Sir Rudolf Ernst Peierls (1907–1995) was a German born British scientist.

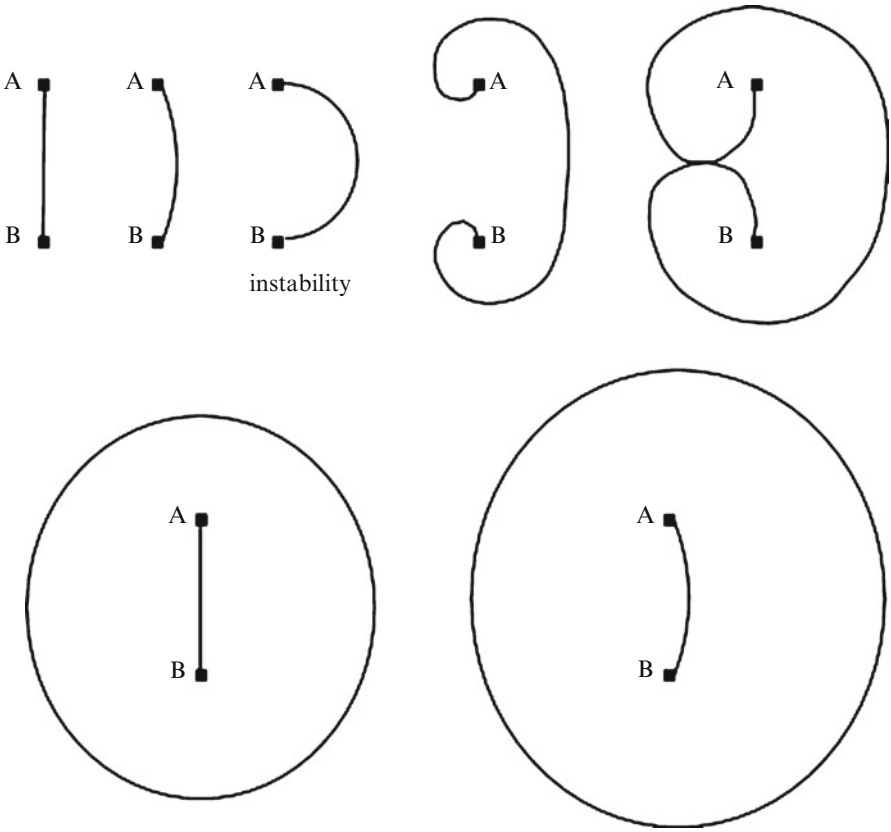


Fig. 3.43 Frank-Read source of dislocations. Successive positions of a dislocation anchored at points A and B and subjected to an increasing force per unit length $f = \tau b$

lines and to follow particular crystallographic orientations. The large increase in energy needed to move a dislocation to the next trough accounts for the very high elastic limits of covalent materials. It is almost impossible to make dislocations slip in these at ambient temperatures, which is why such materials – diamond, carbides, oxides, nitrides, etc. – are used as abrasives or for making cutting tools.

The troughs are much less deep in metals, especially in FCC structures; their presence in BCC metals explains why in these the dislocations tend always to be of screw type, especially when these materials are deformed at low temperature (Caillard 2010a, b).

Peierls (1940) (Nabarro 1967) gave an estimate of the core energy of a dislocation. His method was to regard the two parts separated by the slip plane as semi-infinite elastic media (Fig. 3.45) and to calculate their deformations under the effect of the local continuous inter-atomic forces approximated by a sine function of the inter-atomic distance on either sides of the planes A and B, analogous to the

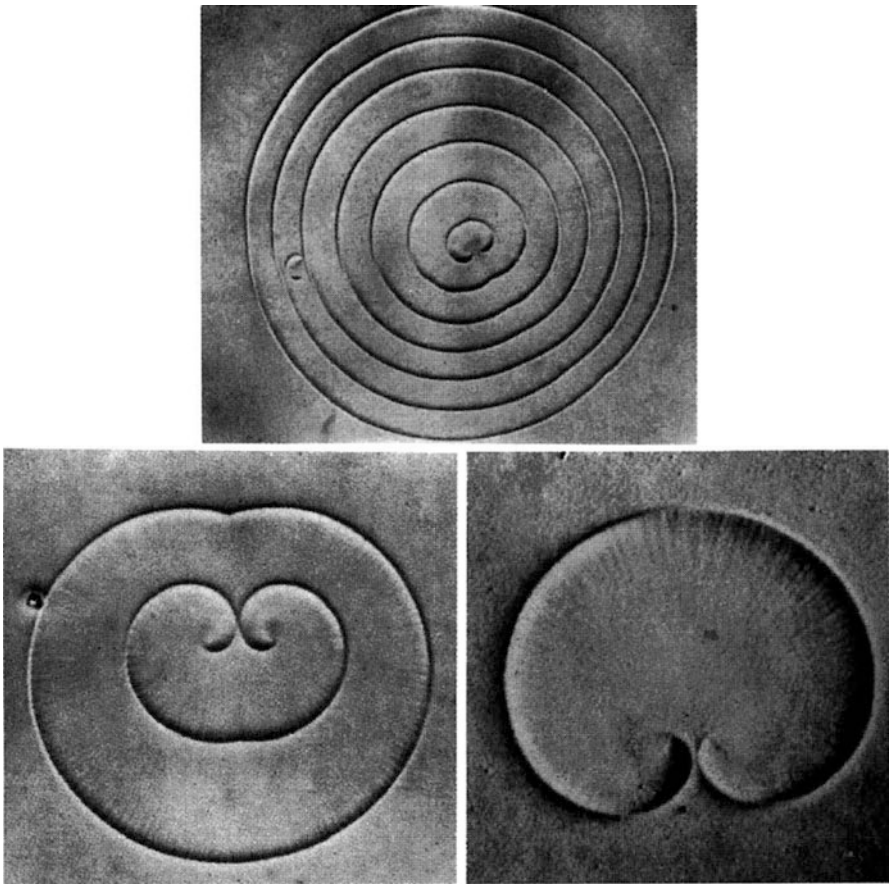


Fig. 3.44 Growth fronts on the surface of electrolytically polished aluminium. There are steps resulting from the emergence of two screw dislocations (Marchini and Wyon 1962)

one used in the calculation of the theoretical shear stress (refer to Sect. 3.3.2.2). The displacements of the atoms of the plane A are (u_1, u_2) and the ones of the atoms of plane B $(-u_1, u_2)$. The relative displacement is thus equal to $2u_1$. The elastic deformation of the upper half is that of an elastic semi-infinite continuum, to the surface of which is applied a shear stress σ_{12}, and is given by:

$$\sigma_{12} = \frac{\mu}{\pi\,(1-\nu)} \int_{-\infty}^{+\infty} \frac{\mathrm{d}u_1\,(\xi)}{\mathrm{d}\xi} \frac{\mathrm{d}\xi}{\xi - x_1} \tag{3.52}$$

This stress is equated to the stress exerted by the B atoms, which is assumed to be a sinusoidal function of period b of the relative displacement and equal to $\mu(2u_1/a)$ when the shear strain u_1/a is small, a being the distance between planes A and B:

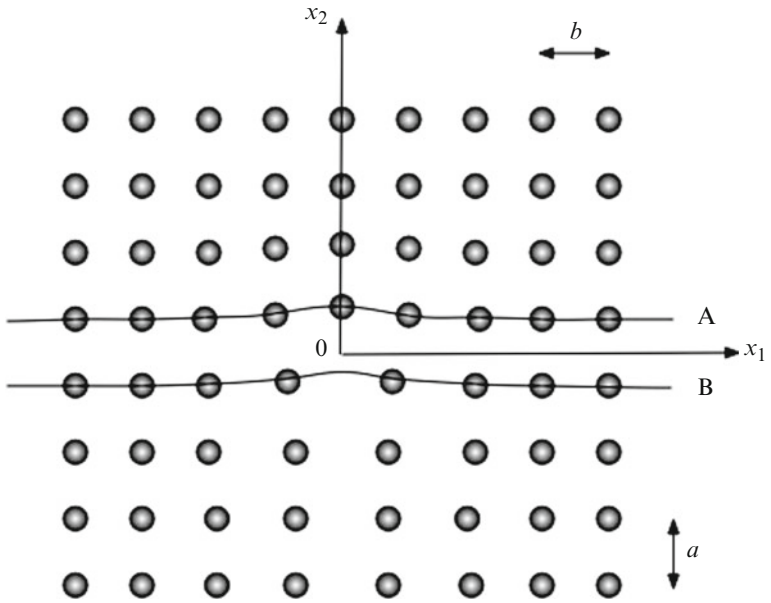


Fig. 3.45 The Peierls's model of a dislocation core

$$\sigma_{12} = -\frac{\mu b}{2\pi a} \sin\left(2\pi \frac{2u_1}{b}\right) \tag{3.53}$$

Hence must be found the solution of the integral equation:

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}\phi\,(\xi)}{d\xi} \frac{\mathrm{d}\xi}{x_1 - \xi} = \frac{1-\nu}{a} \sin 2\pi\phi \tag{3.54}$$

with $\phi = 2u_1/a$

The solution is:

$$u_1 = -\frac{b}{2\pi} \arctan \frac{x_1}{\zeta} \tag{3.55}$$

where $\zeta = \frac{a}{2(1-\nu)}$. 2ζ is considered as the half-width of the dislocation.

For screw dislocations the result is the same, suppressing the Poisson ratio ν in the formulae.

The calculation of the dislocation energy with this model yields the same expression as (3.46) with $r_0 = 2\zeta$. And if we take into account the energy corresponding to the stress field, which annihilates the stresses on the free surfaces (refer to Sect. 3.3.7.1) the best estimate for the dislocation energy per unit length is:

$$E_D = M \log \frac{e^{1/2} r_1}{2\zeta} \tag{3.56}$$

where $M = \mu/2\pi$ for a screw dislocation and $\mu/2\pi(1-\nu)$ for an edge dislocation.

The Peierls's model is not well suited in the case of ionic and covalent crystals. For the first, the slip results in ions of the same sign confronting one another, using a large amount of Coulomb energy; for the second, the bonds are strongly oriented and are quasi-rigid.

Moreover, the assumption of the model is that the forces of interaction between the two parts of the slip plane A and B are spread over a large distance, whereas it yields a narrow core width. This contradiction can be overcome by assuming a flatter inter-atomic potential than the sinusoidal one.

3.3.7.8 Moving Dislocations

So far we considered stationary or slow moving dislocations, so that we could calculate the strain and strain fields around them neglecting completely inertia effects. Observations of dislocations in the electron microscope often show that they move very fast. In order to achieve high strain rates in plasticity, Eq. 3.24 shows that the velocity of dislocations must be quite large. Even for a moderate high strain rate of 10^4 s^{-1}, with a dislocation density of 10^{12} m^{-2}, the dislocation velocity must reach 40 m/s. We then need to reconsider the strain and stress fields of a moving dislocation (Nabarro 1967).

We first consider the case of a screw dislocation. The elastic equation of motion is:

$$\nabla u_3 = \frac{1}{c_2^2} \ddot{u}_3 \tag{3.57}$$

where u_3 is the displacement vector in the direction of the Burgers vector and c_2 is the shear wave velocity $(\mu/\rho)^{1/2}$.

Assuming that $u_3(x_1, x_2)$ is a solution of the static case $\nabla u_3 = 0$, then $u_3[(x_1 - vt)/\beta_2, x_2]$ is a solution of Eq. 3.57, with $\beta_2 = (1 - v^2/{c_2}^2)^{1/2}$. It represents a disturbance travelling in the x_1 direction at a uniform velocity v. Hence, Eq. 3.33, the solution in the static case, becomes:

$$u_3 = \frac{b}{2\pi} \arctan \frac{\beta_2 x_2}{x_1 - vt} \tag{3.58}$$

Now, shifting to polar coordinates, the density of elastic energy $(1/2)\,\mu \left[(\partial u_3/\partial x_1)^2 + (\partial u_3/\partial x_2)^2 \right]$ can be calculated:

$$E_{\mathrm{d}} = \frac{\mu b^2}{8\pi} \frac{1+\beta_2^2}{\beta_2} \log \frac{R}{r_0} \tag{3.59}$$

per unit length of dislocation.

Similarly, the kinetic energy $(1/2)\,\rho \dot{u}_3^2$ is:

$$W_{\mathrm{k}} = \frac{\mu b^2}{8\pi} \frac{1-\beta_2^2}{\beta_2} \log \frac{R}{r_0} \tag{3.60}$$

per unit length.

Hence, the total energy per unit length of the dislocation line, moving at uniform velocity, is:

$$E = E_{\mathrm{d}} + W_{\mathrm{k}} = \frac{\mu b^2}{4\pi} \frac{1}{\beta_2} \log \frac{R}{r_0} \tag{3.61}$$

Considering that to this energy corresponds a force $F_1 = \partial E / \partial x_1$, we can calculate an associated crystal momentum p_{x}:

$$p_{\mathrm{x}} = \int F_1 \mathrm{d}t = \int \frac{\partial E}{\partial x_1} \mathrm{d}t = \int \frac{1}{v} \frac{\partial E}{\partial v} \mathrm{d}v = \frac{Ev}{c_2^2} \tag{3.62}$$

We find equations similar to the ones giving the Lorentz contraction in relativistic dynamics, the shear wave velocity replacing the speed of light. The energy tends to infinity when the velocity of dislocations approaches the shear wave velocity, which is then a limiting velocity.

For an edge dislocation the solution is more complicated. The speed of longitudinal waves $c_1 = [(\lambda + 2\mu)/\rho]^{1/2}$ (λ being the Lamé constant) must be introduced. The displacement $\underline{u}$ in the upper half plane ($x_2 > 0$) is given by:

$$\begin{aligned} \frac{\pi}{b} \frac{v^2}{c_2^2} \bar{u}_1 &= \arctan \frac{\beta_1 x_2}{x_1 - vt} - \alpha^2 \arctan \frac{\beta_2 x_2}{x_1 - vt} \\ \frac{\pi}{b} \frac{v^2}{c_2^2} \bar{u}_2 &= \frac{\beta_1}{2} \log \left[(x_1 - vt)^2 + \beta_1^2 x_2^2 \right] - \frac{\alpha^2}{2\beta_2} \log \left[(x_1 - vt)^2 + \beta_2^2 x_2^2 \right] \end{aligned} \tag{3.63}$$

with $\beta_1 = \left(1 - v^2/c_1^2\right)^{1/2}$ and $\alpha = \left(1 - v^2/2c_2^2\right)^{1/2}$.

The displacement $\underline{u}$ of the lower half plane has a similar solution. We find a superposition of two relativistic disturbances with limiting speeds c_1 and c_2. Their relative proportion depends on the velocity v. As this velocity approaches c_2 the energy is found to go to infinity, not as $1/\beta_2$ in the relativistic manner, but as $1/{\beta_2}^3$. This involves a continuous radiation of energy (Stroh 1962).

We try then to describe, in a semi-qualitative manner, the motion of a dislocation. We expect a screw dislocation to behave like a rod of mass $E/{c_2}^2$ (refer to Eq. 3.62).

When it accelerates it dissipates energy by radiating shear waves. The energy is proportional to log$R$, R being a cut-off radius. For an oscillatory motion of the dislocation it is expected R to be roughly the wavelength of the shear wave. For free motion in the absence of applied stress we expect R to be roughly equal to $c_2 t$, and thus a mass of the dislocation proportional to log$t$. The dislocation ought to slow down as the region of disturbance, which travels with it, expands.

3.3.8 Interaction of Dislocations with Other Dislocations and Interfaces

The stress field of dislocations exerts a force on neighbouring dislocations according to the Peach and Koehler formula (3.29). These interactions, attractive or repulsive, can create special arrangements of dislocations, which are important for strain-hardening. So are the interactions of dislocations with boundaries, which introduce discontinuities in the stress and displacement fields.

3.3.8.1 Interactions Between Dislocations

We consider two dislocations of Burgers vectors $\underline{b}_1$ and $\underline{b}_2$, and we determine the force $\underline{f}_{12}$, which the first one exerts on the second according to the Peach and Koehler formula (3.29). As the stress field of the first dislocation decreases when the distance increases (Eqs. 3.35 and 3.43), in general a torque will be applied on the second one. It is not the case when the two dislocations are parallel. We begin then by studying this configuration, simplifying even more by considering that they lie on the same glide plane (this is always the case for two parallel screw dislocations) a distance d apart. The Peach and Koehler formula reduces to $f_{12} = \tau_1 b_2$ or:

$$f_{12} = M \frac{b_1 b_2}{d} \tag{3.64}$$

where $M = \mu/2\pi$ for a screw dislocation and $\mu/2\pi(1-\nu)$ for an edge dislocation.

Thus the force is repulsive (positive) if the dislocations have the same sign, attractive if the signs are opposite.

Two repulsive parallel dislocations on the same slip plane can remain in equilibrium at a finite distance d, when they create either a stacking fault (such as a Shockley dislocation) or an anti-phase boundary in an ordered crystal. The equilibrium distance d is such that the stacking fault energy, or the anti-phase boundary energy, compensates the repulsive energy of the two dislocations. Measuring this distance d is a way to estimate the stacking fault or the anti-phase, energy. (We will elaborate later when we discuss the Suzuki effect.)

From the expression (3.43b) for the stress field of a dislocation we can easily calculate the interaction force between two parallel edge dislocations with slip

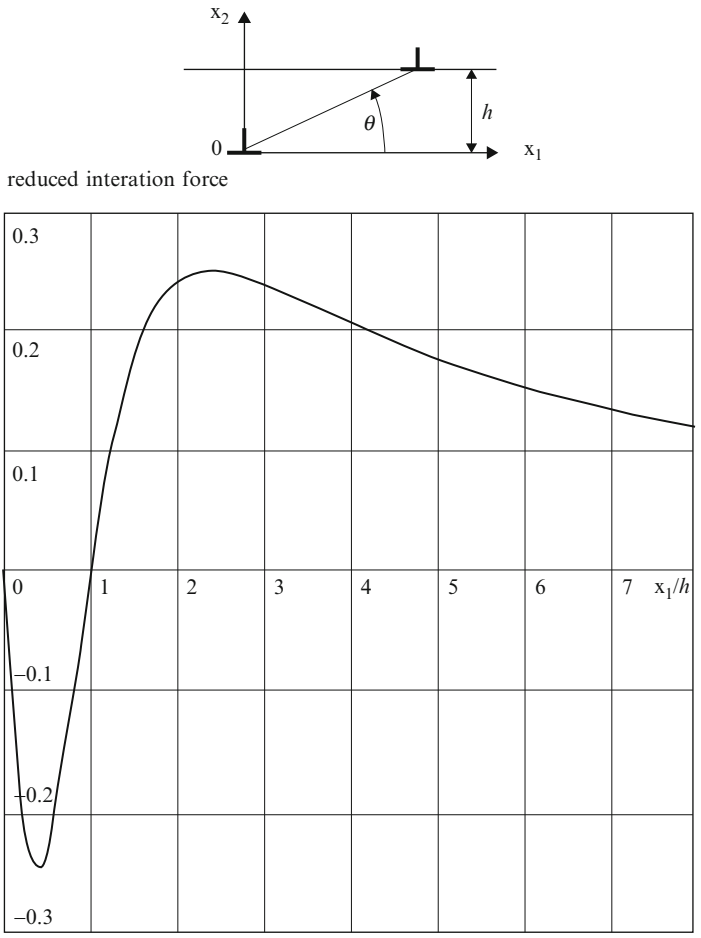


Fig. 3.46 Reduced force, $\frac{2\pi(1-\nu)}{\mu b^2} f_{12}$, as a function of x_1/h for two parallel edge dislocations of the same sign on two parallel slip planes a distant h apart

planes a distance h apart (Fig. 3.46 and Exercise in Volume III):

$$f_{12} = \frac{\mu \underline{b}_1 \cdot \underline{b}_2}{8\pi\,(1-\nu)} \frac{\sin 4\theta}{h} \tag{3.65}$$

If the dislocations have the same signs they repel one another when $|x_1| > |x_2|$; otherwise they will attract and will arrange themselves so that one is above the other ($x_1 = 0$). If the signs are opposite they attract if $|x_1| > |x_2|$ and arrange themselves so that $x_1 = x_2$, forming what is called a *dipole*.

From the Peach-Koehler formula (3.29) the climb force is easily found to be $-\sigma_{11} b_2$.

In general when two dislocations of Burgers vectors $\underline{b}_1$ and $\underline{b}_2$ meet they can interact to create a new dislocation of Burgers vector $-(\underline{b}_1 + \underline{b}_2)$ (Fig. 3.18). The energy of the two dislocations before they combine is proportional to ${b_1}^2 + {b_2}^2$ per unit length, whereas the energy of the recombination is proportional to

$$(\underline{b}_1 + \underline{b}_2)^2 = b_1^2 + b_2^2 + 2\underline{b}_1.\underline{b}_2$$

When the scalar product $\underline{b}_1.\underline{b}_2$ is negative the recombination gains energy and is favourable: the junction is attractive. On the contrary, when $\underline{b}_1.\underline{b}_2$ is positive the junction is repulsive. Consider two dislocations on two different glide planes in a FCC crystal. When they meet, if the junction is attractive, they combine along the intersection of the glide planes to form a stair rod dislocation (refer to Sect. 3.3.4.2). In the Lomer-Cottrell lock, the two dislocations which meet are dissociated for instance in the glide planes (111) and $(11\bar{1})$ as $\frac{a}{2}\left[01\bar{1}\right] = \frac{a}{6}\left[\bar{1}2\bar{1}\right] + \frac{a}{6}\left[11\bar{2}\right]$ and $\frac{a}{2}[101] = \frac{a}{6}\left[2\bar{1}1\right] + \frac{a}{6}[112]$. The two Shockley partials $\frac{a}{6}\left[\bar{1}2\bar{1}\right]$ and $\frac{a}{6}\left[2\bar{1}1\right]$ attract each other and combine to form the sessile dislocation $\frac{a}{6}[110]$.

If, on the contrary, the junction is repulsive, the two dislocations after crossing glide away; each one has acquired a small segment, equal to $\underline{b}_2$ for the first dislocation and to $\underline{b}_1$ for the second. The segment $\underline{b}_2$ on the first dislocation can be decomposed in three components: one along the dislocation line of the first dislocation, one in its slip plane producing a *kink* and the third perpendicular to this slip plane called a *jog* (Fig. 3.47). In a FCC structure, referring to the Thompson tetrahedron (Fig. 3.30), we can consider a dislocation of Burgers vector $\underline{AB}$ in the d glide plane. Crossing with repulsive dislocations of Burgers vectors $\underline{DA}$, $\underline{DB}$, $\underline{DC}$, or $\underline{CD}$, will produce a jog $\underline{D\delta}$. This jog cannot glide, so that the gliding dislocation will stretch a trailing dipole with branches so close that they will annihilate by the emission of vacancies or interstitials. If however several dislocations cross in succession, the jog will be large enough for the dipole to be stable. Dipoles are often observed in electron microscopy.

Plastic deformation, which involves the displacement of a large number of dislocations and thus many crossing of repulsive dislocations, results in the emission of a great number of *vacancies* and *interstitials*. This super-saturation of point defects can be reduced by recombination of vacancy-interstitial pairs. But it also contributes to dislocation climb. The return to equilibrium requires the thermal activated diffusion of these point defects, faster for interstitials than for vacancies.

3.3.8.2 Dislocations Pile-Ups

We consider n parallel dislocations in the same slip plane under a shear stress τ. Each one is pushed by a force τb and we assume that they are held by some obstacle at point $x = 0$. This configuration is a pile-up. They were observed particularly in austenitic stainless steels as shown in Fig. 3.48. We want to find at which distance from each other lie the dislocations and the size L of the pile-up (Fig. 3.49).

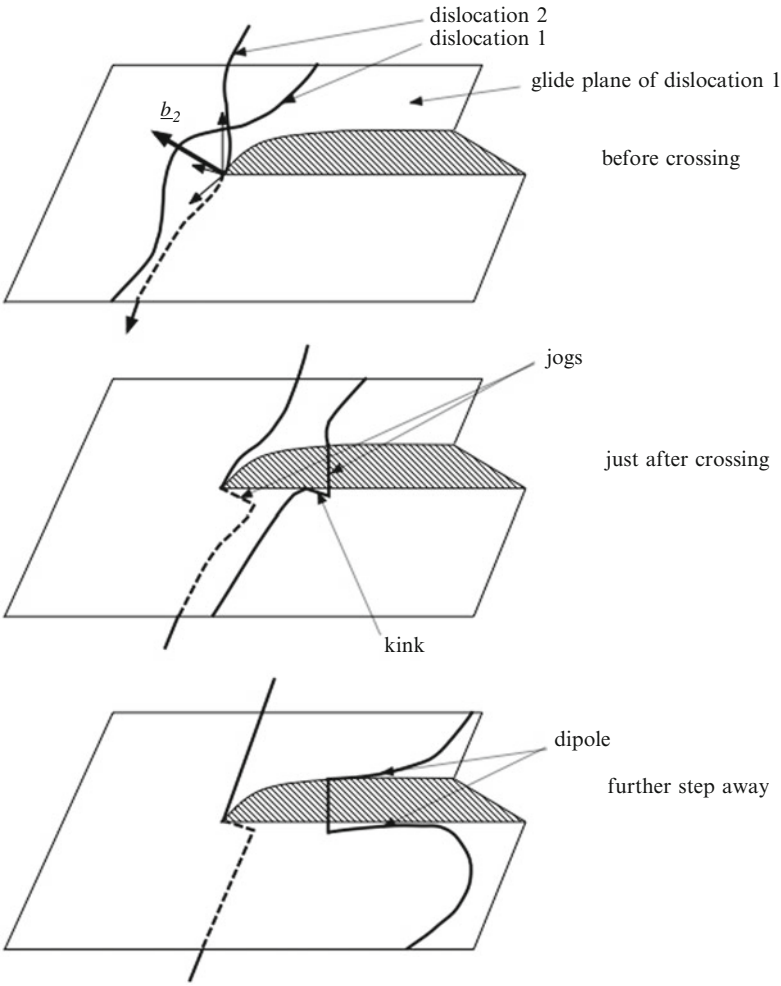


Fig. 3.47 Formation of jogs and kinks at the crossing of two repulsive dislocations

If the j'th dislocation with abscissa x_j is moved a distance $-\delta x$ the following $(n-j)$ also move through this distance, requiring an amount of work $\delta W = (n - j + 1)\,\tau b\,\delta x$ to be done. Hence the j'th dislocation experiences a force $(n-j+1)\tau\, b$ exerted by the $(n-j)$ following dislocations. In particular, the obstacle against which the dislocations are piled up experiences a force $n\tau b$.

The distance $x_2 - x_1$ is determined by the equilibrium between the force that the first dislocation exerts on the second and that exerted on it by the $n-2$ others:

$$\frac{Mb^2}{x_2 - x_1} = (n-1)\,\tau b \tag{3.66}$$

where $M = \mu/2\pi(1-\nu)$ for an edge dislocation, $\mu/2\pi$ for a screw dislocation.

Fig. 3.48 Pile-ups of dislocations in an austenitic stainless steel. Their slip planes are inclined with respect to the thin foil. The bar represents 1 μm

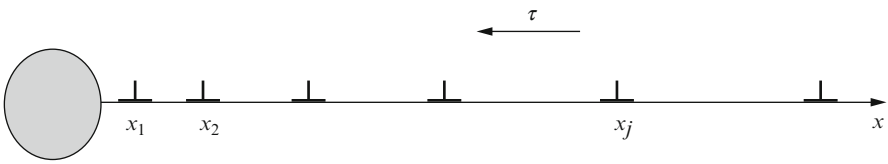


Fig. 3.49 Pile up of dislocations against an obstacle at $x = 0$

As this distance $x_2 - x_1$ is small, we can calculate $(x_3 - x_2)$ on the assumption that the first and second dislocations act together like a single dislocation with Burgers vector $2b$:

$$x_3 - x_2 = \frac{2Mb}{(n-2)\,\tau} \tag{3.67}$$

and continuing, we have in general:

$$x_{j+1} - x_j = \frac{jMb}{(n-j)\,\tau} \tag{3.68}$$

The distance between successive dislocations in the pile-up thus increases with the distance from the head. Taking the distance $(x_n - x_{n-1})$ to be or the order of $L/2$, this gives:

$$n \simeq \frac{\tau L}{2Mb} \tag{3.69}$$

The same result is obtained by an exact calculation of the positions of the dislocations by solving the n equations:

$$\sum_{j=1, j\neq i}^{n} \frac{Mb^2}{x_i - x_j} = \tau b \tag{3.70}$$

It can easily be shown that the first dislocation located at the head of the pile-up is submitted to a stress equal to $n\tau$. This explains the stress concentration produced by dislocations pile-up.

3.3.8.3 Low Angle Boundaries

We have found that two edge dislocations of the same sign on two parallel slip planes can find a stable configuration by aligning one above the other at a distance h. A third edge dislocation of the same sign can on its turn rest on the same alignment. By adding in this way many edge dislocations of the same sign a distance h apart from one another we build a *dislocation wall* forming a low angle boundary. It separates two *subgrains*, which present a tilt disorientation of angle $\theta = b/h$ (Fig. 3.50).

In the same way, a regular arrangement of screw dislocations creates a low angle twist boundary. A regular combination of mixed dislocations can achieve any misorientation of the subgrains (Fig. 3.51).

The calculation of the energy of these configurations gives the value of the energy of the sub-boundary as a function of the misorientation of the subgrains. Here we consider a simple *tilt boundary* formed by regularly spaced edge dislocations a distance h apart; and calculate the energy of a dislocation belonging to this wall. The other dislocations in the wall will screen the one we are considering from the surrounding stress field to a distance of the order of $h/2$; the energy associated with it is therefore from Eq. 3.46, approximately,

$$E_{\mathrm{D}} = \frac{\mu b^2}{4\pi\,(1-\nu)} \log\left(\frac{h}{2r_0}\right) \tag{3.71}$$

where r_0 is the core radius.

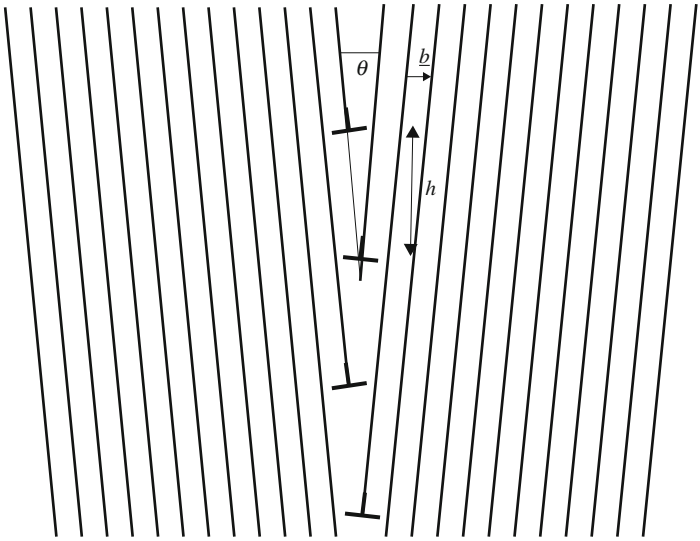


Fig. 3.50 Tilt low angle boundary

As there are $1/h$ dislocations per unit length of the wall, the sub-boundary energy per unit area is:

$$\gamma_{\mathrm{i}} = \frac{\mu b}{4\pi\,(1-\nu)}\theta \log\left(\frac{b}{2r_0\theta}\right) \tag{3.72}$$

Taking $h = 20b$, *i.e.* $\theta = 2.86°$, one finds $\gamma_{\mathrm{i}} = 120$ mJ/m^2 which is a reasonable value.

The calculation can be extended to the case of the *twist boundary* formed by screw dislocations, when the sub-boundary will be in torsion; and even to more complex assemblies of dislocations.

The result shows that the energy of the wall increases with increasing disorientation up to a maximum, γ_{im} equal to $b/2r_0\mathrm{e}$, e being the base of the natural logarithms, after which it falls.

Figure 3.52 shows that the predicted form of the variation agrees with what is found experimentally.

3.3.8.4 Interaction of a Dislocation with a Surface or an Interface

Suppose there is a dislocation close to a free surface. On the latter the stress vector must vanish and therefore the expressions (3.35) and (3.43) for the stress field of the dislocation must be modified. A simple way to do this is to place a fictitious

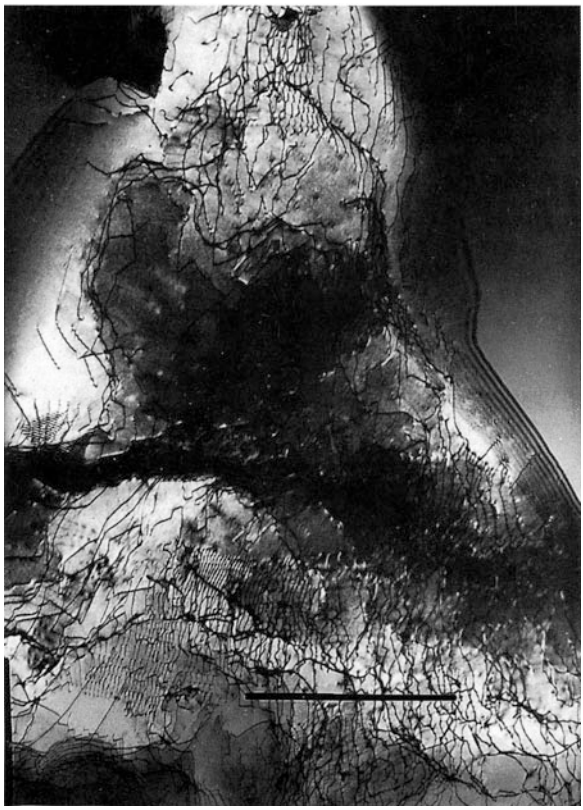


Fig. 3.51 Wall of dislocations starting to form in an austenitic stainless steel annealed for 24 h at 900°C after work-hardening

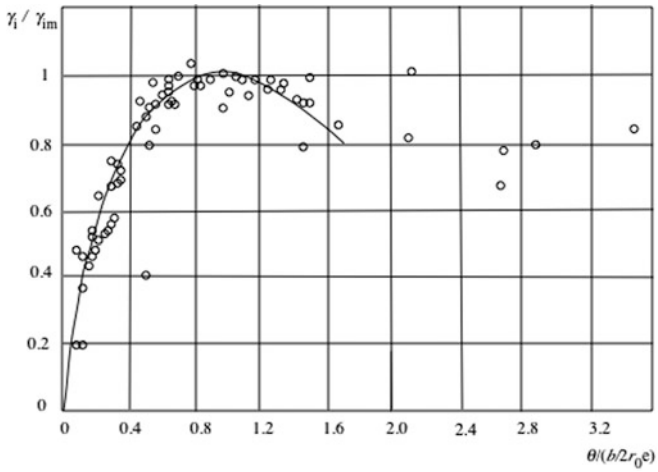


Fig. 3.52 Comparison of the reduced boundary energy γ_i/γ_{im} as a function of the disorientation angle $\theta/(b/2r_0e)$ with the formula (3.65). $(b/2r_0e)$ is treated as an adjustable parameter (After Nabarro 1967)

dislocation in such a way as to cancel the stress vector of the real dislocation on the free surface. If the latter is a screw dislocation parallel to the surface, the fictitious dislocation can be another screw dislocation with Burgers vector of opposite sign, located symmetrically with respect to the surface. The real stress field is the resultant of the superposition of the fields of the two dislocations; the second (fictitious) dislocation is the image of the first, which experiences an attractive *image force* and is therefore attracted by the surface.

The same process applied to an edge dislocation cancels the stress component normal to the surface, but not the shear. It can however be shown that the further stress field that has to be applied in order to cancel this exerts no force on the dislocation, and therefore an edge dislocation also is attracted to the surface by the image force (Head 1953).

The problem becomes more complicated if we consider not a free surface but the interface between two media with shear moduli μ_1 and μ_2. For a screw dislocation parallel to the interface in the medium of modulus μ_1 the solution is given by an image dislocation with Burgers vector $\beta\underline{b}$, symmetrically placed in the medium of modulus μ_1, together with an image with vector $\gamma\underline{b}$ in the medium of modulus μ_2, super-imposed on the real dislocation. β and γ are given by:

$$\begin{aligned} \beta &= (\mu_2 - \mu_1)/(\mu_2 + \mu_1) \\ \gamma &= 2\mu_1/(\mu_2 + \mu_1) \end{aligned} \tag{3.73}$$

Thus a screw dislocation is attracted to the interface if $\mu_1 > \mu_2$ and repelled otherwise. The method that has led to this conclusion is not strictly applicable to an edge dislocation, but it provides a reasonably good approximation.

It follows from these results concerning interactions with surfaces that when a thin sheet of material is examined by the electron microscope a significant number of the dislocations must have disappeared. Further, a surface layer of oxide with modulus greater than that of the metal will generate a force that opposes the slipping out of the dislocations and which will therefore have a hardening effect. This is a partial explanation of the *Rehbinder effect*, that exposure to an active medium makes a material more easily deformable[11] (Rehbinder 1947; Nabarro 1967).

3.3.9 Twinning

Twinning plays a key role in the plastic deformation of alloys, which have few slip systems, such as the HCP materials (Zn, Mg, Ti, Zr, Be). When we described slip systems in HCP (refer to Sect. 3.3.5.2), we drew attention on the difficulty to activate slip in directions out of the basal plane and thus the importance of deformation possibilities offered by twinning. In FCC materials that have low stacking fault energy twinning is facilitated. Furthermore, twinning is part of the mechanism

[11] Petr Alexandrovitch Rehbinder (1898–1972) was a Russian scientist.

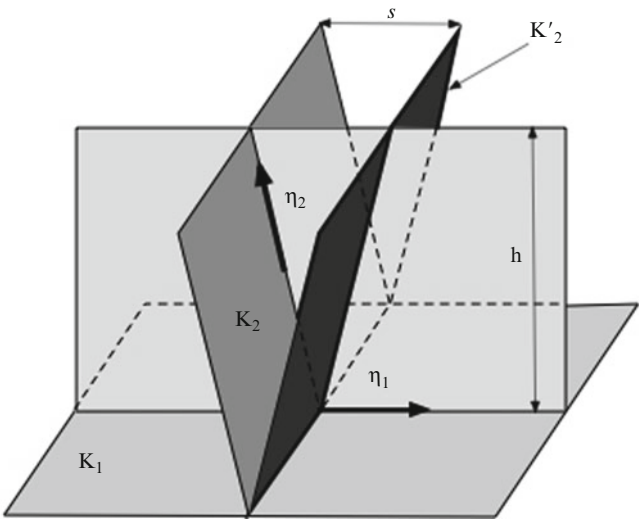


Fig. 3.53 Characteristic elements of twinning

of the martensitic transformation. Many twins are found in stainless steels, in manganese (Hadfield) steels, in modern TWIP (Twinning induced plasticity) steels (see Sect. 3.4.5), in Cu-Al alloys and in nickel-based alloys. In BCC materials, the difficulty to move screw dislocations can be compensated by twinning, which plays an important role in ferritic alloys.

3.3.9.1 Twin Elements

A twin within a crystal is a piece of the same material whose unit cell is related to the unit cell of the parent crystal by symmetry elements. In this section we are interested in twinning as far as it results in plastic deformation. We will limit ourselves to the most common cases of industrial interest, which are found in FCC, BCC and HCP alloys. The symmetry elements reduce then to mirror symmetry: the twin is the image in a mirror of the parent crystal. The twin boundary, also called the composition surface, is the mirror plane. We will first give a brief general description of twinning and we will calculate the strain produced by a twin.

A twin is characterised by the crystallographic planes K_1 and K_2, which remains unchanged, and by the twinning direction η_1 of K_1. The twin results from the shear of the parent crystal in this twinning direction (Fig. 3.53). The shear plane is normal to K_1 and contains the direction η_1, intersecting K_2 along a line whose direction is η_2.

The crystallographic plane K_2 is then flipped to K'_2. The shear is such that $\gamma = s/h$.

Twins of the first kind are such that K_1 and η_2 have rational Miller indexes, while twins of the second kind are such that K_2 and η_1 have rational Miller indexes. In the structures of interest to us all indexes are rational ({111} in FCC and {112} in BCC crystals).

A crystallographic structure can contain K_1 planes of the same kind, of different orientations, such as {111} planes for instance. These are called variants.

3.3.9.2 Twinning Deformation

The formation of a twin produces large displacements requiring calculation of strains in term of the Green-Lagrange tensor. Note that the displacement of individual atoms remains small, but their cooperative movement induces a large deformation.

The coordinates axes are Ox_1 along the η_1 direction, Ox_2 perpendicular to the shear plane and Ox_3 perpendicular to the K_1 plane (Fig. 3.53). Let the initial coordinates of a particular point be a_i of the plane K_2 and the corresponding ones after twinning x_i in plane K'_2.

$a_i = (a_1 = -a_3 s/2h,\ a_2 = 0,\ a_3)$ and $x_i = (x_1 = -a_1,\ x_2 = 0,\ x_3)$. Thus the displacement is $\underline{u} = (u_1 = a_3 s/h,\ u_2 = 0,\ u_3 = 0)$.

The Green-Lagrange tensor $2\Delta_{ij} = \frac{\partial u_i}{\partial a_j} + \frac{\partial u_j}{\partial a_i} + \frac{\partial u_k}{\partial a_i}\frac{\partial u_k}{\partial a_j}$ has only two elements different from zero:

$$\Delta_{13} = s/2h = \tan\theta \quad \text{and} \quad \Delta_{33} = (1/2)\,(s/h)^2 \tag{3.74}$$

The linear dilatation in a direction $\underline{\mathrm{d}a}$ is given by:

$$|\mathrm{d}\underline{x}|^2 - |\mathrm{d}\underline{a}|^2 = 2\Delta_{ij}\,\mathrm{d}a_i\,\mathrm{d}a_j = 2\frac{s}{h}\mathrm{d}a_1\mathrm{d}a_3 + \left(\frac{s}{h}\right)^2 \mathrm{d}a_3{}^2 \tag{3.75}$$

so that:

$$\left(\frac{\mathrm{d}\underline{x}}{\mathrm{d}\underline{a}}\right)^2 = 1 + 2\frac{s}{h}\sin\alpha\cos\alpha + \left(\frac{s}{h}\right)^2 \sin^2\alpha \tag{3.76}$$

with α the angle between the Ox_1 axis and the vector $\underline{\mathrm{d}a}$.

The largest dilatation is thus found in a direction such that:

$$\tan 2\alpha = -\frac{1}{\tan\theta} \tag{3.77}$$

with $\tan\theta = s/2h$. Hence there are two possible directions at $\pi/2$ from each other: $\alpha = \pi/4 - \theta/2$ and $\alpha = 3\pi/4 - \theta/2$ (note that in the absence of Δ_{33} the deformation of the twin would be pure shear and those directions would be $\pm\pi/4$). It is the second direction, which corresponds to the maximum dilatation; it is given by:

$$\frac{|\mathrm{d}x|}{|\mathrm{d}a|} = 1 + \lambda_{\max} = \tan\theta + \left(1 + \tan^2\theta\right)^{1/2} \tag{3.78}$$

3.3.9.3 Twins in Various Crystallographic Structures

In FCC structures we have seen in Sect. 3.3.4.1 that a stacking fault occurs when dense planes {111} are shifted to wrong positions, the stacking becoming ΔΔΔΔ∇ΔΔΔΔ or ΔΔΔΔ∇∇ΔΔΔΔ. A twin is created by the repetition of this, yielding a stacking ΔΔΔΔΔΔ∇∇∇∇∇∇ or PQRPQRQPRQPR. The twin is the image of the parent structure in the mirror plane (111). Thus the twinning elements are $K_1 = (111)$, $\eta_1 = [11\bar{2}]$, $K_2 = (11\bar{2})$, $\eta_2 = [111]$. Figure 3.54 shows the way twinning occurs in the shear plane $(1\bar{1}0)$. The shear ratio *s*/*h* is equal to 0.707 and the angle θ to 35°. The maximum elongation is equal to 0.414.

In BCC the twinning mechanism (Fig. 3.55), in a similar way, can be considered as a succession of stacking faults on {112} planes described in Sect. 3.3.5.2 and Fig. 3.35. In that case $s/h = 1/\sqrt{2} = 0.707$ and $\theta = 35°$.

In HCP are observed various twin systems. Tension twins are $\{10\bar{1}2\}\ <10\bar{1}1>$ and $\{11\bar{2}1\}\ <11\bar{2}6>$; compression twins are $\{11\bar{2}2\}\ <11\bar{2}3>$. The nature of these mechanical twins in HCP crystals depends on the value of *c*/*a* ratio. The twinning shear depends on the ratio *c*/*a*. It can easily be shown that in the case where the twinning plane is $(10\bar{1}2)$, the shear is given by:

$$\Delta_{13} = \frac{3 - c^2/a^2}{c/a\sqrt{3}} \tag{3.79}$$

Thus for $c/a = \sqrt{3}$, the shear strain is equal to zero. When *c*/*a* is larger (smaller) than $\sqrt{3}$, the shear strain occurs in opposite directions, but it is always small. The largest amounts of shear strain occur for materials with extreme values of the *c*/*a* ratio, *i.e.* Be ($c/a = 1.57$) and Cd ($c/a = 1.89$). Twinning shear is then of 0.197 in Be and 0.170 in Cd, while the maximum elongation is of the order of 10% in Be and 6.5% in Cd (Jaoul 2008).

3.4 Hardening Mechanisms

3.4.1 Introduction

In Sect. 3.3 we first described tests on single crystals, which show that plastic deformation takes place when a critical shear stress, much smaller than the theoretical one, is reached. This led us to introduce displacement of dislocations and

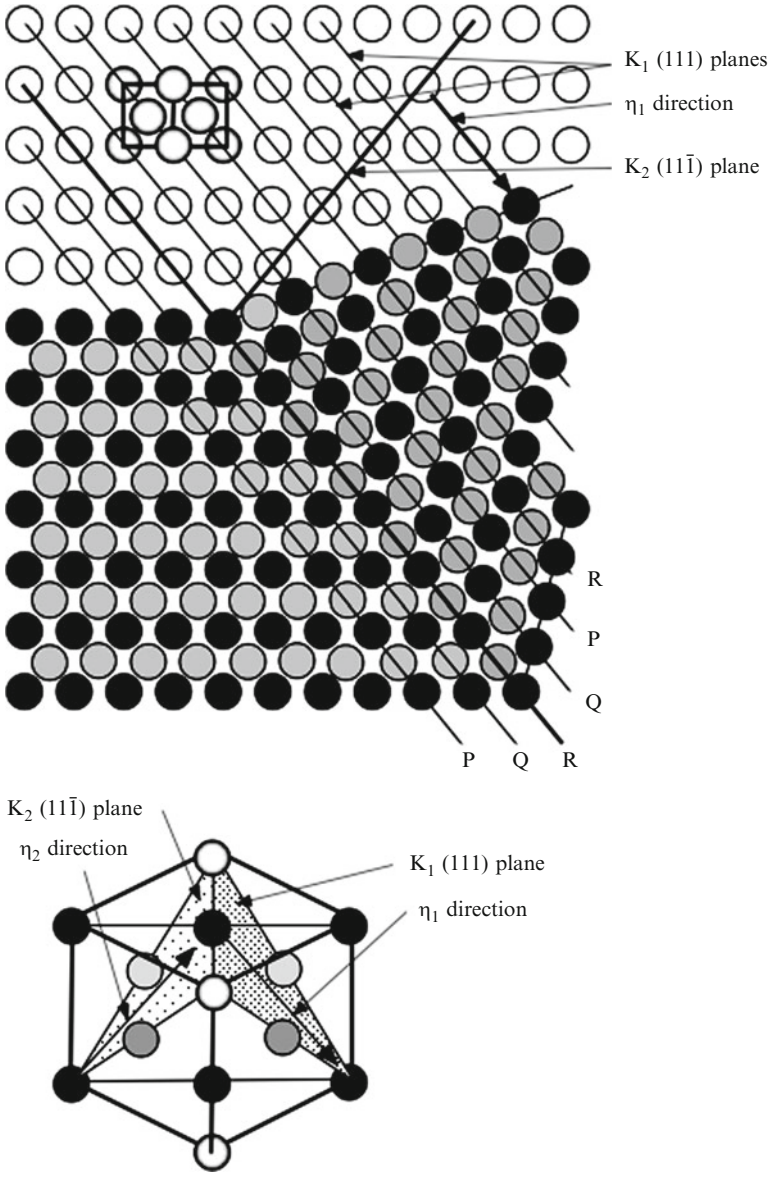


Fig. 3.54 Twinning in FCC crystallographic structure. The plane K_1 is a {111} plane and the η_1 direction is a <112> direction

twinning as needed for the explanation of plastic deformation. We then described in detail these crystallographic defects, providing the tools to understand their behaviour in relation with plastic deformation. However, so far we have not studied the critical stress itself. We now need to know what are its conditioning mechanisms.

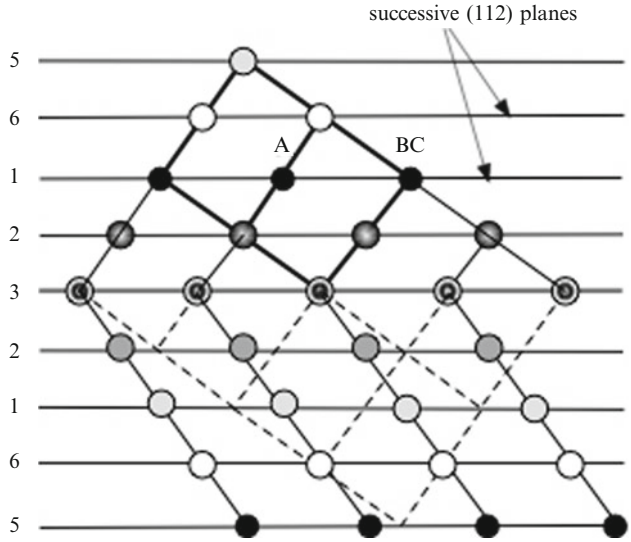


Fig. 3.55 Twinning in BCC crystallographic structure. The plane K_1 is a {112} plane and the η_1 direction is a <111> direction

It will allow us to understand how the yield strength and more generally the proof strength vary, how they can be increased. When, in Chap. 4, we will study time-dependent phenomena, it will be the activation energies concerning these mechanisms which will come into play. Later, in the second volume, this will also help in the mastering of damage and fracture. Thus, these mechanisms are essential ingredients in materials design, the art of conceiving materials for a given purpose.

Understanding of the nature and effects of the various obstacles is also needed for a sound formulation of constitutive equations, which enter in the calculation of stress and strain distributions in components and structures.

The movement of dislocations, the creation and development of twins (or more generally of transformed phases) meet obstacles of various natures. They constitute the root of hardening mechanisms that we must now study. We can picture a dislocation as an extended rope, which moves in a mountainous landscape: it meets hills and peaks and on top of those rocks of various heights. The applied force on the dislocation makes it climb these various obstacles. Temperature forces the rope to jump up and down, so it can help to overcome the smaller obstacles. At 0 K all obstacles must be overcome by the applied force only. At high enough temperature, remain only the big obstacles denoted athermal. The various obstacles can be classified in several categories as shown in Table 3.5.

At this stage, even though strictly speaking, plastic deformation is time independent, we will find convenient, when dealing with a particular mechanism, to give the elements entering the thermally activated strain rate equation. The formalism of

Table 3.5 Classification of hardening mechanisms

Types of obstacles	Type of mechanism	Representative examples
1. Localised	Thermally activated	Repulsive dislocation trees
		Solute atoms
		GP zones
	Thermally activated or athermal	Attractive junctions
		Coherent precipitates
		Radiation damage
	Athermal	Incoherent precipitates
		Long range stress fields
2. Linear	Thermally activated	Peierls force
		Cross-slip
		Cottrell-Lomer dissociation
	Thermally activated or athermal	Suzuki unlocking
	Athermal	Short range order
		Long range order
3. Volumetric	Athermal	Thermoelastic
		Phonon scattering
		Phonon viscosity
		Electron viscosity

thermal activation of plastic deformation will be treated in detail in Chap. 4 of this book. Let here simply state that the strain rate can in general be written:

$$\dot{\varepsilon} = \dot{\varepsilon}_0 \exp\left(-\frac{Q_0 - \sigma b A^*}{\mathrm{k}T}\right) \tag{3.80}$$

where k is the Boltzmann[12] constant, T the absolute temperature, σ the applied stress, Q_0 the activation energy in the absence of stress to overcome a particular obstacle and A^* the activation area associated with it. It is the area that a dislocation sweeps to overcome the obstacle in question.

The obstacles of types 1 and 2 can be further classified as: first, obstacles that come from the crystallographic nature itself, second, obstacles constituted by foreign atoms. Volumetric obstacles of type 3 cause viscous behaviour that will be briefly treated in Chap. 4.

We will proceed by increasing size of obstacles. So in the first category we will treat in succession the obstacles provided by atomic binding, by other dislocations and by grain boundaries. In the second category, atoms dispersed in solid solution, precipitates and inclusions. Afterwards, we will study twinning and transformation plasticity. These mechanisms will be illustrated by characteristic examples of industrial materials. Finally, we will see how, based on hardening mechanisms, the constitutive equations can be formulated.

[12]Ludwig Boltzmann (1844–1906) was an Austrian physicist.

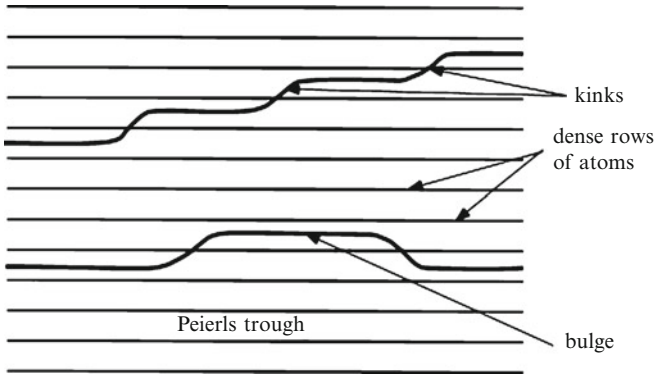


Fig. 3.56 Kinks

3.4.2 Obstacles of Crystallographic Nature

3.4.2.1 The Peierls-Nabarro Force[13]

In Sect. 3.3.7.7, we gave the definition and properties of a *Peierls trough* (Fig. 3.45). We also explained how Peierls calculated the half width ζ of a dislocation (Eq. 3.55). He then calculated the variation of the energy as the dislocation is displaced out of the trough (Peierls 1940; Nabarro 1947):

$$E_{\mathrm{PN}} = \frac{\mu b^2}{4\pi\,(1-\nu)}\left[1 + 2\,(\cos 4\pi\alpha)\exp\left(-\frac{4\pi\zeta}{b}\right)\right] \tag{3.81}$$

where α is the portion of the Burgers vector b that the centre of the dislocation is moved, and $\zeta = a/2(1-\nu)$. For a screw dislocation ν is taken equal to zero.

The corresponding force acting on the dislocation has an amplitude:

$$\tau_{\mathrm{PN}} = -\frac{1}{b}\frac{\mathrm{d}E_{\mathrm{PN}}}{\mathrm{d}\alpha}\bigg|_{\max} = \frac{2\mu}{1-\nu}\exp\left(-\frac{4\pi\zeta}{b}\right) \tag{3.82}$$

The amplitude is the smaller, the larger a/b, that is the denser the slip plane. This model yields a very low critical shear stress of the order of $10^{-5}\,\mu$ for a close-packed lattice. It is even smaller for partial dislocations.

A dislocation, which is not perfectly aligned with a dense row of atoms, is composed of segments lying on a Peierls trough joined by *kinks* (Fig. 3.56). It is

[13] In 1967 Rudolph Peierls claimed that this force should have been called the Orowan-Nabarro force (Rosenfield et al. 1968).

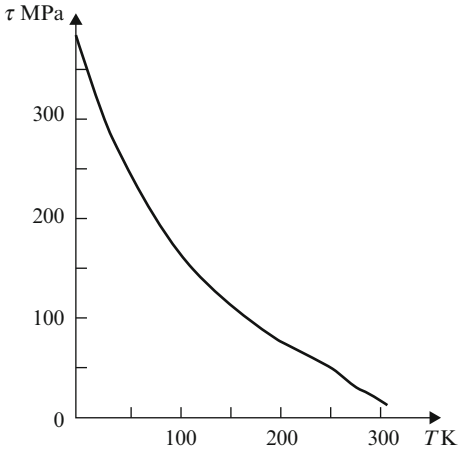


Fig. 3.57 Critical resolved shear stress on {110} planes in pure iron as a function of temperature (According to Caillard 2010)

then much easier to displace the dislocation by the sideway propagation of kinks, than by rigid jump as a whole, yielding a still lower critical shear stress.

Read (1953) calculated the energy of a kink to be:

$$E_{\mathrm{k}} \simeq \frac{2.8\mu b^3}{\pi} \exp\left(-\pi \frac{\zeta}{b}\right) \tag{3.83}$$

The activation energy, when a strong Peierls force places down dislocations in the troughs, corresponds to the creation of two kinks of opposite signs and thus is $2E_{\mathrm{k}}$, while the activation area would be of the order of a few b^2.

The yield strength of pure BCC metals is characterised by a large increase when the temperature is lowered and by a dissymmetry between tension and compression (Fig. 3.57). Screw dislocations are found to be difficult to move whereas this is needed for plastic deformation to take place. It is usually linked to control by the Peierls stress. In Sect. 3.3.7.7, we saw that it has been considered that a particular dissociation of the core could explain the particular behaviour of BCC metals, but that recent observations and calculations cast much doubt on this model. Screw dislocations in pure iron propagate on {110} planes, and at low temperature by jerky motion (Caillard 2010). Figure 3.56 shows the critical resolved shear stress on {110} plane of pure Fe as a function of the temperature. A definite break is observed at 250 K. The interpretation is based on two mechanisms in parallel: above 250 K for a screw dislocation to move a double kink (Fig. 3.56) must be nucleated, which requires an energy in the absence of stress of 11 meV per distance of line b; this corresponds to a classical model. (The kink energy E_{k} as given by Eq. 3.83 is equal to 17 meV). But below 250 K the critical event would be the creation of a bulge of metastable glissile dislocation; the energy needed to extract the sessile dislocation from the Peierls trough so as to give it this glissile yet unknown configuration is 26 meV per b, of the order of the Peierls energy.

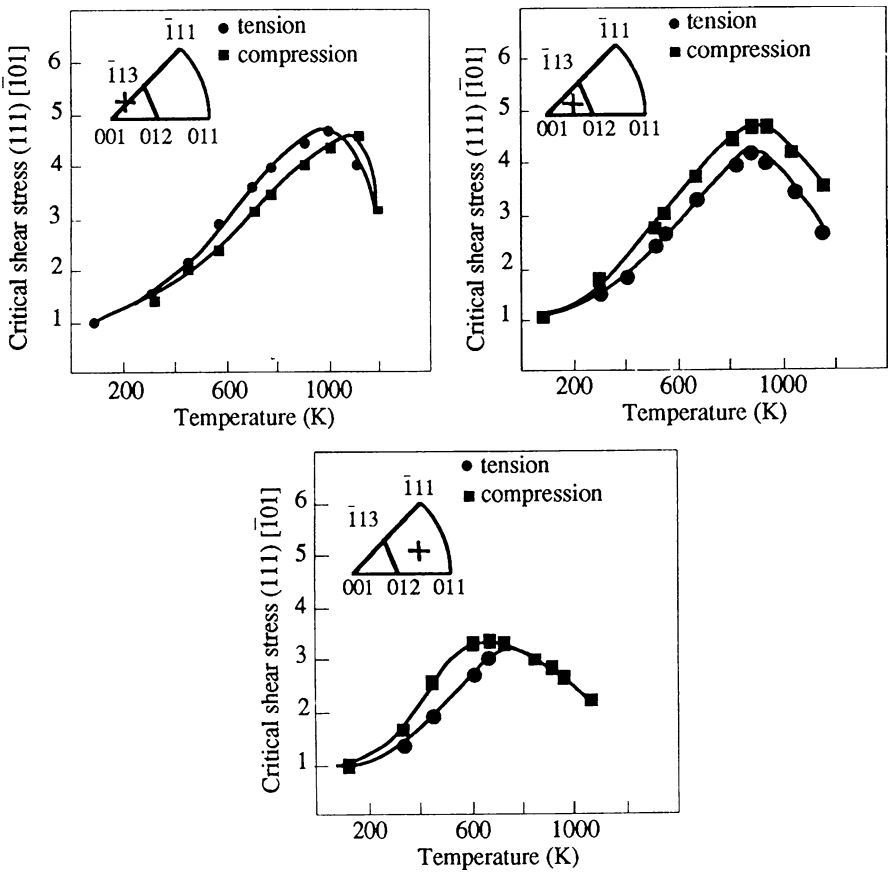


Fig. 3.58 Critical resolved shear stress for Ni_3Al (Pope and Ezz 1984)

This effect of anisotropy on the dissociation of dislocations is encountered in other crystal structures, such as Ll_2 (see Annex 1), a pseudo-FCC, phase γ', of approximate composition Ni_3Al; this is responsible for the hardening of nickel-based superalloys (Fig. 3.58). For further discussion see Gil Sevillano (1993).

3.4.2.2 Interaction with Other Dislocations; Strain-Hardening or Work-Hardening

The dislocations in a crystal that has not been work-hardened form a 3-dimensional network called the Frank network (Fig. 3.20). The mean distance between dislocations is $l = \rho_D{}^{-1/2}$.

The mutual interactions of the dislocations result in their arranging themselves so that they occupy a position of minimum energy on the potential surface. The interaction between dislocations was studied in Sect. 3.3.8.1. Moving a dislocation in its slip plane involves making it cross the hills created by its interactions

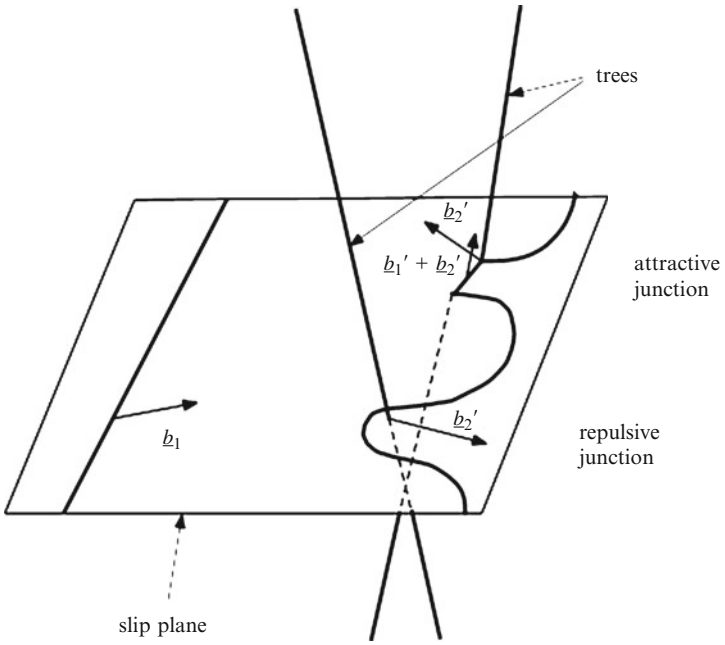


Fig. 3.59 Crossing of the forest of dislocations by a dislocation of Burgers vector $\underline{b}_1$ forming attractive and repulsive junctions

with the other dislocations, among which we distinguish those which are in slip planes parallel to the slip plane of the first from those which cross this plane: we call these latter *trees*, and they form the *forest*. It is clear that the critical slip stress will be greater, the greater the density of dislocations; and that the work-hardening will result from the multiplication of dislocations during plastic deformation.

(a) *Interaction with dislocations parallel to the slip plane*

It is easy for screw dislocations to change their slip planes and thus to choose a path between the dislocations of the Frank network. From Eq. 3.64 the shear stress opposing this is of order $(\mu b/2\pi)\sqrt{\rho_D}$. For edge dislocations Eq. 3.65 gives the minimum as $(\mu b/8\pi)\sqrt{\rho_D}/(1-\nu)$.

Edge dislocations become blocked by forming dipoles; it is the screw dislocations that move.

(b) *Interaction with the forest.*

Let $\underline{b}_1$ be the Burgers vector of the dislocation that slips and $\underline{b}_2$ that of one of the trees. The junction of the two is said to be attractive if the scalar product $\underline{b}_1.\underline{b}_2$ is negative, repulsive if it is positive (Fig. 3.59). At an attractive junction the two

dislocations can combine to form a dislocation with Burgers vector $\underline{b}_1 + \underline{b}_2$ and energy $(\mu/2)(\underline{b}_1 + \underline{b}_2)^2$. This reduces the total energy, and the combination is stable. We can imagine the attractive junctions as forming the anchor points required for the Frank-Read multiplication process (Sect. 3.3.7.6). Their average distance apart will be $2l$, since on average one in every two trees will be attractive, and from Eq. 3.50, for the activation stress for the source, the crossing force will be $0.5\,\mu b\sqrt{\rho_D}$. Actually, attractive junctions are broken by a weaker force, which for FCC structures has been estimated as $(\mu b/4)\sqrt{\rho_D}$.

The crossing of a repulsive junction, in contrast, requires only the creation of a "jog", which represents a much weaker force (refer to Sect. 3.3.8.1 and to Fig. 3.47). Taking the distance over which the dislocation interacts with such a repulsive tree as $3b$, the area swept out, in other word the *activation area*, is of the order of $3lb$ and the work done by the applied stress is $\tau\ b.3lb$. This must equal the energy of the jog, the *activation energy*, which is $0.1\mu b^3$, so $\tau = (\mu b/30)\sqrt{\rho_D}$.

We can summarise the results by saying that the stress required to move a dislocation in the middle of the forest is:

$$\tau_c = \alpha\mu b\sqrt{\rho_D} \tag{3.84}$$

with α between 1/3 and 1/4. This critical stress increases as the density of dislocations increases in the course of plastic deformation: it is this that explains the associated *work-hardening* (Figs. 3.60 and 3.61).

(c) *Microdeformation; depressed modulus of elasticity*

So long as the stress does not exceed the Peierls-Nabarro force the dislocations do not move and the material remains perfectly elastic; but as soon as they start to move they form arcs between anchor points and sweep out an area A, generating a further deformation. However, if the stress remains below the instability threshold $\mu b/l$ for these arcs then the dislocations will return to their original positions when the load is removed: the crystal is still elastic. The result is an apparent decrease in elastic modulus. Small hysteresis loops appear on the stress/strain curve. This dissipative phenomenon is marked by an increase in the *internal friction* when the amplitude of the deformation is increased. The onset of instability of the arcs coincides with the micro yield strength; this is difficult to measure because it corresponds to very small deformations (see Sect. 5.2.2.6).

(d) *The tensile stress-strain curve for FCC single crystals*

The tensile stress-strain curve of FCC single crystals exhibits three main stages (Fig. 3.62) when their orientation is right. However, this is a simplifying picture as more stages than three can be distinguished.

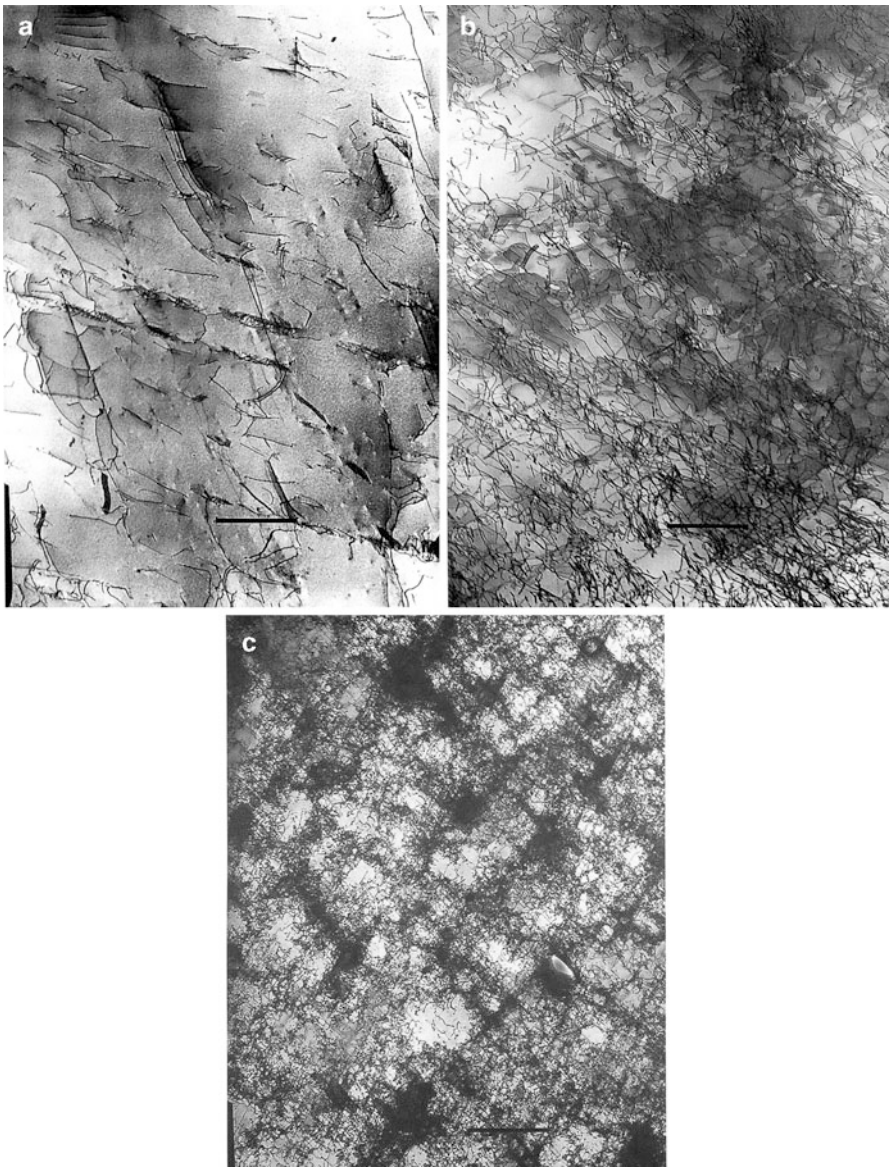


Fig. 3.60 Multiplication of dislocations in the course of deformation resulting in work-hardening. (**a**) Annealed Hastelloy (the bar corresponds to 1 μm), (**b**) the same alloy work-hardened, (**c**) after 15% deformation the dislocations begin to cluster

Stage I, often called the *easy-slip* stage, is characterised by a low *work-hardening modulus*, the slope of the stress-strain curve; it corresponds to the activation of only a single slip system. This stage occurs only with particular orientations of a FCC

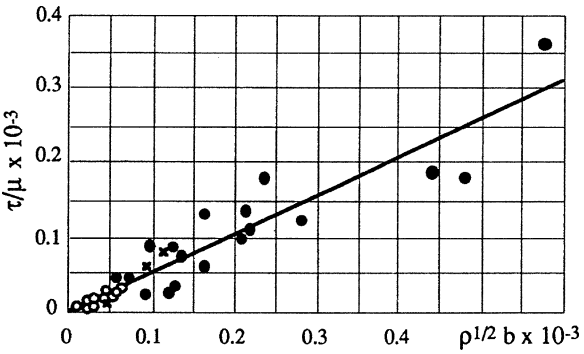


Fig. 3.61 Work-hardening in FCC metals as a function of dislocation density

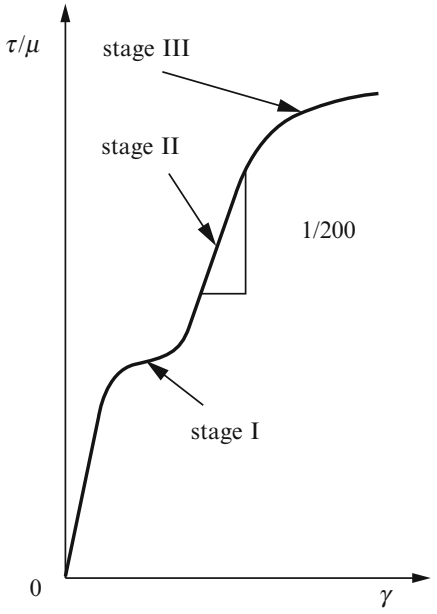


Fig. 3.62 Typical stress-strain curve of a FCC single crystal exhibiting three deformation stages

crystal but predominates in the case of hexagonal single crystals. The slip lines are long and straight.

Stage II, corresponding to the activation of secondary systems, is characterised by a constant work-hardening modulus of the order of $\mu/200$, much greater than that in Stage I. The secondary dislocations interact with the primaries to form *Lomer-Cottrell locks* (refer to Sect. 3.3.4.2) and the hardening increases greatly. The slip lines are straight but are shorter than in Stage I.

Stage III, corresponding to increasingly frequent cross slips, is characterised by a falling work-hardening modulus. Sometimes at still higher strains a fourth stage is distinguished at least (see *e.g.* Ryen and Laukli 2006; Rollett and Kocks 1993).

However, these explanations of the form of the curve are based on a view of the movement of the dislocations that has proved to be too simple: observations have

shown that the structures are much more complex. The dislocations quickly form first *tangles* and then *cells*, the latter of much smaller dimensions than the slip lines. The mean distance d between slip lines and their mean length L can be measured on the surface of a crystal that has been subjected to plastic deformation. The mean size λ of the cells can be found by examining thin sections with the transmission electron microscope. The average length L of slip lines is of the order of the dimension of the grain size: 10–100 μm; that of their mean distance λ is of the order of 1 μm; that of the dislocation cells is of the order of 0.1 μm.

As strain-hardening increases the level of stress, cross-slipping becomes easier and we need to study this phenomenon, which involves the stacking faults.

The tangles and cells are easily destabilised by a change in the stress tensor; as a result the cells are destroyed, but they are re-formed elsewhere. The small distance between dislocations within the cell walls makes the annihilation of dislocations easy; this phenomenon is responsible for *dynamic recovery*, which counteracts the work-hardening.

(e) *Model of work-hardening in relation with slip lines*

We now show how the rate of work-hardening $h = \mathrm{d}\tau/\mathrm{d}\gamma$ ($\simeq \mu/200$) can be determined and the available data on the slip lines can be related.

The flow stress τ_c is related to the dislocation density ρ_D by the equation

$$\tau_c = \alpha \mu b \sqrt{\rho_D} \tag{3.85}$$

and therefore

$$\mathrm{d}\tau_c/\mathrm{d}\rho_D = (1/2)\,\alpha \mu b/\sqrt{\rho_D} \tag{3.86}$$

The increase $\mathrm{d}\gamma$ of the plastic deformation derives from the movement of the dislocations; so we have to calculate $\mathrm{d}\rho_D/\mathrm{d}\gamma$.

Suppose a source of dislocations has generated n loops that have accumulated at the end of a slip line. These will have given rise to a plastic deformation

$$\gamma = nL\Lambda b/V \tag{3.87}$$

where Λ is the length of the dislocations and V is the volume occupied by a slip line, that is $L\Lambda d$. The rate of increase in plastic deformation is then

$$\mathrm{d}\gamma = (b/d)\ \mathrm{d}n \tag{3.88}$$

Since the dislocation density is $\rho_D = n/Ld$ we have:

$$\mathrm{d}\gamma = Lb\mathrm{d}\rho_D \tag{3.89}$$

Substituting these relations 3.86 and 3.89 into $h = \mathrm{d}\tau/\mathrm{d}\gamma$, we get

$$\frac{h}{\mu} = \frac{\alpha}{2}\sqrt{\frac{d}{nL}} \tag{3.90}$$

On the other hand, the spacing d between the slip lines is related to the possibility of dislocations to move on parallel slip planes; so from Eq. 3.64

$$d = \frac{\mu b}{4\pi\tau} \tag{3.91}$$

From this relation and using $\rho_D = n/Ld$ in Eq. 3.85, we get:

$$n = \frac{1}{(4\pi\alpha)^2}\frac{L}{d} \tag{3.92}$$

Hence Eq. 3.90 can be written:

$$\frac{h}{\mu} = 4\pi\alpha^2\frac{d}{L} \simeq \frac{1}{200} \tag{3.93}$$

This shows that it is expected that the length L of the slip lines would be about 600 times their spacing d and that the number of dislocations n would be about 15, so that the height of the slip lines nb would be of the order of 40 nm. These are the right orders of magnitude (refer to Fig. 3.7).

(f) *Initiation of dislocations cells*

If the distance between adjacent attractive trees were constant and equal to $l = 2/\sqrt{\rho_D}$ the flow strength would be given by Eq. 3.85. But in fact this distance is distributed statistically. In the forest of dislocations there are hard zones that are thicker. The flow stress is the stress that has to be applied to the dislocation to make it cross the more sparsely planted areas. The dislocation, which is slipping leaves a loop around each hard zone. The passage of a succession of dislocations across the forest builds up an accumulation of loops at the hard points, which form the first stages of the dislocation *cells*, stabilised by the arrival of further trees from secondary slip planes, creating Lomer-Cottrell locks. Figure 3.63 shows a thin slice of deformed material in a numerically simulated distribution of dislocations. Bundles of dislocations can be observed.

The slip of the dislocations continues until the passages between the hard zones become choked with accumulated loops. This occurs when the number n of loops is of the order of λ/l, and therefore about 15 as estimated in the previous paragraph.

(g) *Cross-slip*

We have seen (Sect. 3.3.4.1) that the dissociation of dislocations in FCC materials introduces a small ribbon of stacking faults, along which the structure is close-

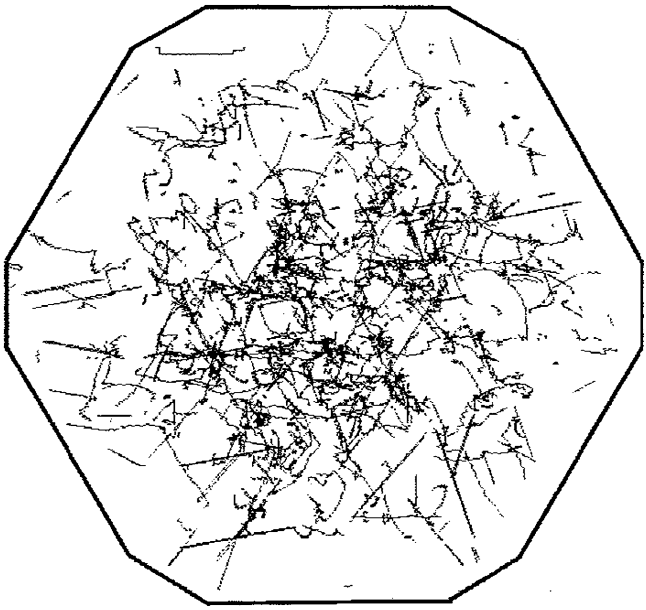


Fig. 3.63 Thin slice of deformed material in a numerically simulated distribution of dislocations in FCC metal (Verdier et al. 1998)

Table 3.6 Stacking-fault energy and separation of Shockley partials at room temperature

	γ_f	a_0	b	μ	d
Metal	mJ/m^2	nm	nm	GPa	nm
Aluminium	166	0.410	0.286	26.1	1.0
Copper	45	0.367	0.255	48.3	5.5
Gold	32	0.408	0.288	27.0	5.6
Nickel	125	0.352	0.249	76.0	3.0
Silver	16	0.409	0.289	30.3	12.4

packed hexagonal (CPH). The energy γ_f of the stacking faults is related to the difference in Gibbs free energy between the FCC and CPH phases. If a dislocation (1/2)[110] is broken down so as to give $(1/6)[211] + (1/6)\left[12\bar{1}\right]$ (*cf.* Sect. 3.4.2.4) the two imperfect dislocations will repel one another since $\underline{b}_1.\underline{b}_2$ is positive, with a force

$$f_{12} = \frac{\mu b_1 b_2}{2\pi w}\left(\cos\theta_1 \cos\theta_2 + \frac{\sin\theta_1 \sin\theta_2}{1-\nu}\right) \tag{3.94}$$

where w is their distance apart and θ_1, θ_2 the angles made by their Burgers vectors with the original dislocation.

This repulsion f_{12} is counterbalanced by an attractive force, since the energy of the stacking fault is proportional to w: at equilibrium $f_{12} = \gamma_f$, the fault energy per unit area (Table 3.6).

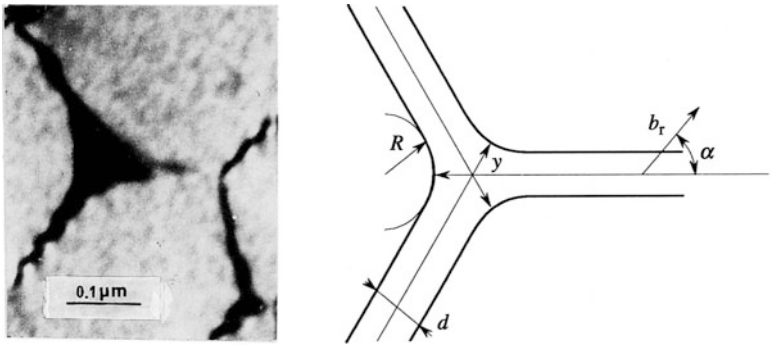


Fig. 3.64 A triple node of dissociated dislocations in the (111) plane, for an austenitic stainless steel showing how the stacking fault energy can be determined. Similar dissociations can be seen in Fig. 3.48 (Lecroisey 1971; Lecroisey and Thomas 1970)

The dislocations are more highly dissociated, the lower the stacking fault energy.

Fault energies are measured by observing particular configurations of dislocations, such as the one just described. The method most commonly used is the direct observation of triple nodes satisfying $2{\cdot}10^{-4} < \gamma_f/\mu b < 5{\cdot}10^{-3}$. Figure 3.64 is an electron micrograph of such a node; this corresponds to the reaction of three perfect dislocations ($\underline{b}_T$): (a/2)$[1\bar{1}0]$, (a/2)$[0\bar{1}1]$, (a/2)$[10\bar{1}]$, in the same plane (111), each dissociated into two partials with Burgers vectors $\underline{b}_p$. Theoretical treatment developed by Brown and Tholen (1964) enables γ_f to be deduced from measurements of the inner and outer radii y and R.

It has been shown that in a number of FCC materials (*e.g.* austenitic stainless steels, Fe-Mn austenitic alloys, Co-Ni alloys) the stacking fault energy is strongly dependent on temperature (see *e.g.* Lecroisey and Thomas 1970). This effect largely contributes to the temperature dependence of the work-hardening rate observed in these materials, when they are deformed at various temperatures close to room temperature (Brown and Tholen 1964; Brown 1964).

Measuring the size of tetrahedra of stacking faults is another way to determine the stacking fault energy.

The dissociated dislocations can slip in only one plane, that defined by the Burgers vectors $\underline{b}_1$ and $\underline{b}_2$ (Fig. 3.65). When meeting an obstacle, they can avoid it by recombining for a certain distance, which requires energy, and then, when it is a screw, by cross-slipping and dissociating again in the secondary slip plane. The energy required is the larger, the greater the dissociation. Dislocations in metals with a low stacking fault energy do not cross-slip easily. Thus metals with low stacking fault energy have very straight slip lines. This process can take place on two different (111) planes in FCC metals, from the basal plane to a prism plane in HCP and on two different (110) plans in BCC.

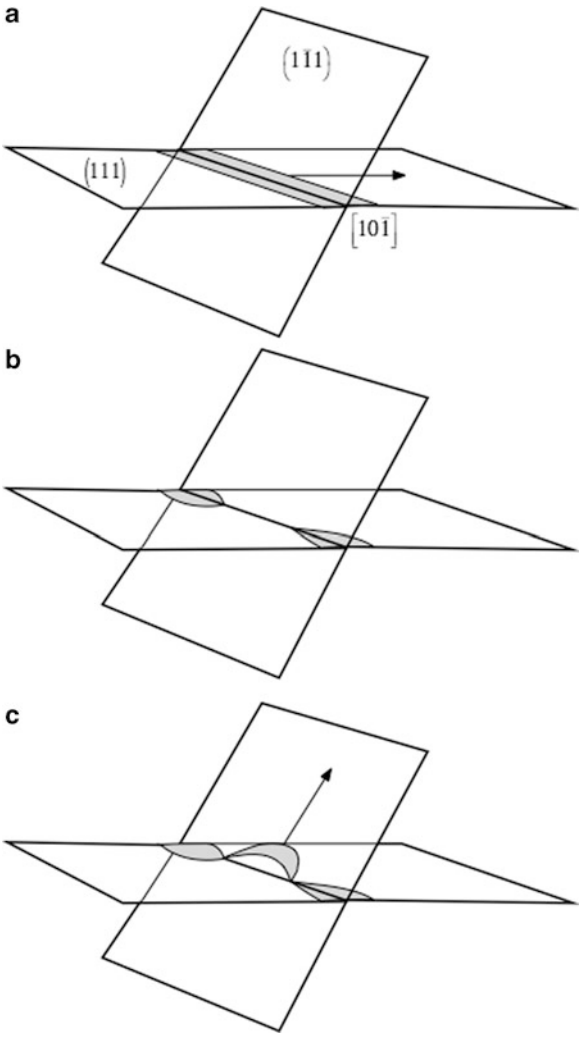


Fig. 3.65 Cross-slip of a dissociated dislocation (**a**) Dissociated dislocation gliding on (111) plane, (**b**) recombination along the intersection $[10\bar{1}]$ with $(1\bar{1}1)$ plane, (**c**) cross-slip in $(1\bar{1}1)$ plane

The calculation of the activation energy is complicated because it results from a minimisation of the constriction energy, the increase of the line energy when it bows in the secondary slip plane and the work of the shear stress (Friedel 1964).

(h) *Dynamic recovery*

The multiplication of dislocations represented by the relation 3.89: $\mathrm{d}\rho_{\mathrm{D}} = (1/bL)\mathrm{d}\gamma$ is in fact compensated for by the *dynamic recovery* process. This can be viewed as the annihilation of pairs of dislocations of opposite signs when they

approach one another on parallel slip planes closely enough, within a distance y say. For a dislocation which moves a distance L there will then be an area $a = 2yL = (2y/b)(\mathrm{d}\gamma/\mathrm{d}\rho_\mathrm{D})$ within which annihilation is possible, and within which it will meet $(1/2)a\rho_\mathrm{D} = (y/b)\rho_\mathrm{D}(\mathrm{d}\gamma/\mathrm{d}\rho_\mathrm{D})$ dislocations of opposite sign. The reduction in density, expressing the dynamic recovery, is therefore obtained by setting this number equal to 1:

$$\mathrm{d}\rho_\mathrm{D} = -\frac{y}{b}\rho_\mathrm{D}\mathrm{d}\gamma \tag{3.95}$$

It follows that Eqs. 3.89 and 3.85 for the work-hardening must be modified to

$$\mathrm{d}\rho_\mathrm{D} = \left(\frac{1}{bL} - \frac{\rho_\mathrm{D} y}{b}\right)\mathrm{d}\gamma \tag{3.96}$$

$$\frac{h}{\mu} = \alpha\left(\frac{\alpha\mu b}{L\tau_\mathrm{c}} - \frac{y\tau_\mathrm{c}}{\alpha\mu b}\right) \tag{3.97}$$

These results indicate that the work-hardening will be linear if both the mean free path of the dislocations and the annihilation distance y vary like $1/\tau_\mathrm{c}$, which seems reasonable.

(i) *Composite model of a material containing cells of dislocations.*

Plastic deformation leads to a very heterogeneous distribution of dislocations, with cells in the walls of which the density is very high. These cells become very well defined after cyclic deformation, when the crystal can be considered as consisting of two phases, one hard and the other soft, working in parallel under the effect of the stress. The slip planes are in fact continuous between the phases, and the shear strain γ varies only slightly. The mean stress can be written

$$\tau = f_\mathrm{wl}\tau_\mathrm{wl} + f_\mathrm{cl}\tau_\mathrm{cl} \tag{3.98}$$

where the subscripts wl and cl denote the cell wall and interior, respectively, and the f's are the respective volume fractions ($f_\mathrm{wl} + f_\mathrm{cl} = 1$). The diagram of Fig. 3.66 shows the macroscopic behaviour of this "composite" as a function of the separate behaviours of the walls and the cells.

Let the dislocation densities in the walls and in the cells be respectively $\rho_{\mathrm{D(wl)}}$, $\rho_{\mathrm{D(cl)}}$; the mean dislocation density is

$$\bar{\rho}_\mathrm{D} = f_\mathrm{wl}\rho_{\mathrm{D(wl)}} + f_\mathrm{cl}\rho_{\mathrm{D(cl)}} \tag{3.99}$$

If the dislocations were distributed homogeneously the flow stress would be

$$\bar{\tau}_\mathrm{c} = \alpha\mu b\sqrt{\bar{\rho}_\mathrm{D}} \tag{3.100}$$

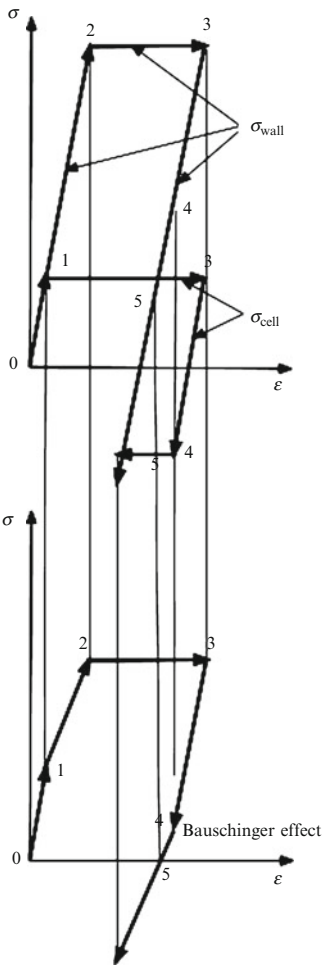


Fig. 3.66 Mugrabi's composite model for a material containing cells of dislocations. The upper diagram refers to the cell walls and interiors separately, the lower to the combination (Mughrabi 1988; Gil Sevillano 1993)

but in fact:

$$\tau_{\mathrm{D(wl)}} = \alpha\mu b\sqrt{\rho_{\mathrm{D(wl)}}}$$
$$\tau_{\mathrm{D}(c\mathrm{l})} = \alpha\mu b\sqrt{\rho_{\mathrm{D(cl)}}} \tag{3.101}$$

so that we have approximately:

$$\tau_c = \alpha\mu b\left(f_{\mathrm{wl}}\sqrt{\rho_{\mathrm{D(wl)}}} + f_{\mathrm{cl}}\sqrt{\rho_{\mathrm{D(cl)}}}\right) \tag{3.102}$$

and

$$\bar{\tau}_{\mathrm{c}}^2 - {\tau_{\mathrm{c}}}^2 = (\alpha\mu b)^2 f_{\mathrm{wl}}\, f_{\mathrm{cl}}\left(\sqrt{\rho_{\mathrm{D(wl)}}} - \sqrt{\rho_{\mathrm{D(cl)}}}\right)^2 \tag{3.103}$$

showing that the heterogeneity lowers the flow stress.

Figure 3.66 shows several stages in the loading-unloading curve. Up to point 1 the behaviour of both cell walls and cell interiors remain elastic. At point 1 and point 2 respectively the cell interior and the cell walls begin to deform plastically. At point 3 elastic unloading begins, the cell interiors go to compression, and at point 4 reach the yield strength in compression, of equal magnitude to the one in tension, and deform plastically: plastic yielding of the "composite" begins even before complete unloading. This is the *Bauschinger effect*.[14] The cell walls are still under tension at this point. Residual stresses will remain after the load has been removed (point 5). This can be expressed otherwise by saying that the work-hardening is *kinematic*. Thus, this is linked to the heterogeneity, whilst *isotropic work-hardening,* proportional to $\sqrt{\rho_D}$, is produced by a homogeneous multiplication of the number of dislocations (Exercise in Volume III).

(j) *Recovery and recrystallisation*

In spite of dynamic recovery, dislocations keep accumulating in the course of deformation. Work-hardening is a treatment that can often be used with advantage, for example for hardening copper tubes, drawn wires for tires, pre-stressed steel bars. However, too strong a work-hardening can in particular impede operations of shaping by plastic deformation; dislocations must then be partly removed with the help of temperature by *annealing*.

The first effect is to accelerate the removal of pairs with opposite signs that are close to each other, such as those that form dipoles. This gives a small reduction in the dislocation density which is responsible for the *recovery* and so, combined with the removal of the point defects formed during the work-hardening, in particular by the dragging of the jogs, brings about the first stage in reducing the hardness.

In the next stage the dislocations arrange themselves so as to minimise their interaction energy. The dislocation climb, made possible by the migration of vacancies to the core, causes them to become equidistant from one another, so forming walls of dislocations, as shown in Figs. 3.50 and 3.51 (refer to Sect. 3.3.8.3). These walls introduce a misorientation and separate the grains into *subgrains*: this is *polygonisation*. In Sect. 3.3.8.3 we determined the energy of such dislocations walls or *sub-boundaries* (Fig. 3.52). The calculation becomes invalid when the dislocations are too close from one another and the misorientation increases. With such strong misorientations we should not be speaking of sub-boundaries, but of boundaries proper. Their structure can be described by saying that where two crystals of different orientations join there are very many atomic sites in quasi-coincidence; for certain orientations, twins in particular, all are. These correspond to minima on the curve of boundary energy against misorientation angle.

Finally, in the last stage of the annealing one of the subgrains that happens to be more perfect may grow at the expense of the others, and form a new, recrystallised grain; and after *recrystallisation* there will be an entirely new structure of grains in which the dislocation density has returned to a low value (Fig. 3.67). The size of the recrystallised grains is an increasing function of the annealing temperature and a

[14] Johann Bauschinger (1834–1893) was a German engineer.

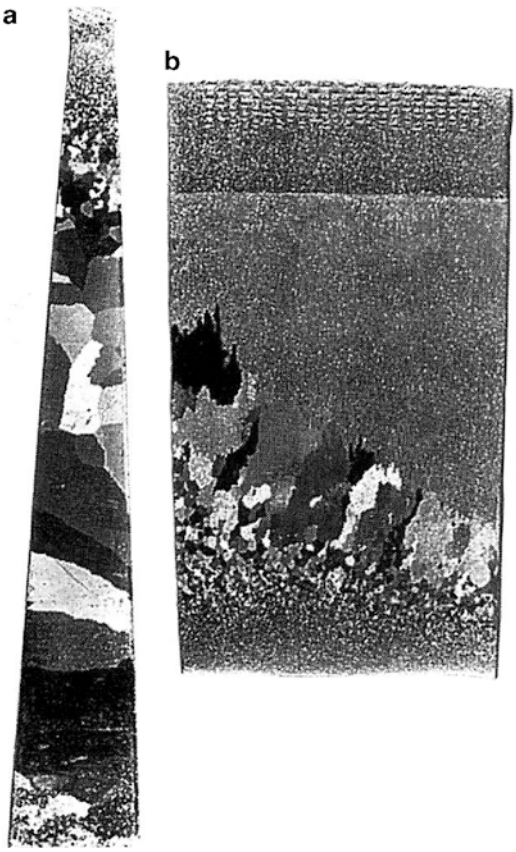


Fig. 3.67 Macrographs of annealed aluminium test pieces. (**a**) The trapezoidal shaped tensile elongated test piece contained a strain gradient. (**b**) Annealing under a temperature gradient

decreasing function of the original strain (Figs. 3.67, 3.68 and 3.69); the latter must be above a threshold at which recrystallisation can occur. Annealing in the critical conditions can be profitable for the fabrication of single crystals. As Fig. 3.69 shows, recrystallisation reduces the hardness considerably.

It can be noted that there is a correlation between the recrystallisation temperature and the melting point (Fig. 3.70).

Recrystallisation can take place while an alloy is being deformed at high temperature. This is called *dynamic recrystallisation*. Two types are observed: *continuous* and *discontinuous recrystallisation*. The first one results from the migration of grain boundaries in contrast to the second one, which occurs by the nucleation and growth of new grains. This difference is reflected in the evolution of the stress in a torsion test. In continuous recrystallisation, the stress first rises to a peak and then keeps decreasing. In discontinuous recrystallisation, the stress oscillates.

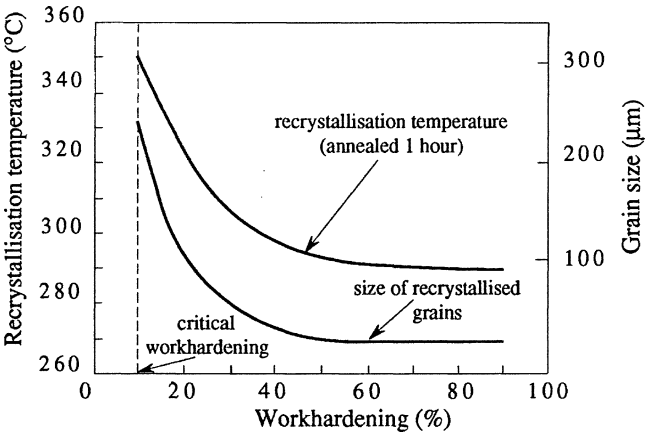


Fig. 3.68 Recrystallisation temperature of aluminium and size of the recrystallised grains

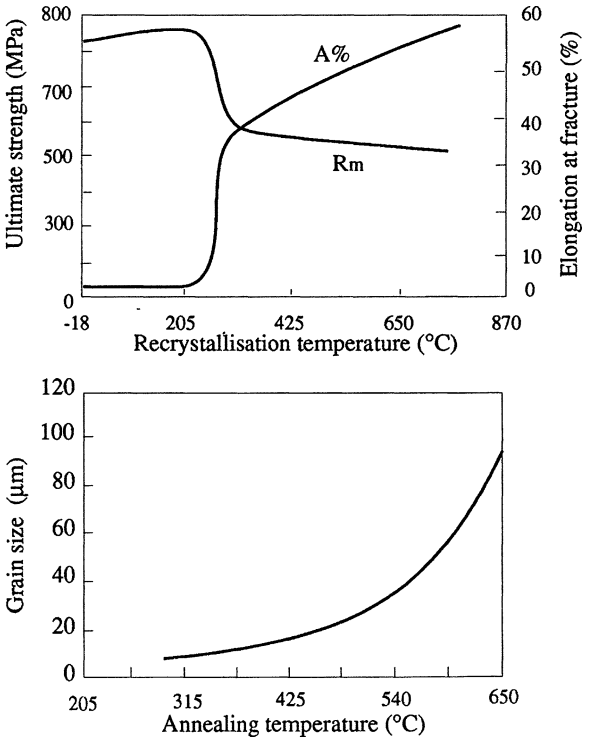


Fig. 3.69 Change of mechanical properties of brass and recrystallised grain size as a function of the recrystallisation temperature

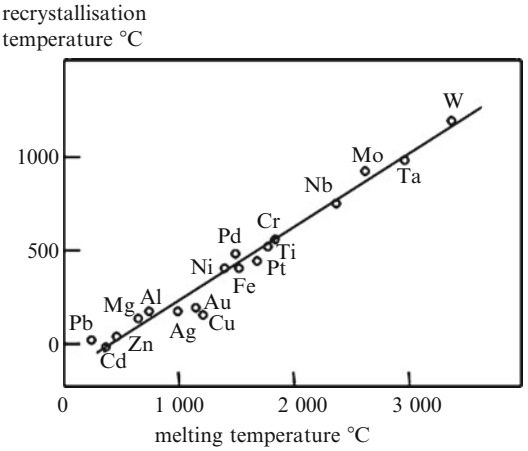


Fig. 3.70 Correlation between the recrystallistion temperature and the melting point

3.4.2.3 Interaction of Dislocations with Grain Boundaries

(a) *The Hall-Petch law*

Materials usually consist of grains with different crystallographic orientations, separated by grain boundaries. When a polycrystal is stressed the displacement of the dislocations starts in those crystals that are most favourably oriented with respect to the shear stress. Sources of dislocation are activated when this reaches a critical value – given by Eq. 3.51 – and then emit loops, which pile up against the grain boundaries. This continues until the reaction exerted on the source stops it from working. The accumulation of dislocations on the grain boundary induces stress concentration on the adjoining grain and, when it reaches a critical condition, this can activate sources there (Fig. 3.71). From there on plastic deformation extends in all the grains. The yield strength of the polycrystal is then the critical stress needed for this mechanism to operate. Of course, the number of dislocations, which can accumulate in the first grain to deform, and consequently the stress concentrations in the adjoining grains, are the smaller the smaller the grain size. We then expect that the yield strength will be a decreasing function of the grain size d. Indeed, it was found experimentally by Hall and Petch[15] (Hall 1951; Petch 1953) that the yield strength R_p followed the law:

$$R_p = \sigma_i + k_y d^{-1/2} \tag{3.104}$$

[15]Eric Ogivie Hall (1925–2009) was a New-Zelander born Australian metallurgist. Norman James Petch (1917–1992) was a British metallurgist. The Hall-Petch equation is named after their independent contributions in the early 1950s to the study of the influence of grain size on the yield strength of polycrystals.

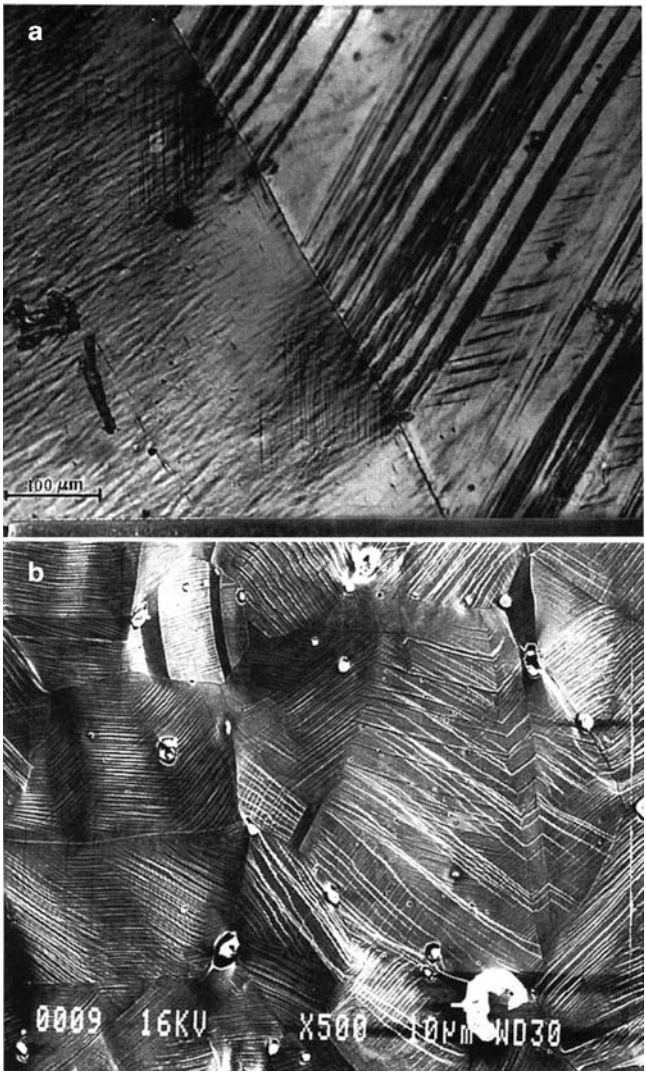
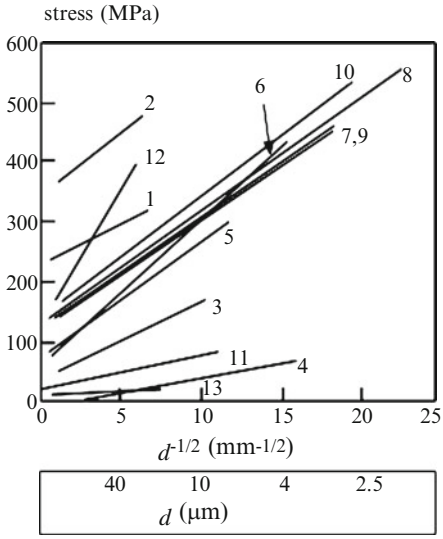


Fig. 3.71 (**a**) Slip bands in the copper grain on the *right*, causing stress concentrations in the grain on the *left*. These are relieved by secondary slips in the latter. (**b**) Slip lines in a copper polycrystal work-hardened to 20%

where σ_i is the internal stress needed to move the dislocations in the first place, and k_y is the Petch factor. This is known as the Hall-Petch law (Fig. 3.72) (Table. 3.7).

The internal stress σ_i is strongly dependent on temperature, but k_y is not.

Finally, note that the sub-boundaries and the boundaries between twins can constitute obstacles analogous to the grain boundaries and can lead to relations of the Hall-Petch type in which grain size is replaced by the mean distance between these obstacles. For further study see Gil Sevillano (1993).



1 304 steel annealed (R_{p1})
2 304 steel strain hardened (R_{p1})
3 brass (R_p)
4 cupper (R_p)
5 ferrovac (R_p)
6 0.05 C steel (R_p)
7 0.09 C steel (R_p)
8 0.13 C steel (R_p)
9 0.15 C steel (R_p)
10 0.20 C steel (R_p)
11 nickel (R_p)
12 molybdenum (R_p)
13 aluminium (R_p)

Fig. 3.72 Hall-Petch relation

Table 3.7 Values of the Petch factor k_y

Alloy	σ_ι (MPa)	k_y (MPa$\sqrt{\text{m}}$)
BCC		
Mild steel R_e	71	0.74
R_p	294	0.39
Molybdenum R_e	108	1.77
R_p	392	0.53
Niobium R_e	69	0.04
FCC		
Copper $R_{p0.5}$	26	0.11
Aluminium $R_{p0.5}$	16	0.07
Aluminium 3.5% Mg R_e	49	0.26
Silver R_e	37	0.07
Brass 70/30 R_e	45	0.31
R_{p20}	337	0.34
HCP		
Zinc $R_{p0.5}$	32	0.22
$R_{p17.5}$	72	0.36
Magnesium $R_{p0.2}$	6.9	0.03
Titanium R_e	78	0.40
Zirconium $R_{p0.2}$	29	0.25

Recent studies on nanocrystalline materials (grain size of the order of 10–100 nm) have shown that the Hall-Petch law (3.104) does not work for these grain sizes. The yield strength decreases with the grain size, which is the opposite of the effect predicted from Hall-Petch equation (see *e.g.* Sanders et al. 1997).

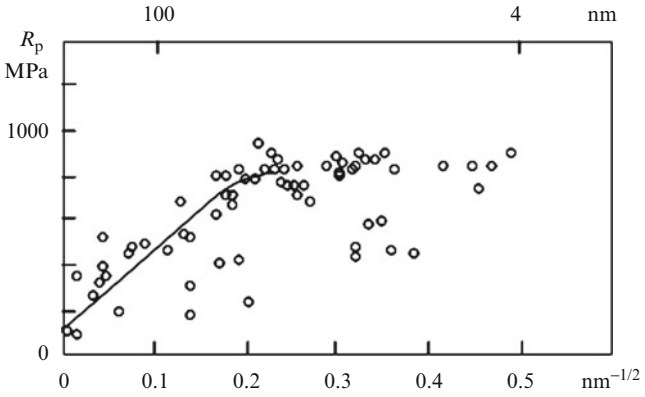


Fig. 3.73 Yield strength of copper as a function of the grain size (From Meyers et al. 2006)

Figure 3.73 shows that the hardness in nanocrystalline Cu no longer increases with decreasing grain size below about 15 nm. While this may indeed be the case, it is possible that changes in synthesis conditions used to prepare the smaller grain sizes are counteracting the strengthening from further grain refinement. It is therefore difficult to separate inherent mechanical properties from processing. The behaviour of nanocrystalline materials is discussed later.

We must now try to calculate the critical stress to see if the Hall-Petch law can be justified and several models have been used.

(b) *Pile-ups model*

The first model, which we study is that of simple pile-ups of dislocations, as described in Sect. 3.3.8.2, is at the origin of the stress concentration in the adjoining grain (Fig. 3.49). Such pile-ups can be observed in materials of low stacking fault energy, so that the slips are very planar, such as in austenitic stainless steels (Fig. 3.48). We demonstrated that at the head of a pile-up the stress reaches $n\tau b$, and that the number n of dislocations, which can be accumulated before the source ceases to be activated, is given by formula (3.69). We slightly modify these expressions by replacing the shear stress τ by the effective stress τ_{eff}. It is equal to $\tau - \tau_i$, τ_i being the *"friction" stress*, the stress needed to move dislocations in the first place. Now equating $n\tau_{\text{eff}}b$ to the stress needed to activate a source in the adjoining grain τ_s, replacing n by the value given by Eq. 3.69 with the size of the pile-up L taken as half the grain size d, we obtain:

$$\frac{d}{4Mb}(\tau_Y - \tau_i)^2 = \tau_s \tag{3.105}$$

where τ_Y is the yield strength of a grain in the polycrystal. This can justify the Hall-Petch law.

(c) *Model of the mean free path of dislocations (Johnson model)*

Despite the above analysis, several observations seem to indicate that very high stresses are not necessary for the activation of sources in the grain boundaries. Now, slip bands consisting of several slip lines can produce stronger stress concentrations on the grain boundary. But, we can also confirm the Hall-Petch law from the relation (3.85)

$$\tau_{\mathrm{y}} - \tau_{\mathrm{i}} = \alpha\mu b\sqrt{\rho_{\mathrm{D}}} \tag{3.106}$$

by assuming the mean distance travelled by the dislocations to be proportional to the grain size, say $l_{\mathrm{D}} = \beta d$; then with $\gamma = \rho_{\mathrm{D}} b l_{\mathrm{D}} = \rho_{\mathrm{D}}\beta b d$ we have

$$\tau_{\mathrm{Y}} - \tau_{\mathrm{i}} = \alpha\mu\sqrt{\frac{b\gamma}{\beta}}d^{-1/2} \tag{3.107}$$

This model, due to Johnson, gives a Petch factor that varies like the square root of the deformation. It has been improved by Ashby (1970), who took into account the density ρ_{G} of geometrically necessary dislocations: with $\rho_{\mathrm{G}} = 4\gamma/bd$ (from Eq. 3.17) we have

$$\tau_Y - \tau_i = \alpha\mu\left[b\gamma\left(\frac{1}{4} + \frac{1}{\beta}\right)\right]^{1/2} d^{-1/2} \tag{3.108}$$

(d) *From grain to overall behaviour*

Strain incompatibilities between different grains can become very important in plastic deformation. Von Mises has shown that these can be eliminated by introducing *five independent slip systems.* We shall take up this question again in Sect. 3.5.3.6, with Taylor's model. Crystals that do not have five such systems, such as beryllium, which does not slip easily except in the basal and prismatic planes, compensate for this lack by twinning, but they have only poor ductility.

In FCC metals, on the other hand, there are many ways in which slip systems can be combined to produce a given slip; Taylor made a detailed study of this. The work $\sigma\ \mathrm{d}\varepsilon$ done in a plastic deformation $\mathrm{d}\varepsilon$ must equal the sum of that done in all the slip systems, that is:

$$\sigma\mathrm{d}\varepsilon = \tau_{\mathrm{Y}}\sum\nolimits_i \mathrm{d}\gamma_i \tag{3.109}$$

where τ_{Y} is the shear stress, assumed to be the same in the various systems, and the γ_i are the shear strains.

We get $\sigma/\tau_{\mathrm{Y}} = \mathrm{m}_{\mathrm{T}}$, the mean value of the *Taylor factor;* in the absence of texture this is 3.067 for FCC, when all the slip systems combine to minimise the work. m_{T}

is different from m_S, the mean of the inverse of the Schmid factor for single crystals, which for FCC structures is 2.24.

To derive from the curve $\tau = f(\gamma)$ for a single crystal the tensile curve $\sigma = f(\varepsilon)$ for a polycrystal, we can put $\sigma = m_T\tau$ and $\gamma = m_T\varepsilon$. In these conditions the relations derived above for a grain lead to the *Hall-Petch* relation (3.104) between the proof strength R_p and the grain size.

(e) *Nano-structured materials*

The Hall-Petch law displaying increase of the yield strength when the grain size is decreased incites to fabricate materials with as small a grain size as possible (Meyers et al. 2006; Saada and Durras 2009). The extrapolation down to nano-grains would predict yield strengths of the order of 4,000 MPa. Various techniques have been developed to obtain nano-structured materials (Tham et al. 2007). We will not quote them all. One of them is to use *severe plastic deformation*. Very large drawing ratios allow to obtain eutectoid carbon steel wires used in tires with yield strength that can reach 3,000 MPa. Equal channel angular pressing (ECAP) (Valiev 1996) consists in using in several passes an extruder the channel of which is bent to an angle which can be as large as 90° (Fig. 3.74). This can produce 100 nm grain size metals. Electrolytic deposition can produce nano-structured materials of smaller grain size (Erb 1995; El-Sherik and Erb 1995). We omit the techniques, which require sintering of powders, which makes difficult the complete elimination of porosities.

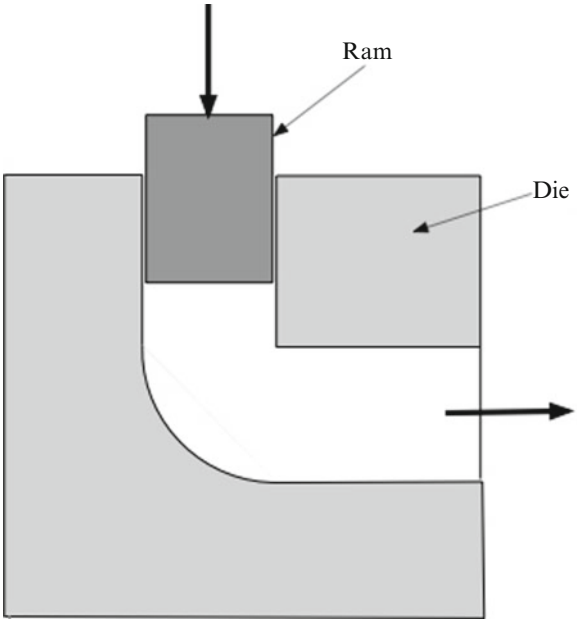


Fig. 3.74 Sketch of the equi-channel angular pressing (ECAP) process for grain refinement

Various considerations prevent extrapolating the Hall-Petch law to nano-grains. In such materials, considering that the thickness of grain boundaries δ is of the order of a few inter-atomic distances, the volume fraction of grain boundaries $3\delta/d$ can reach 60% for a grain size of 5 nm. Such grains are too small to store dislocations. Pile-ups are not possible. Furthermore work-hardening from interaction between dislocations is absent. However, it must be noticed that a dislocation crossing a grain of nanometric size produces a very large strain ε: as n dislocations produce a strain $\varepsilon = nb/d$, if the grain size is of the order of $20b$, one dislocation produces a strain of 5%. Thus, the initial slope of strain-hardening is very large (Saada 2005).

Indeed the yet too scarce experimental results demonstrate that below a grain size of the order of 25 nm the Hall-Petch law is no longer valid. The scatter of data, which might be due to the presence of porosities in some of the experimented alloys, does not allow discriminating between a plateau and a negative slope in the Hall-Petch curves (Fig. 3.73 as an example).

Various models have been proposed for the deformation of nano-grains. Grain boundaries can emit dislocations, and numerical simulations show that this is the essential mechanism for small grain sizes. From Eq. 3.51 the stress to activate a source of a $20b$ grain size in steel would be 5,800 MPa, too high in view of experimental results even considering local stress concentrations. However it is less for partial Shockley dislocations. In crossing the grains they would leave stacking faults and even twins and this was observed in aluminium (Chen et al .2003; Liao et al .2003); it was also found by numerical simulations. Another model (Carlton and Ferreira 2007) treats the absorption of dislocations by grain boundaries, showing that, whereas these are impervious for large grains, making pile-ups possible, the probability of absorption increases as the grain size decreases; this is due to the fewer number of dislocation core atoms, which for absorption must jump in the grain boundary, when the dislocation length is decreased.

Lastly, the predominance of grain boundaries in nano-structured materials eases deformation by *grain boundary sliding* and by *Coble*[16]*creep*, which will be studied in Chap. 4 of this Volume.

It remains that decreasing the grain size to very low values is beneficial for increasing the yield strength (Fig. 3.72). As examples, a yield strength of 900 MPa was reached for a 10 nm grain size of electro-deposited nickel (Robertson et al. 1999). Using 16 passes of ECAP (Equal-Channel Angular Pressing), the yield strength of a 6061 aged aluminium alloys was raised from 90 MPa up to 306 MPa (Tham et al. 2007). We mention also at this point that decreasing the grain size is also beneficial when considering other properties, such as fatigue and brittle fracture resistance. This will be studied in Volume II, devoted to damage mechanisms.

At the end of this section devoted to grain size effect it should be admitted that, if it is well accepted that grain size refinement is the best method to improve the yield strength and decrease the ductile-to-brittle transition temperature in ferritic steels (see Volume II), the physical basis of the scale effect represented by the Hall-Petch equation is not yet firmly established. This is likely due to the poor knowledge of

[16]Robert L. Coble (1928–1992) was an American materials scientist.

the mechanisms describing the crossing of grain boundaries by dislocations. This fundamental problem can only be addressed with special grain boundaries, such as twin boundaries, as shown later, in Sect. 3.4.4.

3.4.3 Interaction of Dislocations with Foreign Atoms

We have been looking at three mechanisms for hardening materials: the Peierls force (which depends essentially on the binding force of the crystal), work-hardening and the effect of grain size. These provide some scope for affecting the hardness of a pure material, and we now show how suitable alloying can give us much greater flexibility in modifying the mechanical properties, particularly by heat treatment. The foreign atoms in a crystal can enter either in solid solution or as precipitates or particles introduced on manufacture.

3.4.3.1 Solid Solution Strengthening

Atoms in solid solution can interact with dislocations in three ways, determined by the effects of size, modulus and chemistry, respectively. This last effect has to do with the stacking fault energy. We will discuss this after studying the solid solution hardening due to size effects.

(a) *Effect of size*

In general, a solute atom will have a different radius $a_{0S} = (a_0 + \Delta a_0)$ from that (a_0) of the solvent atoms and will therefore be surrounded by a dilatation field, which has an energy of interaction with the stress field of a dislocation. For this reason an atom that dilates the network will tend to place itself in the dilatation zone close to the edge dislocations. More precisely, the change in volume ΔV due to an atom with size factor $\eta = \Delta a_0/a_0$ is $\Delta V = 4\pi {a_0}^3 \eta$ and the interaction energy is $\sigma_m \Delta V$, where σ_m is the hydrostatic stress, equal to the mean stress. From Eq. 3.43, giving the stresses around the dislocations, we have

$$\sigma_m = -\frac{1+\nu}{1-\nu}\frac{\mu b}{3\pi r}\sin\theta \tag{3.110}$$

so the interaction energy is:

$$U = -\sigma_m \Delta V = \frac{4}{3}\frac{1+\nu}{1-\nu}\frac{\mu b}{r}(\sin\theta)\,{a_0}^3\eta = \frac{1}{6}\frac{1+\nu}{1-\nu}\frac{\mu b^4}{r}\eta \tag{3.111}$$

for $\theta = \pi/2$ and $a_0 = b/2$.

For copper, with crystallographic parameter $a_0 = 3.615 \times 10^{-10}$ m, $\mu = 48.3 \times 10^9$ Nm^{-2}; taking $r = 2a_0$ and $\theta = \pi/2$, this gives $U = 1.7\eta$ eV.

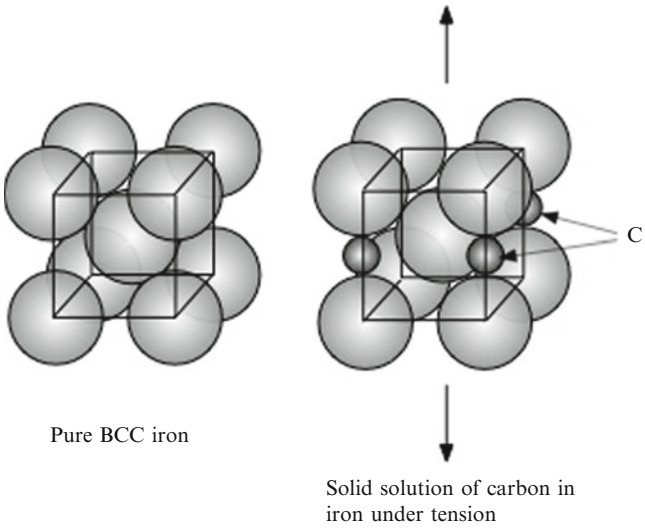


Fig. 3.75 Solid solution of carbon in BCC iron under tension; the lattice is changed to body centred tetragonal (A5)

An atom, which only dilates the lattice does not interact with a screw dislocation. However, certain atoms in solution also distort the lattice, a case of great practical importance being that of *carbon* and *nitrogen* in solid solution in *α-iron* (Fig. 3.75). Being very small – the radius of the carbon atom is 0.77×10^{-10} m, whilst that of BCC iron is 1.241×10^{-10} m – they can be inserted between the iron atoms, forming a solid solution. The carbon atoms can occupy the positions (½,0,0), (0,½,0), (0,0,½); under tensile stress in the direction [001] the (0,0,½) positions become the most favourable and the carbons migrate there. Under compression, on the contrary, they leave this position. If the crystal is cycled through tension and compression at a frequency corresponding to the jumps of these atoms from one position to the others, there is a strong absorption of energy, which can be measured by means of a torsion pendulum. Following the decrement of the oscillations, at 1 Hz is measured the corresponding peak of the *internal friction*, called *Snoek's peak* (Snoek 1941), in the neighbourhood of the ambient temperature (refer to Sect. 5.2.2.3).

Here the distortion brought about by the carbon atoms generates an interaction energy not only with edge dislocations but also with screw dislocations; the level is a few electron volts (see Exercise in Volume III).

(b) *Effect of modulus*

The effect here is due to the difference in elastic modulus between the matrix and the solute. In the absence of any effect of size the deformation energy is proportional to the elastic modulus; it can be shown that it is of second-order relative to the size effect and therefore can often be neglected.

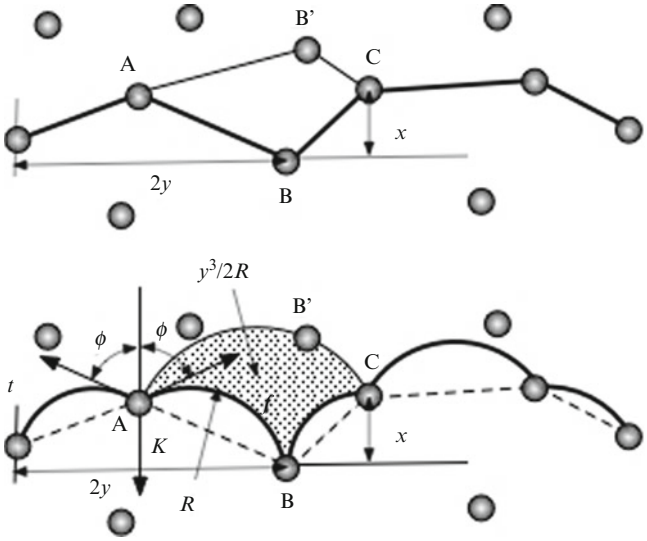


Fig. 3.76 Dislocation interacting with solute atoms close to its slip plane. The top sketch shows the equilibrium position of the dislocation and how it jumps from one atom to the next when the force is high enough. The bottom sketch shows the model of Friedel for the displacement of the dislocation

(c) *Solid solution strengthening*

Here we calculate the stress needed to move the dislocations in the solid solution.

As a result of the interaction with the solute atoms the equilibrium position of a dislocation zigzags between those in its slip plane (Fig. 3.76). The first step is to determine this position. It results from the compromise between the gain in energy by solute atom pinning and the increase of line energy. As shown by Eq. 3.102, the interaction energy decreases as the inverse of the distance of the atom to the dislocation, and we will take into account only those atoms, which are at a distance of the order of the Burgers vector.

If the concentration of solute atoms is c, there are c/b^2 of them, which, being at least at a small enough distance, can interact with the dislocation. Averaging the distance x and y displayed in Fig. 3.72, there are $1/xy$ solute atoms per unit area of slip plane, which interact with the dislocation: $\frac{1}{xy} = \frac{c}{b^2}$. The gain in energy per unit length of the dislocation due to the pinning effect of the atoms is $E_1 = \frac{U}{y} = \frac{Uc}{b^2}x$. On the assumption that x/y is small, the energy of the line increases by

$$E_2 = \frac{\mu b^2}{2}\left[\left(x^2+y^2\right)^{1/2} - y\right] / \; y \simeq \frac{\mu b^2}{4}\left(\frac{x}{y}\right)^2 = \frac{\mu c^2 x^4}{4b^2} \tag{3.112}$$

Minimising $E_2 - E_1$, replacing U by the expression 3.111 with $r = b$, we find the equilibrium zigzag width x_e:

$$x_e{}^3 = \frac{U}{\mu c} = \frac{1}{6}\frac{1+\nu}{1-\nu}\frac{\eta}{c}b^3 \tag{3.113}$$

and

$$y_e{}^3 = 6\frac{1-\nu}{1+\nu}\frac{1}{\eta c^2}b^3 \tag{3.114}$$

and we can check that x_e/y_e is indeed small.

Under the effect of an applied stress the dislocation jumps from atom to atom. It moves from ABC to AB'C, corresponding to energy minima, passing through AC corresponding to a maximum. The work done by the stress must equal this change in energy, which is $2y_e(½E_1 - E_2)$. Friedel (1964) suggested that the area swept in making this jump is on average equal to the area per atom in the slip plane b^2/c. Thus the work is:

$$\tau_c b\frac{b^2}{c} = \frac{E_1 - 2E_2}{y_e} \tag{3.115}$$

and therefore:

$$\tau_c = \frac{Uc}{2b^3} \tag{3.116}$$

which, with the value of U from (3.111), gives:

$$\tau_c = \frac{\mu\eta c}{4} \tag{3.117}$$

In this calculation we have assumed, for simplicity, that the segments of the dislocation between successive solute atoms are straight lines: this is only approximately true, since if t is the line tension the stress induces a curvature $1/R = \tau b/t$, as we showed in Sect. 3.3.7 (Eq. 3.51).

Let K be the maximum force an atom can exert on the dislocation. From Fig. 3.77, the equilibrium conditions give:

$$2t\cos\phi = K \tag{3.118}$$

and we have also

$$\tau bL = K \tag{3.119}$$

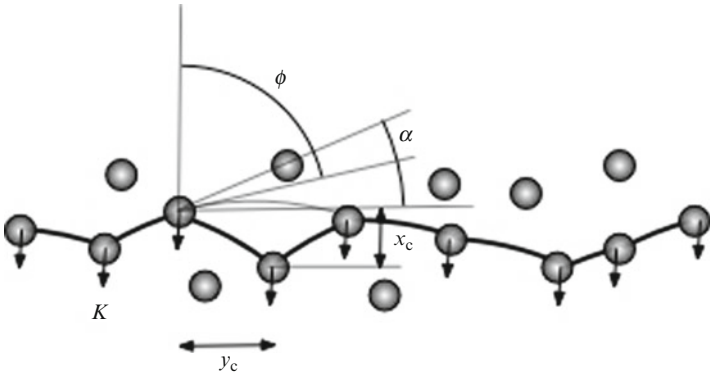


Fig. 3.77 When the atomic concentration is high the angle $\alpha = x_e/y_e$ becomes greater than the complement of the angle ϕ. The dislocation moves by being unpinned from all atoms at the same time

With radius of curvature R, and assuming that ϕ is close to $\pi/2$, which is so if t is much greater than K, the area in question is $L^3/2R$, and Friedel's assumption gives

$$L^3/2R = b^2/c \tag{3.120}$$

Also $L = 2R\cos\phi$. Eliminating ϕ, R and L from the relations and remembering that $t = ½\,\mu b^2$, we find:

$$\tau_c = (K^{3/2}/b^3\sqrt{\mu})\sqrt{c} \tag{3.121}$$

To find the value of the force K we consider an edge dislocation, which slips in a plane distant b from the solute atom. From (3.111), with $a_0 = b/2$ and $\nu = 1/3$, the energy is:

$$U = \frac{\mu b^4}{3}\frac{b}{\xi^2 + b^2}\eta \tag{3.122}$$

where ξ is the projection on the slip plane of the distance between the dislocation and the solute atom. The force exerted on the atom by the dislocation is $K = \partial U/\partial\xi$ and its maximum value, at $\xi = b/\sqrt{3}$, is:

$$K = \frac{\sqrt{3}}{8}\mu b^2\eta \tag{3.123}$$

This is much less that the line tension.

If we replace K in (3.121) by the value we have now found we get, for the critical stress τ_c,

$$\tau_c = 0.1\mu\eta^{3/2}\sqrt{c} \tag{3.124}$$

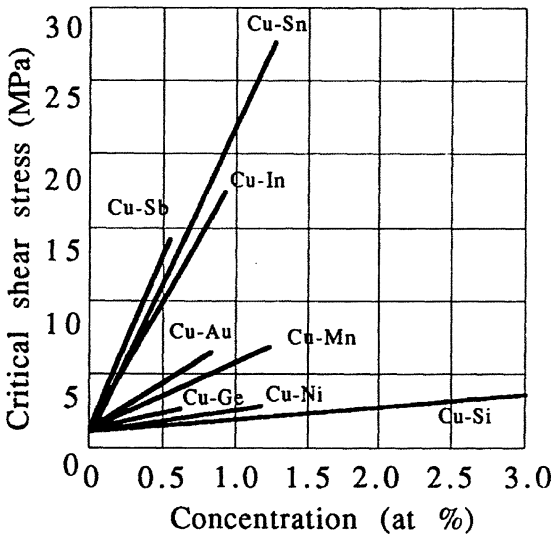


Fig. 3.78 Effect of the concentration of solute atoms on the critical shear stress of copper

However, when the atomic concentration is high, or the pinning of dislocations by solute atoms weak, the average angle of deviation $\alpha = x_e/y_e$ can be greater than the complement $K/2t$ of the angle ϕ, in which case the dislocation cannot jump from one atom to the next, as we just considered, without getting unpinned from the others (Fig. 3.77). The dislocation now must overcome all the obstacles as a whole. The force needed is equal to the sum of the resistance of the $1/y_e$ atoms per unit length to which it is pinned.

The critical concentration is such that $x_e/y_e = K/2t$, that is 0.09η.

When the concentration is larger than this critical value, the critical shear stress is:

$$\tau_c = \frac{1}{y_e}\frac{K}{b} = 0.15\mu\eta^{4/3}c^{2/3} \tag{3.125}$$

These relations, giving a square root dependency of the critical shear stress with respect to the atomic concentration of solute atoms below a critical value and a power 2/3 dependency above, are supported by experimental observations (see Fig. 3.78 for instance); so are the dependencies in power 3/2 and 4/3 respectively with respect to the size factor (Fig. 3.79). But this holds only at very low temperatures; as we shall see later, thermal agitation can destroy hardening completely. For further study see Reppich (1993).

(d) *Hardening by "atmospheres"*

When the interaction energy is high and the solute atoms can diffuse easily, as is the case for carbon and nitrogen in iron, they group themselves around dislocations so as to form what are called *atmospheres* (see Exercises in Volume III). This

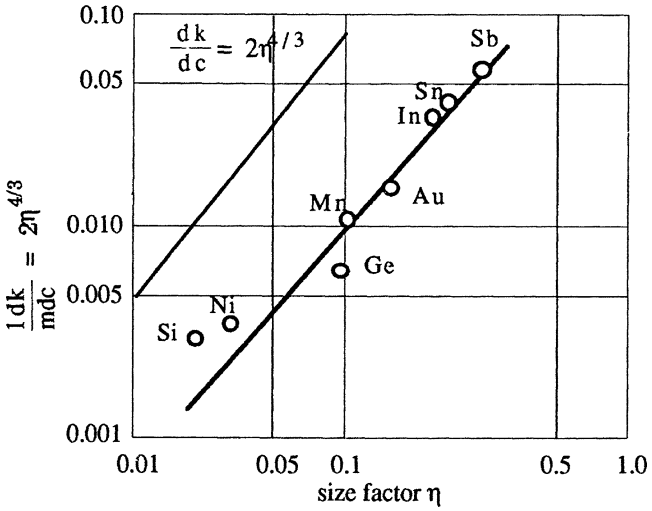


Fig. 3.79 Effect of the size factor on the yield strength of copper

anchors the dislocations very firmly and it is found that a large stress is needed to detach them: it is easier to create new dislocations in stress-concentration zones, for example in the regions where the heads of a test piece join the stem. These dislocations can then slip, to form pile-ups and activate anchored dislocation sources by the Hall-Petch mechanism described in Sect. 3.4.2.3. The process starts an avalanche that runs from grain to grain through the entire polycrystal, forming what is called the *Lüders* or *Piobert-Lüders band*[17] (Fig. 3.80), which can be propagated by a much smaller stress than that needed to start its formation. This is shown by a yield drop in the tensile stress-strain curve (Fig. 3.81), followed by a plateau corresponding to the invasion of the test piece by Piobert-Lüders bands. During this time the deformation is concentrated on the periphery of these bands; they can form patterns on the surface of mild steel that has been annealed at around 300°C, giving it an unattractive "orange peel" appearance, something that clearly is to be avoided.

We now study the kinetics of this hardening mechanism. If D is the diffusion coefficient for carbon and nitrogen atoms, their rate of migration $\underline{v}$ satisfies

$$\underline{v} = \frac{-D\,\underline{\underline{\text{grad}}}U}{\text{k}T} \tag{3.126}$$

[17]Guillaume Piobert (1793–1871) was a French mechanical engineer and scientist who observed and described these bands in 1842 when studying shocks on plates. W. Lüders published his observations in 1860.

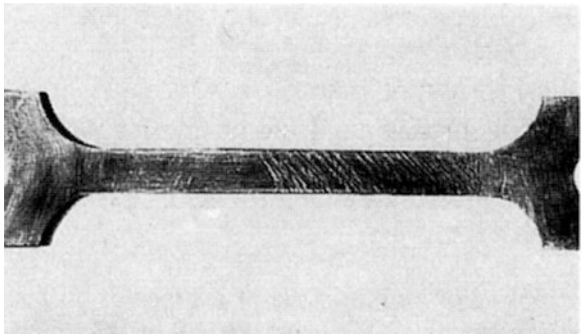


Fig. 3.80 Piobert-Lüders bands on a mild steel test piece

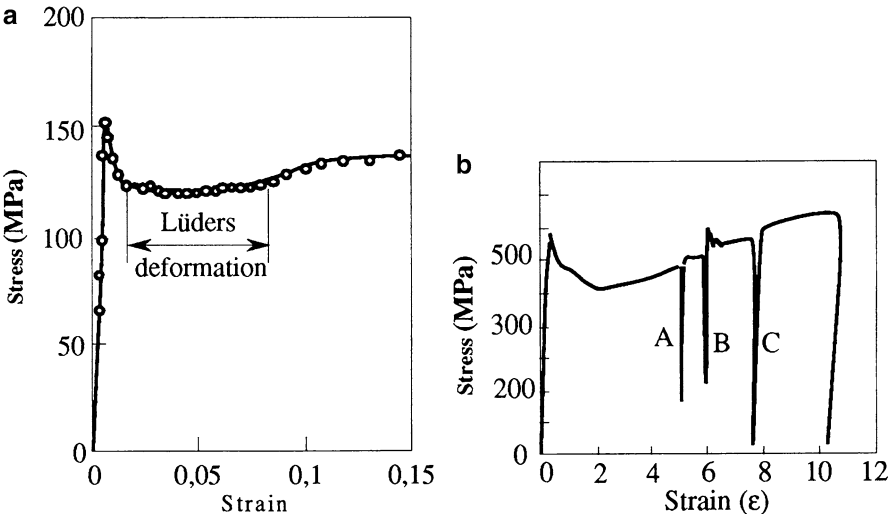


Fig. 3.81 Yield drop (**a**) iron single crystal and (**b**) mild steel: A unloading followed immediately by reloading; B unloading followed by 100°C annealing for 10 min; C unloading followed by 100°C annealing for 1 min

From (3.111) giving the interaction energy U between a solute atom and a dislocation, and ignoring the angular variation, we get

$$v = \frac{D}{\mathrm{k}T}\frac{A}{r^2} \tag{3.127}$$

where $\mathrm{A} = 2.3\mu {a_0}^3 \Delta a_0 = 2.3\mu {a_0}^4 \eta$.

After a time t the atoms, that have migrated to the core of the dislocation, were less than a distance r_t. If their migration rate is $v(r)$ this is expressed by:

$$\int_{t_r}^{0} \frac{-\mathrm{d}r}{v(r)} = 1 \tag{3.128}$$

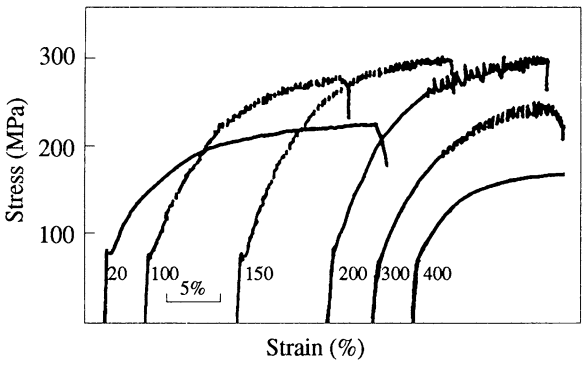


Fig. 3.82 Portevin-Le Chatelier effect in mild steel, strained at 10^{-4} s^{-1} at various temperatures

whence:

$$r_t = \left(\frac{3DAt}{kT}\right)^{1/3} \tag{3.129}$$

From this we find that, if the concentration of solute atoms in the solid solution is c_0, that on the dislocation line is:

$$c = \frac{\pi r_t^2 b c_0}{b^3} = \frac{\pi c_0}{b^2}\left(\frac{3ADt}{kT}\right)^{2/3} \tag{3.130}$$

Thus the hardening kinetics, which is the same as that of the formation of the atmospheres, follows a $t^{2/3}$ law.

This agrees with the results of observations, and the activation energy is found to agree well with that of the diffusion of carbon and nitrogen atoms.

(e) *Portevin-Le Chatelier effect*

If the temperature is high enough and the rate of plastic deformation not too great, the solute atoms, as they diffuse, can recapture the slipping dislocations. More precisely, the dislocations continue to slip until they meet an obstacle, when they are stopped and anchored by the diffusing atoms; to make them move again requires an increase in the stress. The process is repeated, giving a series of kinks on the tensile curve. This is the Portevin-Le Chatelier[18] effect, which can be seen in experimentally-observed curves such as Fig. 3.82 (Le Chatelier and Portevin A 1923).

[18] Albert Portevin (1880–1962) was a French metallurgist. Henry Le Chatelier (1850–1936) was a French chemist.

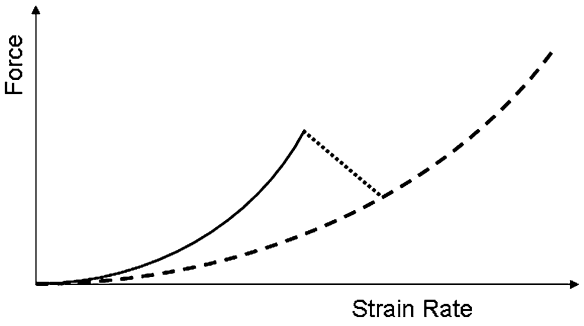


Fig. 3.83 Sketch showing the strain rate dependence on the force to move a dislocation. (**a**) Dragging force in the presence of Cottrell atmospheres; (**b**) Lattice friction force

Mild steels with a BCC structure extended at normal strain rates ($\sim 10^{-4}$ s^{-1}) and at intermediate temperatures ($\sim$300°C) present a succession of yield points or *serrations* which makes them brittle (Fig. 3.82). If during the plateau which follows the initial yield point (due to Piobert-Lüders bands associated with *static ageing* (Sect. 3.4.3)) the zones from which the Lüders bands have started have had the time to age, the deformation after the end of the initial plateau will not be homogeneous: it will reinitiate in the non-aged zones, then propagate into the aged zones forming new Lüders bands and new *serrations* on the stress-strain curve. This type of ageing is named *dynamic ageing* to differentiate it from *static ageing* which produces only one yield point at the onset of plastic deformation. Blue brittleness ($\sim$300°C) should appear when the ageing time (Eq. 3.130) becomes shorter than the time to propagate a Lüders band (Friedel 1964).

Portevin-Le Chatelier effect has also been observed close to room temperature in aluminium alloys with a FCC structure, and in many austenitic iron- or nickel-based alloys, such as Inconel 718 (Ni-20Cr-20Fe-5Nb-3Mo-2.5Ti-2Al). The substitutional solutes in these alloys diffuse much too slowly for repeated yield points to be produced from the beginning of the stress-strain curves. The vacancies created by deformation are then probably present in enough concentration to accelerate sufficiently the diffusion of the solute atoms (Friedel 1964).

One way of representing the source of the Portevin-Le Chatelier effect is shown in Fig. 3.83 where two regimes for propagating a dislocation at a given strain rate are considered. At low strain rate the dislocation moves with its atmosphere of impurities (carbon and nitrogen atoms in BCC crystals; substitutional solutes and eventually carbon atoms in FCC crystals). This regime corresponds to a dragging force. At high strain rates, the solute atoms have not enough time to diffuse during the displacement of the dislocations and to form atmospheres around them. In this regime the force to move the dislocations is only due to the lattice friction force. In between these two extreme regimes there is a domain of strain rate where the force to move a dislocation decreases with strain rate as indicated schematically in Fig. 3.83. This regime is thus accompanied by a negative strain rate effect, which produces a chaotic phenomenon.

In FCC alloys deformed at a given strain rate, the strain at the onset of *serrations* on the stress-strain curves is an increasing function of temperature, as shown in

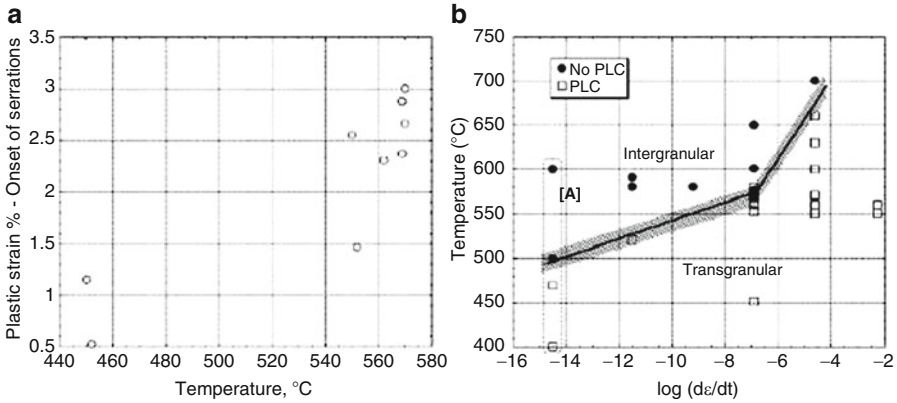


Fig. 3.84 Portevin-Le Chatelier effect in Alloy 718. (**a**) Evolution of the onset of serrations (critical plastic strain) as a function of temperature; (**b**) Fracture modes (transgranular *vs.* intergranular) and plastic flow modes (existence of Portevin-Le Chatelier (PLC) effect)

Fig. 3.84a. This critical strain appears at higher temperatures when the strain rate is increased (Fig. 3.84b). The influence of the PLC effect on the mechanical behaviour of materials is still largely discussed. The negative strain rate dependence of the flow stress may be the source of strain localisation which leads to a reduction in ductility. Recently the "beneficial" influence of the Portevin-Le Chatelier effect has been underlined in Alloy 718 tested in air environment (Garat et al. 2008). At low strain rate and in air this material gives rise to intergranular fracture due to environmental effect when tested around 500–600°C. As shown in Fig. 3.84 this mode of failure disappears when the strain rate is sufficiently low such as the Portevin-Le Chatelier effect is suppressed. In the presence of *serrations*, the strain rate can be locally sufficiently large to annihilate the deleterious effect of the interactions between deformation and intergranular oxidation.

(f) *Chemical (Suzuki) effect*

We have seen in Sect. 3.4.2.2(g) that cross-slip of dislocations is related to the stacking fault energy. If we dissolve in the metal a solute that stabilises the HCP phase relative to the FCC this will lower the stacking fault energy. When the solute atoms migrate to the dislocation, they stabilise the dissociation. A strong interaction energy results (for example, a reduction in γ_f from 45 to 6 mJ m^{-2} when 30% of Zn is dissolved in Cu) and an increase in the yield strength, related to the segregation of foreign elements on the fault. This is called the *Suzuki effect*[19] (Suzuki 1952). These alloying effects, which reduce the stacking-fault energy, can increase the hardness considerably in FCC metals by inhibiting Stage 3 of the work-hardening process.

[19]Hideji Suzuki was a Japanese scientist.

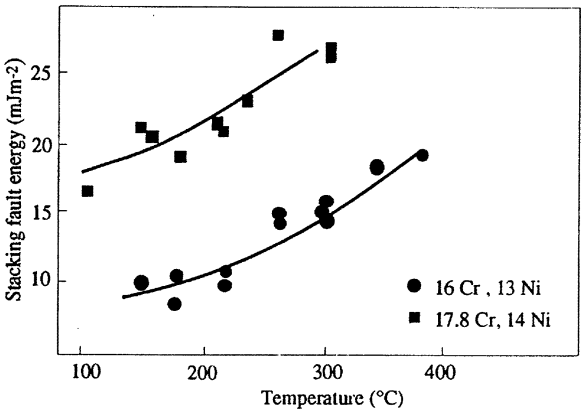


Fig. 3.85 Variation with temperature of the stacking fault energy of two austenitic stainless steels (Lecroisey and Thomas 1970)

It provides a method of reinforcement that is used with many FCC industrial alloys for which the stacking-fault energy at ambient temperatures is low: for example, stainless steels of type 18Cr-12Ni, or Hadfield (20Mn-5Cr) steel, for which $\gamma_f \approx 20$ mJ m^{-2}.

However, the method is effective only at low temperatures. In the materials we have been discussing the stacking fault energy is strongly dependent on temperature, as is shown by the results for stainless steels given in Fig. 3.85 and as indicated previously (see Sect. 3.4.2.2(g)). This dependence is related to the entropy change $\Delta S^{\text{fcc}\rightarrow\text{hcp}}$ associated with the FCC→HCP transition, since a fault corresponds locally to the structure of the HCP phase, and it can be shown, from basic considerations, that $\mathrm{d}\gamma_f/\mathrm{d}T = -(8/\sqrt{3})a^2 \mathrm{N}\Delta S^{\text{fcc}\rightarrow\text{hcp}}$ where N is Avogadro's number.

3.4.3.2 Hardening by Precipitates

(a) *Various types of precipitates and of interactions with dislocations*

Spreading out precipitates or particles of a second-phase in a matrix is one of the most efficient ways to harden it. Precipitation can be achieved in alloys systems in which the limit of solubility of foreign atoms increases when the temperature is raised. Aluminium alloys provide a typical example. The Al-Cu phase diagram shows that the solubility of Cu falls from 5% by weight at 550°C to almost zero at ambient (Fig. 3.86). The precipitation stages start with a solid solution at high temperature, and include very quick cooling by quenching, so as to obtain a solid solution that is supersaturated at ambient temperature, and then ageing at ambient temperature or annealing at a higher temperature causing precipitates to form (see Annex 2).

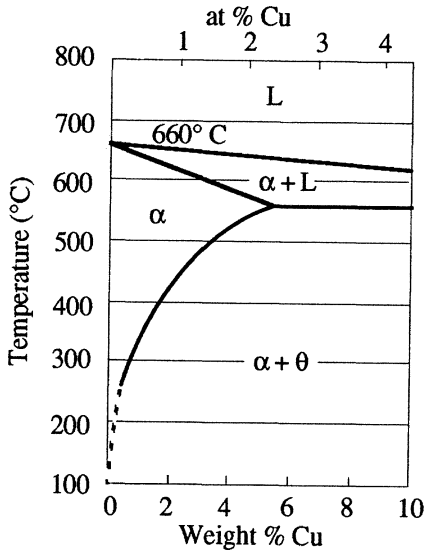


Fig. 3.86 Al-Cu equilibrium phase diagram

A 4% Cu alloy quenched from 550°C evolves as follows (see Figs. 3.87 and 3.88):

- At first the Cu concentration fluctuates, the atoms clustering in certain regions to form *Guinier-Preston (GP) zones* (Preston 1938) of very small size. Guinier and Preston,[20] independently, observed these zones by X-ray diffraction. In AlCu alloys they are one-atom thick disks (Fig. 3.89) (Jouffrey and Dorignac 1992). These clusters give then rise to precipitates θ', which, whilst coherent, have a well-defined interface with the matrix; their lattice is continuous with that of the aluminium.
- As the precipitates grow in size the coherence tends to be lost as a result of dislocations being created on the surfaces: Fig. 3.90 gives an example of this semi-coherent intermediate state.
- The final stage is the formation of precipitates of $CuAl_2$, which are not coherent.

The shape of the GP zones as well as that of the precipitates depends on the alloying element. In Al Ag for instance they are spherical.

In some alloys, such as Ni-based alloys, the precipitates of the phase γ' (Ni_3TiAl) are ordered.

Irradiation can produce foreign atoms by transmutation, for instance atoms of He. It also creates point defects, in particular vacancies. These He atoms and these vacancies are supersaturated and they go back to equilibrium by clustering. Thus, He bubbles and cavities are formed in a similar way as precipitates (refer to Sect. 3.4.3.2(h)).

[20] André Guinier (1911–2000) was a French physicist; George Dawson Preston (1896–1972) was a British physicist.

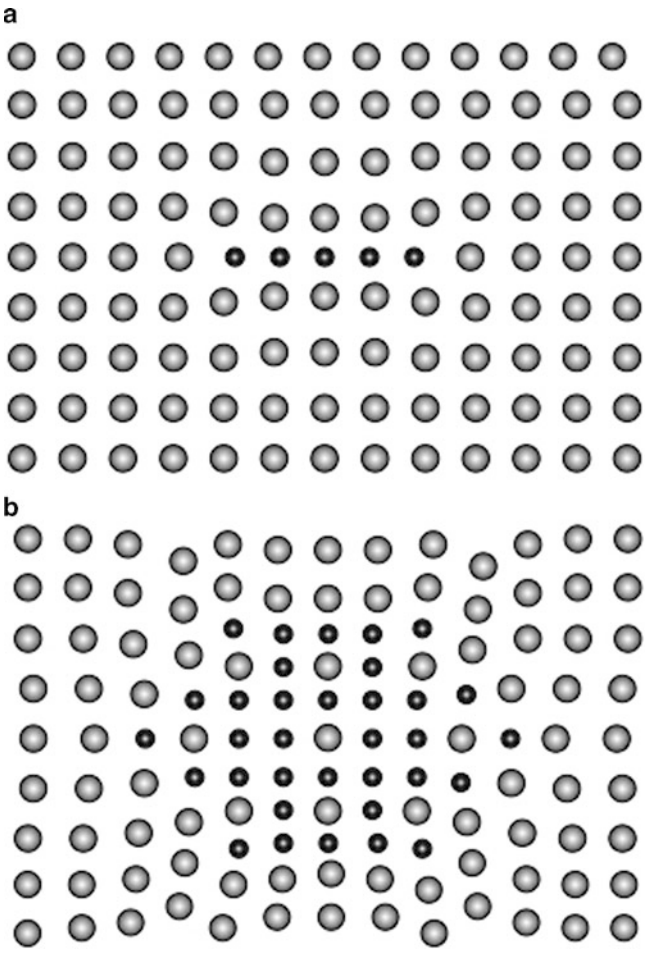


Fig. 3.87 GPI and GPII zones

Various techniques are available to incorporate particles in a metal matrix: powder metallurgy, mechanical alloying. A great number of combinations of inclusions and matrix are possible, for instance aluminium reinforced by SiC or Al_2O_3 particles. This has the advantage that the particles can have a better temperature stability than precipitates, but the disadvantage that it can be difficult to disperse them evenly (see Chap. 1 and Annex 2).

Thus there are two kinds of precipitation hardening: *coherent precipitates* and *incoherent* ones or *particles*. The dislocations can penetrate in the first but not in the second, which they must by-pass by the *Orowan mechanism*. It consists in the bending of the dislocation until the radius of curvature reaches the critical value,

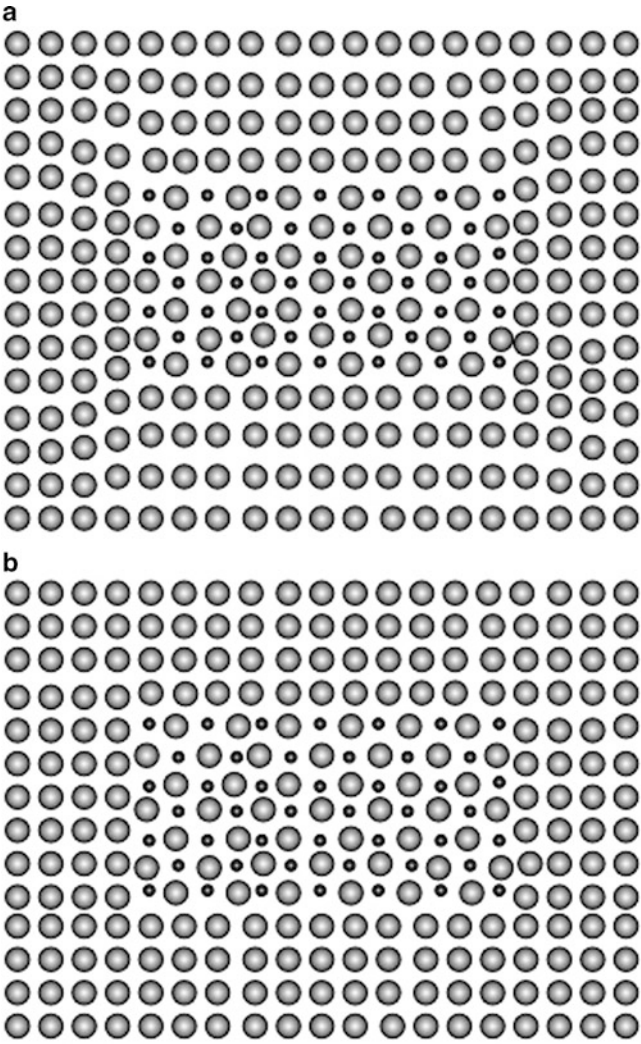


Fig. 3.88 Coherent and incoherent precipitates

followed by the expansion of the loop leaving circular rings around the obstacles (Figs. 3.91 and 3.92) (see also Bacon et al. 2009 and Chap. 4).

Cross-slip is another way to overcome these obstacles, as will be discussed later.

Coherent precipitates can be sheared by the crossing of dislocations (Fig. 3.93).

(b) *Shearing of coherent precipitates*

The penetration of a dislocation in a coherent precipitate is hampered by several factors. Each one of them can be represented by a resisting force $\underline{K}$ acting at the

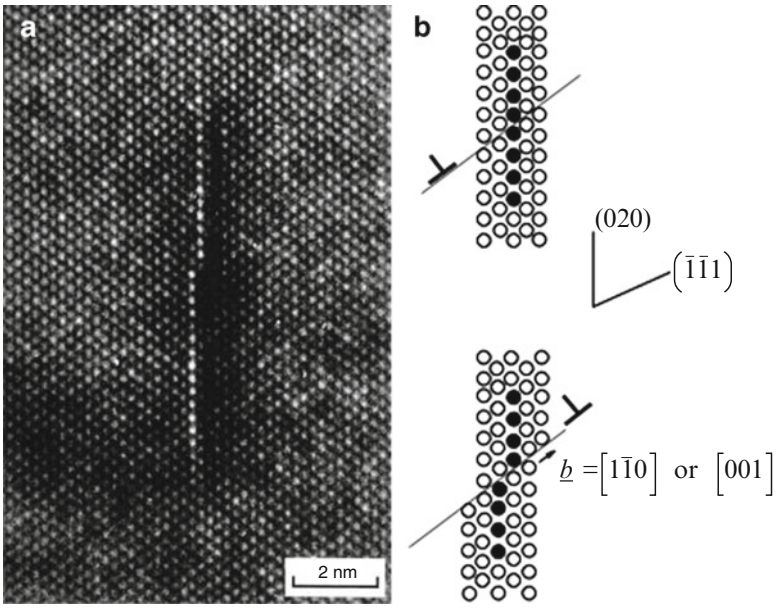


Fig. 3.89 High resolution electron microscopy of an AlCu alloy exhibiting a GP zone, which was split by the crossing of a dislocation (Adapted from Karlik and Jouffrey 1996)

point where the dislocation leaves the precipitate. A detailed treatment is given in Kocks et al. (1975). These resisting mechanisms are the following:

- The Peierls friction stress τ_{PN}.
- If the precipitate is ordered, the passage of a dislocation creates an antiphase boundary which has an energy γ_A per unit of surface area: this is especially so for the very important case of *nickel-based superalloys* containing ordered precipitates of the phase γ' Ni_3Al (Fig. 3.93), and some Al-Li alloy containing ordered $\delta' Al_3Li$ precipitates.
- Since the precipitate and the matrix will have different specific volumes, there will be a size effect characterised by the parameter $\eta = \Delta a_0/a_0$.
- The core energy of the dislocation can be different in the interior of the precipitate, in which case the line tension will be changed
- The precipitate and the matrix can have different elastic moduli; this includes the case of cavities, for which the modulus is clearly zero.
- The passage of a dislocation shears the precipitates and produces a step of height b at the interface, with an interface energy γ_i.

We need to be able to calculate the value of $\underline{K}$, the resistance of the precipitate, in each of the preceding cases. These values are listed in Table 3.8.

In this table W is the width of the particle, that is its diameter in the case of a spherical particle, and θ the angle at which the dislocation enters the spherical precipitate (Fig. 3.94).

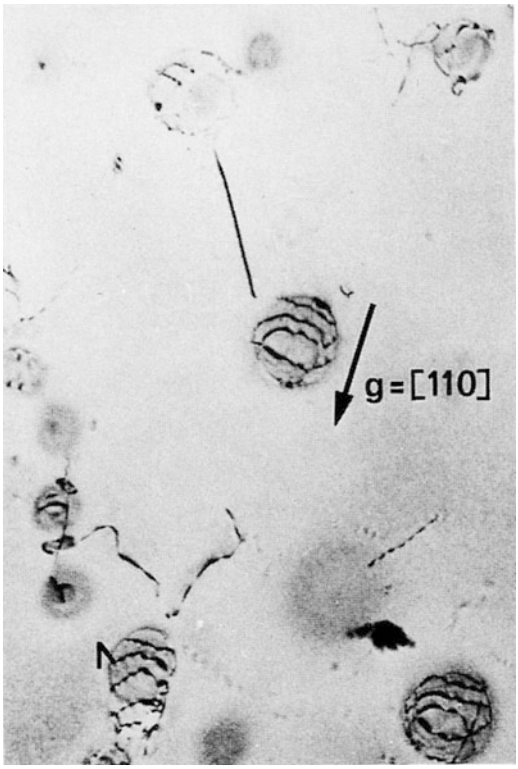


Fig. 3.90 Precipitates of NiAl in a ferritic FeCrNiAl alloy, after tempering at 715°C for 16 ½ days, showing dislocations in the semi-coherent interfaces

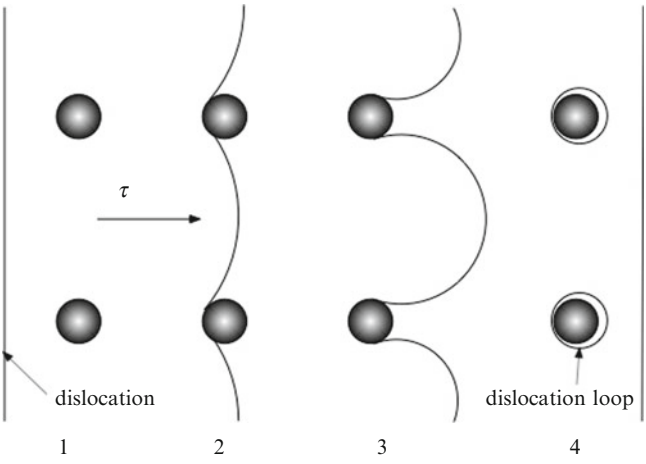


Fig. 3.91 Sketch of the Orowan mechanism

As an example we take the creation of a step on the interface by the passage of a screw dislocation. In Fig. 3.94 K_x and K_y are the forces exerted at the point where the dislocation enters the precipitate; in the case we are considering the increase in energy due to the step is $by\gamma_i$ and therefore

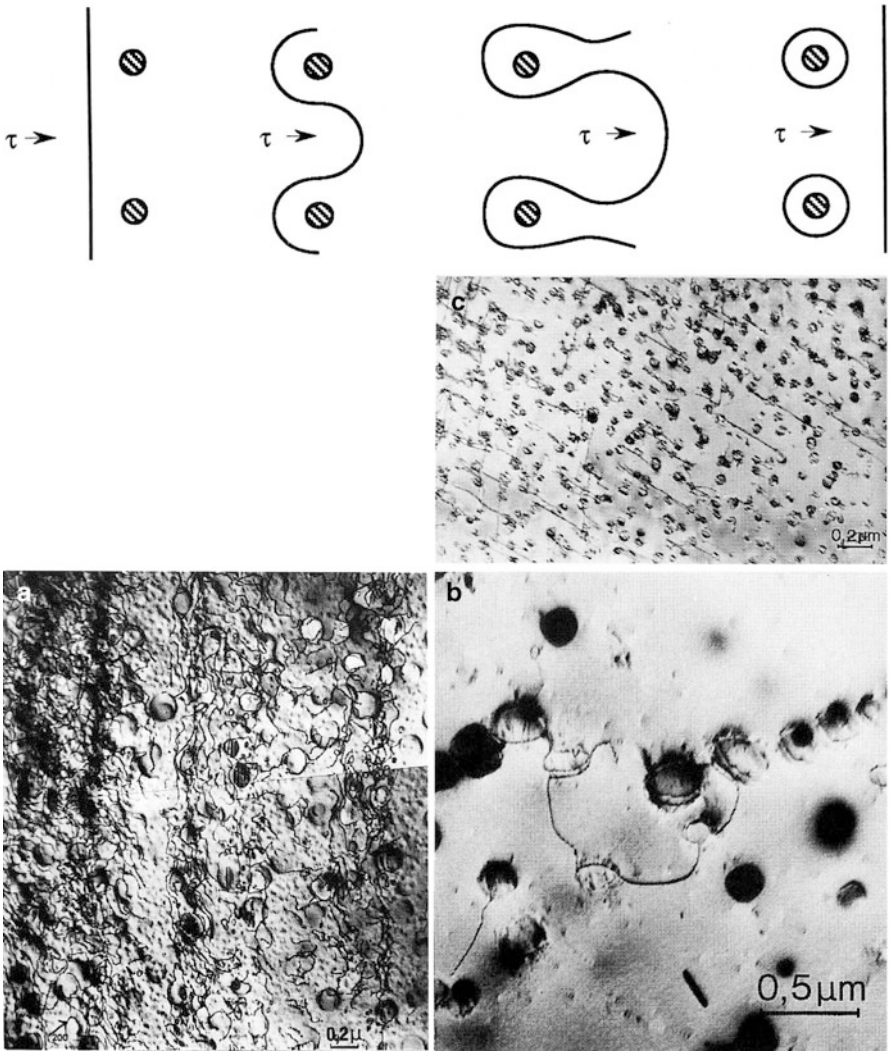


Fig. 3.92 Orowan mechanism (**a**) γ' particles in a nickel-based alloy (waspaloy) strain-cycled ($\Delta\varepsilon_p/2 = 0.2\%$) at 650°C, (**b**) particles in ferritic stainless steel (FeCrNiAl) aged at 650°C for 6 h and hardened by precipitation of the NiAl phase, followed by work-hardening

$$K_y = \frac{d(by\gamma_i)}{dy} = b\gamma_i$$
$$K_x = 0 \tag{3.131}$$

We now seek the stress needed to shear a precipitate, which we assume to be spherical of radius r. From Fig. 3.94, if $K_y(\theta)$ is the resistance offered by this

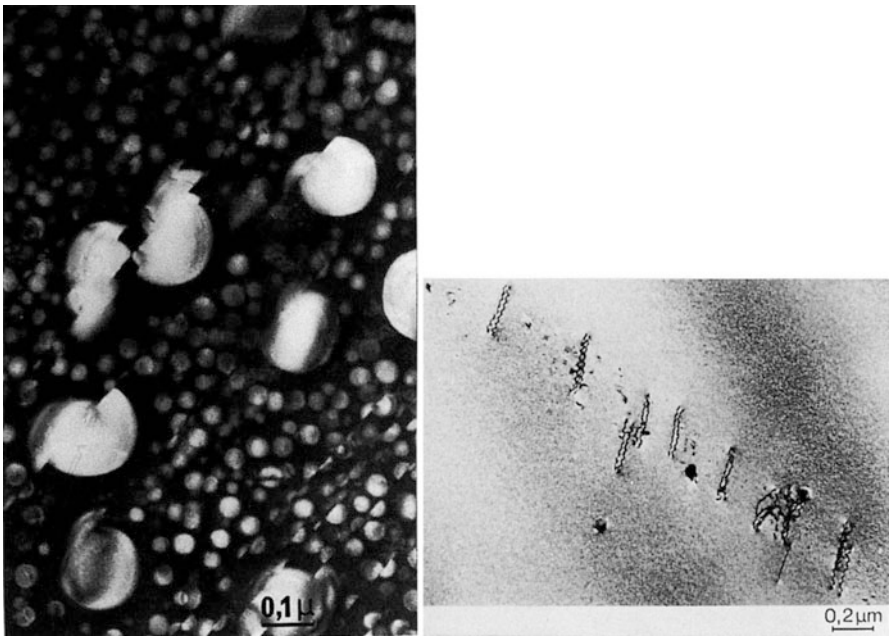


Fig. 3.93 (**a**) Shearing of precipitates of γ' in a nickel-based superalloy (waspaloy (18.7Cr, 14Co, 1.45Al, 8.9Mo, Bal Ni)) aged at 600°C for 10 min and tensile strained by 0.8%, (**b**) The dislocations are grouped by pairs so as to reduce the area of the antiphase boundaries created by shearing the precipitates

Table 3.8 Individual particle resistances (Kocks et al. 1975)

Mechanism Order of magnitude	Resistance to edge dislocation	Resistance to screw dislocation	Activation energy
Friction stress $\mu/100$	$K_y = \tau_{PN} b\,(W/2)\cos\theta$	$K_y = \tau_{PN} b\,(W/2)\cos\theta$	$\frac{\pi}{4}\tau_{PN} b W^2$
Disordering $\gamma_A \approx \mu b/100$	$K_y = \gamma_A\,(W/2)\cos\theta$	$K_y = \gamma_A\,(W/2)\cos\theta$	$\frac{\pi}{4}\gamma_A W^2$
Misfit stresses $\eta \approx 1/100$	$K_{y\max} \approx \lvert\eta\rvert\,\mu b W$	Small	$\approx \mu b W^2\,\lvert\eta\rvert$
Core energy difference $\Delta t \approx t/100$	$K_x = \Delta t_{edge}$	$K_x = \Delta t_{screw}$	$\lvert\Delta t\rvert\,W$
Modulus difference $\lvert\Delta\mu\rvert \approx \mu/2$	$K_x \approx \Delta\left(\frac{\mu}{1-\nu}\right)\frac{b^2}{5}$	$K_x \approx \Delta\mu\frac{b^2}{5}$	$\lvert\Delta\mu\rvert\,b^2 W$
Interface step $\gamma_i \approx \mu b/100$	$k_x = b\gamma_i \operatorname{sign}(-y)$	$K_y = b\gamma_i$	$K_y = 2bW\gamma_i$

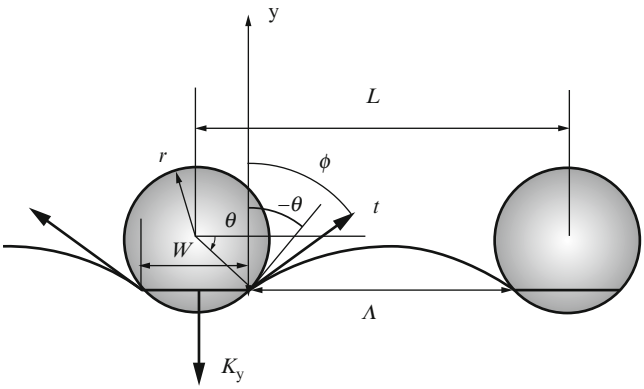


Fig. 3.94 Shearing of a coherent precipitate by a dislocation pushed by a force τb in the direction y

obstacle and t the line tension, the equilibrium equation at the point where the dislocation enters the obstacle is:

$$K_y - \tau b r \cos\theta = t \cos\phi \tag{3.132}$$

If a dislocation is anchored between two points distant of Λ

$$2t \cos\phi = \tau b \Lambda \tag{3.133}$$

Hence the elimination of $2t\cos\phi$., since $\Lambda = L - 2r\cos\theta$, yields $\tau = 2K_y/Lb$.

Now the angle ϕ, which the dislocation makes with the y direction where it enters the precipitate, must remain larger than $-\theta$. Otherwise the dislocation wraps around the precipitate. From the preceding Eqs. 3.132 and 3.133, it means that:

$$\cos\theta \geq \left(\frac{t}{K_y} + \frac{2r}{L}\right)^{-1} \tag{3.134}$$

If the term between braces in this equation is smaller than 1, this condition cannot be fulfilled. The precipitate cannot be sheared.

When shearing occurs, the critical shear stress is given by:

$$\tau_c = \frac{2K_y(\theta)_{max}}{Lb} \tag{3.135}$$

When the second member of the condition (3.135) is larger than 1, the dislocation penetrates in the precipitates until $\phi = -\theta$, and then it wraps around it and we deal with a non-shearable precipitate, a case we are to study in the next paragraph.

However, as in the case of isolated atoms, we need to modify the relation we have just derived to take account of the statistical distribution of the precipitates. Friedel's

statistics equates the area $\Lambda^3/2R$ swept by a dislocation each time it escapes from a precipitate to the mean area per precipitate, L^2; this gives:

$$\tau_c = \frac{2t(\cos\phi)^{3/2}}{Lb} \tag{3.136}$$

$\cos\phi$ being given by Eq. 3.132.

Dislocation movement through random arrays of obstacles has been simulated numerically by Foreman and Makin (1966). The shear stress required to move the dislocation through the array has been determined for the complete range of breaking angles ϕ from 0 to $\pi/2$ and the results have been expressed in an empirical form. The dislocation is observed to move often by an "unzipping" mechanism involving consecutive breakaway from obstacles on the dislocation (Figs. 3.95a, b). For weak obstacles ($\phi \to \pi/2$) the critical shear stress is appreciably less than for a regular square array of the same density and corresponds with the Friedel relation for the spacing of obstacles along a dislocation, which remains straight (Fig. 3.95a). For strong obstacles ($\phi \to 0$) the dislocation motion is characterised by encircling movements and the formation of many closed loops, as observed in Fig. 3.95b. The critical shear stress rises to a value of 0.81 $\mu b/L$, where L is the average spacing. This expression is similar to the equation given later for the *Orowan mechanism.* The results of these numerical simulations are reproduced in Fig. 3.95b where the critical shear stress σ is plotted as a function of breaking angle ϕ for a random array of 10,000 obstacles.

In the presence of particle shearing, when K_y is proportional to the particle size (see Table 3.8), the breaking angle ϕ, given by Eq. 3.134, decreases when the particle size increases for a given volume fraction. The results shown in Fig. 3.95b indicate that the critical shear stress increases with particle size, following roughly a parabolic law. The maximum shear stress occurs when $\theta = 0$ in Fig. 3.94, *i.e.* when $K_y = \gamma_A r$, where r is the particle radius, for disordering is the obstacle. The corresponding shear stress is given by:

$$\tau_c = 2\gamma_A r\,/bL \tag{3.137}$$

(c) *By-passing of precipitates and particles*

A dislocation that slips in a crystal containing precipitates, which cannot be sheared, can by-pass them by what is called the *Orowan mechanism* (Orowan 1948) (Figs. 3.91 and 3.92), analogous to the Frank-Read source, leaving a loop around the precipitate. The stress needed for this is given by Eq. 3.51, with l replaced by $L-2r$, where r is the radius of the precipitate:

$$\tau_{OR} = \frac{\mu b}{L-2r} = \frac{\mu b}{2r} f_v^k \tag{3.138}$$

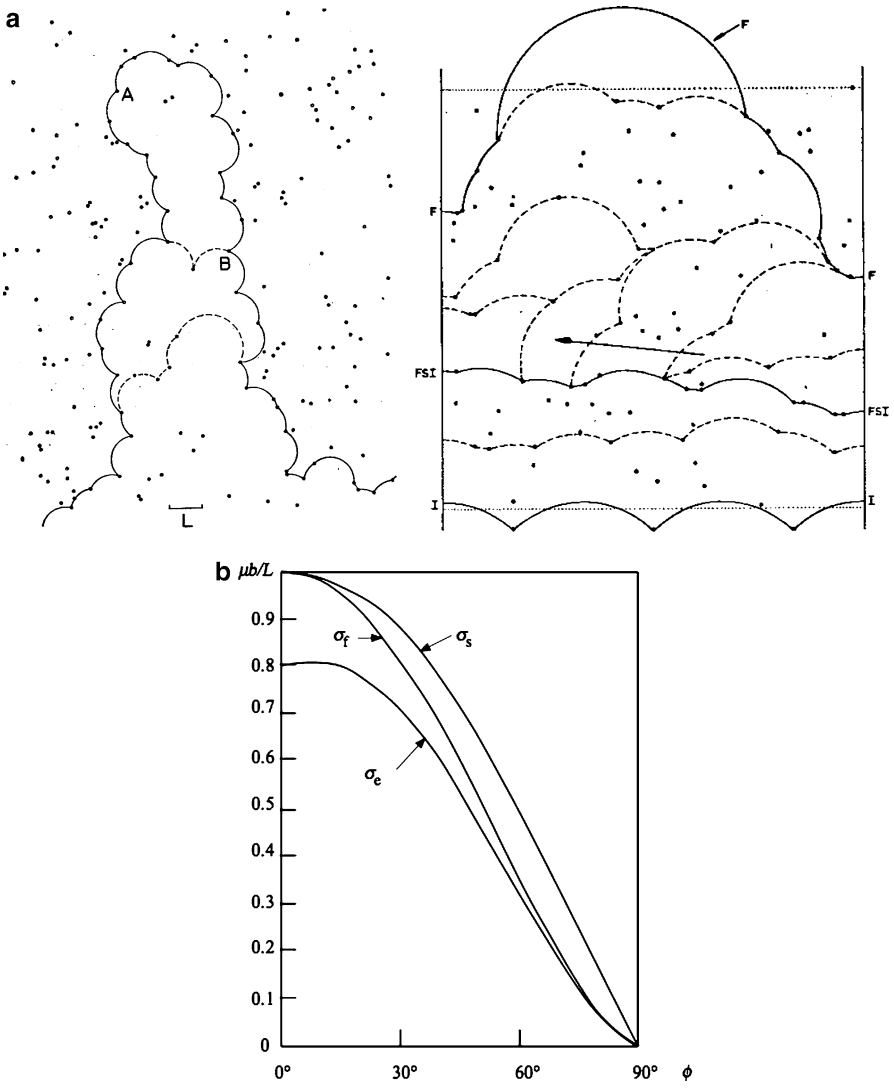


Fig. 3.95 (**a**) Numerical simulation of dislocation movement through an array of random obstacles; (a *left*) breaking angle $\phi = 65.5°$ showing initial (I) and final (F) positions and the position at the final stress increase (FSI). *Broken lines* show some intermediate positions and illustrate unzipping (see *arrow*) from the FSI position; (a *right*) breaking angle $\phi = 5°$. Solid line indicates the position reached under a stress of 0.7 $\mu b/L$. The removal of obstacles A or B would allow the dislocation to penetrate further in the array (Foreman and Makin 1966). (**b**) Variation of critical shear stress σ for a random array of 10,000 obstacles and 1% stress increment (*solid circles*), and for an array of 1,000 obstacles with 2% increment (*open circles*). The square lattice σ_s, the Friedel stress, σ_f and the numerical results, σ_e are shown (Foreman and Makin 1966)

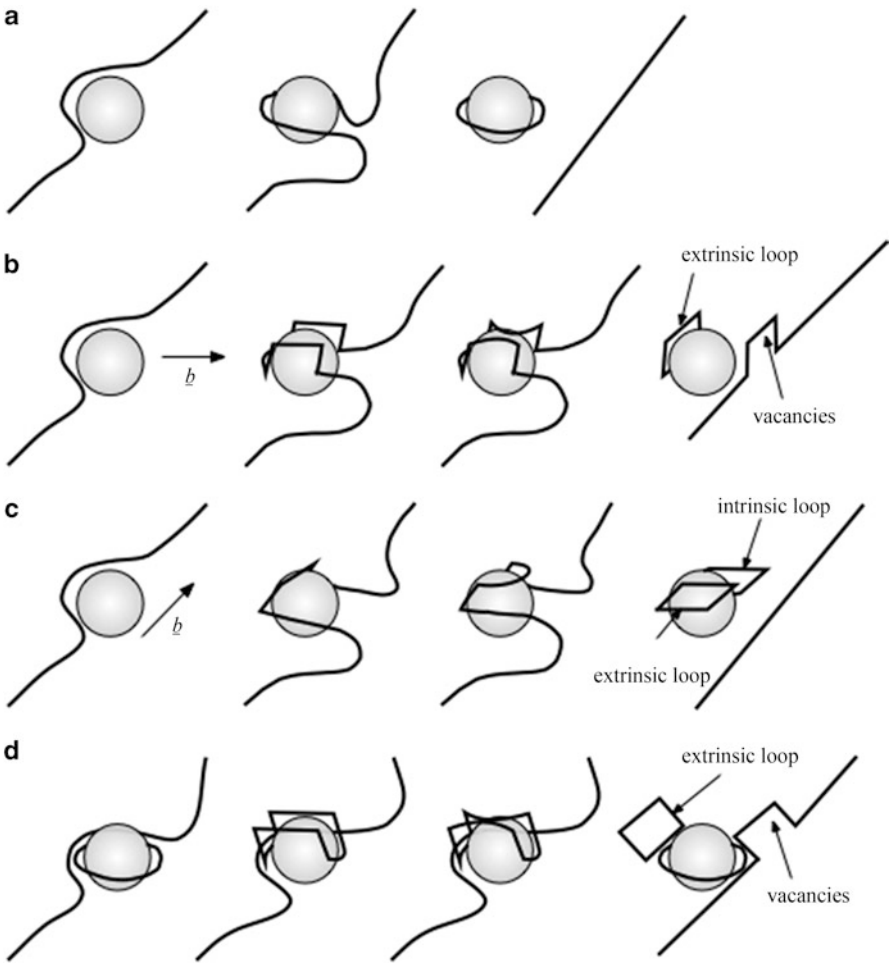


Fig. 3.96 (**a**) Dislocation by-passing a precipitate by the Orowan mechanism. (**b**) and (**c**) Dislocation by-passing a precipitate by cross-slip leaving intrinsic and extrinsic prismatic dislocations loops whose Burgers vector is that of the dislocation which slips. (**d**) As case (**a**) but followed by cross-slip of the loop

which will occur for $\phi = 0$ ($\cos\phi = 1$), corresponding to $K_y = t$ (see also Chap. 4). f_v is the volume fraction of precipitates or particles and the power k depends on the relation between $2r/L$ and the volume fraction.

Cross-slip is another mechanism, which allows by-passing of dislocations around precipitates or inclusions (Fig. 3.96). This leaves intrinsic and extrinsic prismatic dislocation loops such as the ones described in Sect. 3.3.4.2.

The numerical simulation by Foreman and Makin (1966) have shown that this expression overestimates the critical shear stress by about 20% when the

particles are randomly dispersed. The above expression compared with the equation developed earlier for particle shearing allows to predict that the transition between both mechanisms will occur for a critical particle size, r_t given by:

$$r_t \simeq t\,/\gamma_A \tag{3.139}$$

where t is the line tension of the dislocation.

For a given volume fraction of particles, this theory predicts that for small sheared particles ($r < r_t$) the critical shear stress increases with the particle size according to a parabolic law up to a maximum value of the order of $(t\,/2rb)\,\sqrt{f_v}$ and then decreases according to the Orowan law which predicts that, for small volume fraction, $\tau_{OR} \simeq (\mu b\,/2r)\,\sqrt{f_v}$.

The expression derived for r_t predicts that, for instance in nickel-based superalloys strengthened by γ' precipitates, the transition radius is of the order of 100 nm, using $\gamma_A = 40$ mJ/m^2. This value for r_t is of the order of experimental values observed in these materials. Further details are given in Chap. 4 in this volume.

(d) *Evolution of the yield strength of aluminium alloys by ageing*

In the course of *ageing* of aluminium alloys the volume fraction f_v of the precipitates varies; the mean distance between the precipitates is $\Lambda = L(1-f_v)$, since the linear fraction is the same as the volume fraction.

Initially, f_v increases and L remains constant; the precipitates are coherent and are sheared, and the yield strength increases. When all the atoms have been precipitated L starts to increase, with f_v remaining constant. This is over-ageing: according to the Orowan mechanism the yield strength decreases (Fig. 3.97). Noting from Table 3.8 that $K_y \approx \mu br/100$, evaluating the transition radius r_t for the shift from particles shearing to particles by-passing from Eq. 3.134 and the yielding shear stress by Eq. 3.135, it can be shown that $r_t \approx 100\,t/\mu bL$, so that $r_t/b \approx 50$, and that the yield stress at the transition reaches $(b/L)\mu$ (Fig. 3.97).

(e) *Work-hardening*

Work-hardening varies greatly according as the precipitates are or are not sheared. With shearing, the passage of one dislocation eases that of the next, for example as a result of the order being destroyed by shearing (Fig. 3.93); this gives a low level of work-hardening. On the other hand it is very large when at each passage the Orowan mechanism leaves a dislocation loop around the particles (Fig. 3.98).

The dislocation loops around the particles are of the geometrically necessary class (*cf.* Sect. 3.3.3.5). Using the results previously obtained for the density of geometrically necessary dislocations around *spherical inclusions* of diameter D (3.15) and for work-hardening (3.85) we find:

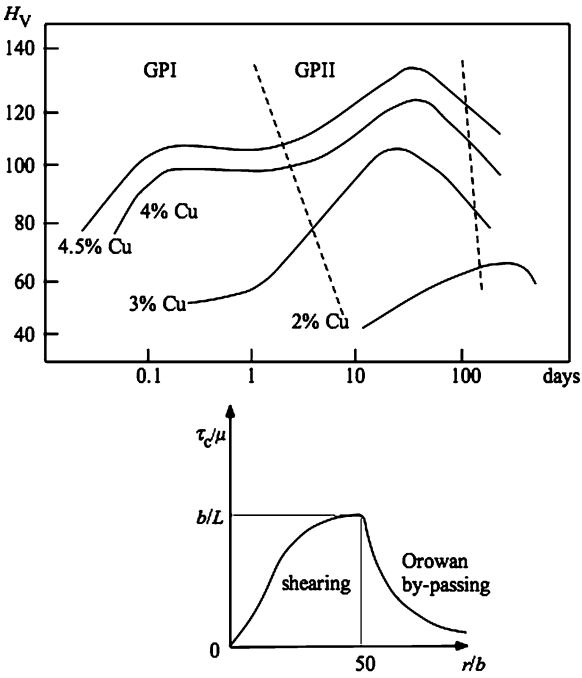


Fig. 3.97 Variation of the hardness of AlCu alloys with annealing time at 130°C

$$\tau = \tau_i + \alpha\mu b\left[(288\pi)^{1/3}\frac{\gamma}{bD}f_v^{2/3}\right]^{1/2} \simeq \tau_i + \mu f_v^{1/3}\left(\frac{b}{D}\gamma\right)^{1/2} \tag{3.140}$$

for a shear strain γ.

When a carbon steel is cooled, the carbon content of the austenite increases until the eutectoid composition (0.8% C) is reached. Platelets of alternating cementite (Fe_3C) and ferrite (αFe) then form, giving a structure called *pearlite*. This is an example of reinforcement by *platelets*, which, because of the covalent bonds of the cementite, are quasi-rigid. Deformation of pearlite is accompanied by the introduction of geometrically necessary dislocations. From formulae (3.17) and again (3.85), we find:

$$\tau = \tau_i + 2\alpha\mu\left(\frac{b}{\Delta}\gamma\right)^{1/2} \tag{3.141}$$

Δ being the distance between the platelets of cementite.

However, relation 3.140 was derived on the assumption that the dislocations accumulated randomly in a crystal in response to the effects of anchoring and multiplication, described at the time. Geometrically necessary dislocations are not in fact stacked randomly, so the case for using Eq. 3.85, which has led to Eq. 3.140,

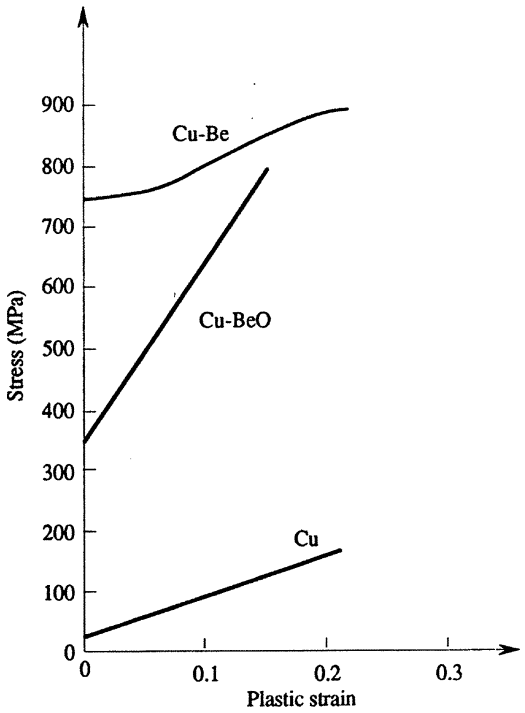


Fig. 3.98 Work-hardening of single crystals of copper or CuBe alloy hardened by particles of BeO or by precipitates

predicting the parabolic variation of work-hardening, could be questioned. This emphasises the difficulty of relating the macroscopic and microscopic aspects of plasticity.

Even so, the value of the density ρ_G of the geometrically necessary dislocations obtained in this way can be a valuable help in fixing orders of magnitude, especially in regions in the crystal in which the presence of particles is expected to have a strong effect. With a shear strain γ, $\rho_G \approx 10(\gamma\ /bD)f_v^{2/3}$ for spherical particles of diameter D, and $\rho_G = 4\gamma/b\Delta$ for platelets a distance Δ apart. The quantities $D/f_v^{2/3}$ and Δ both have the dimension of length; we denote them by λ_G and compare them with the distance d between the slip lines in a crystal in which there are no particles and in which dislocations accumulate only randomly. The conclusion is that the presence of particles will have a strong effect on work-hardening only when $\lambda_G << d$, that is $\rho_G >> \rho_S$, where ρ_S is the density of randomly-stacked dislocations (*cf.* Fig. 3.99).

Equations 3.140 and 3.141 considered predict *parabolic work-hardening* for the two kinds of reinforcements. But Fig. 3.98, for the alloy Cu-Be containing particles around which the dislocations wrap, shows that in fact it is more like linear. In what follows, therefore, we shall look again at the nature of work-hardening in alloys hardened by precipitates. In so doing we shall take the opportunity to sketch the relation between behaviour on the macroscopic scale and plasticity on the

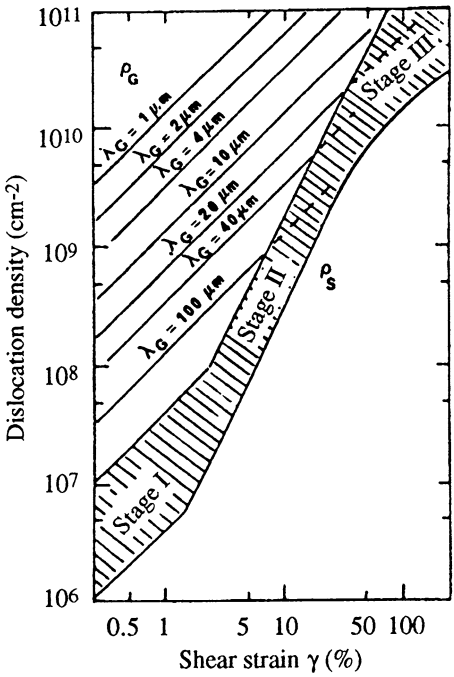


Fig. 3.99 Densities of geometrically necessary dislocations ρ_G and of randomly distributed dislocations ρ_S in a FCC crystal as a function of deformation. Note that for small deformations ρ_G can be much greater than ρ_S whilst the reverse can be true at large deformations

microscopic scale, and to indicate what contribution to the solution of the problem might be made by the mechanics of inclusions developed in the previous chapter (Sect. 2.7).

(f) *Kinematic hardening of heterogeneous materials*

As we have shown, the effect of precipitates surrounded by dislocations is to increase greatly the work-hardening rate in a material subjected to monotonic increasing loading (Fig. 3.98). Most often, and at least for small deformations, say less than a few percent, the hardening is effectively linear. Under cyclic loading the alloys concerned show a strong *Bauschinger effect*, as in Fig. 3.100. Thus we can say that work-hardening in these heterogeneous materials is essentially kinematic and linear, representing a particular case of what is called the *Prager model* of plasticity (Sect. 3.5.2.4). At greater deformations it is seen to become parabolic and of a more isotropic nature.

The kinematic nature of the work-hardening of these materials can be explained quite simply in physical terms, qualitatively at least; the explanation emphasises the *reversible aspect of microscopic-scale plasticity* in such conditions.

Consider the hysteresis loop (b) of Fig. 3.100. The first effect of the shear stress τ is the by-passing of the particles in state A and the formation of Orowan loops (Fig. 3.101). Using a rheological model of parallel spring-and-friction-element, the friction element is released at the Orowan stress τ_{OR} and from A to B, on the first

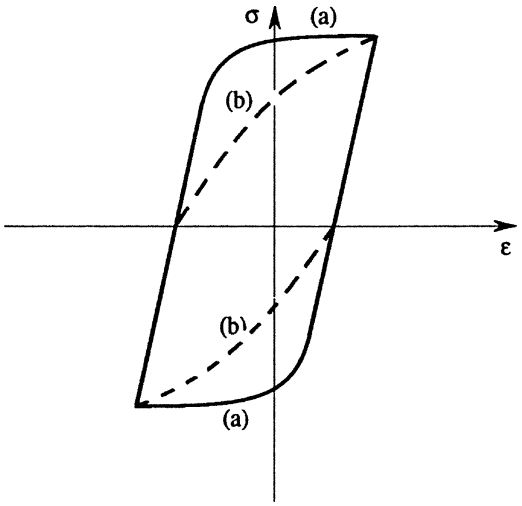


Fig. 3.100 Hysteresis stress-strain loop in an alloy with precipitates (**a**) sheared; (**b**) by-passed by dislocations

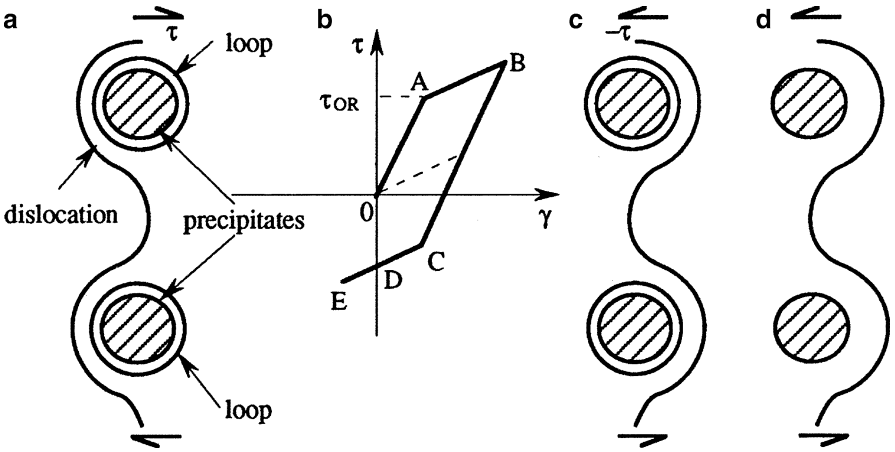


Fig. 3.101 Interaction of dislocations and precipitates with loops which vanish when the dislocations return (steps c and d). Associated hysteresis loop

half of the hysteresis loop, dislocation loops build up around the particles, which are subjected not only to the applied stress but also to that exerted by these loops. Thus internal stresses develop, corresponding in the model to the tension in the spring. To a first approximation the Orowan stress, corresponding to isotropic hardening, does not change, since the effective size of the particles is affected very little by the surrounding loops, which remain very close to the particles. When the direction of the stress is reversed, as in Fig. 3.101c, d, this internal stress is steadily reduced and the density of the loops falls as the unpinned dislocation returns to make contact with the particle. The geometry is such that there is an attraction between the loops

and the dislocation returning towards the particle, since the two parts have opposite signs. At point D of the hysteresis loop, the dislocation loops corresponding to the outward journey (AB) have vanished and the initial state is re-established. New dislocation loops are then formed, corresponding to DE, and the process is repeated.

If such reversibility or such memory is to be significant the dislocation loops must remain in one plane, the more so the less the deformation and the stacking fault energy. Otherwise the internal stresses which accompany directly the applied macroscopic deformation will be relaxed by the mechanisms of cross-slipping, as indicated schematically in Fig. 3.96. In this way the "perfect memory" effect is lost.

(g) *Elastic inclusion model of work-hardening*

Means for a more quantitative modelling of the problem are provided by the mechanics of inclusions. Several models have been suggested, all based on the concepts developed in the treatment of inclusions in Chap. 2 (Sect. 2.7). In explaining them here we shall confine our treatment to the case of pure plastic shear, γ_{p}.

The strain incompatibility between the inclusion and the matrix, that is the eigenstrain $\underline{\underline{e}}^{\mathrm{F}}$, reduces to $e_{13}^{\mathrm{F}} = -\gamma_p/2$. Thus, although there are shear loops around the particles (Fig. 3.101a), the behaviour can be analysed in essentially elastic terms. In this sense we can speak of dealing with a solid having *perfect memory*, the loops around the precipitates providing the quantitative expression of incompatibility in deformation between these and the matrix.

We have shown (Sect. 2.7.2, Eq. 2.159) that in a spherical inclusion subjected to an eigenstrain a homogeneous internal stress $\sigma_{13}(\mathrm{int})$ develops. This is balanced by an external stress field $\sigma_{13}(\mathrm{ext})$, whose mean value we denote by $\bar{\sigma}_{13}$. Thus we have:

$$\sigma_{13}\,(\mathrm{int}) = \frac{1}{15}\frac{7-5\nu}{1-\nu}\mu\gamma_{\mathrm{p}}, \quad \bar{\sigma}_{13}(\mathrm{ext}) = -\frac{1}{15}\frac{7-5\nu}{1-\nu}\mu\frac{f_{\mathrm{v}}}{1-f_{\mathrm{v}}}\gamma_{\mathrm{p}} \tag{3.142}$$

where f_{v} is the volume fraction of the particles.

The distribution of σ_{13} is drawn schematically in Fig. 3.102. This shows that the stresses are very strongly *polarised*, an effect that is at the very origin of kinematic work-hardening (refer to Sect. 3.5.2.4(b) for similar effects). Extending this formalism to the case of tension/compression deformation, and with the same notation as used in Sect. 3.2.2 for macroscopic plasticity, we get the diagram of Fig. 3.103, with $R = \sigma_{\mathrm{OR}}$ (the Orowan stress) and $X \approx \mu f_{\mathrm{v}}\varepsilon^{\mathrm{p}}$.

3.4.3.3 Irradiation Hardening

Bombardment of a crystal with neutrons creates zones in which the atoms have suffered large displacements. Some will be in interstitial positions, leaving compensating vacancies, and such point defects can migrate if the temperature

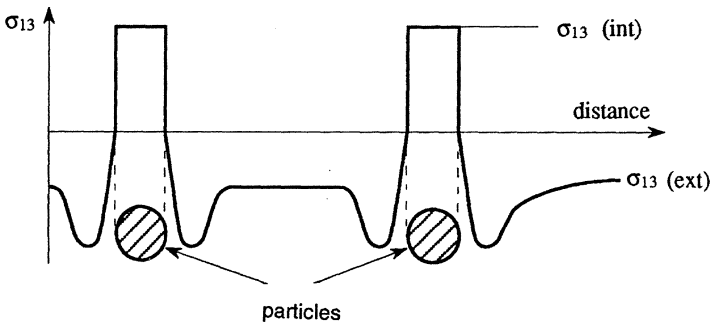


Fig. 3.102 Stress generated by the dislocation loops in the particles (int) and in the matrix (ext)

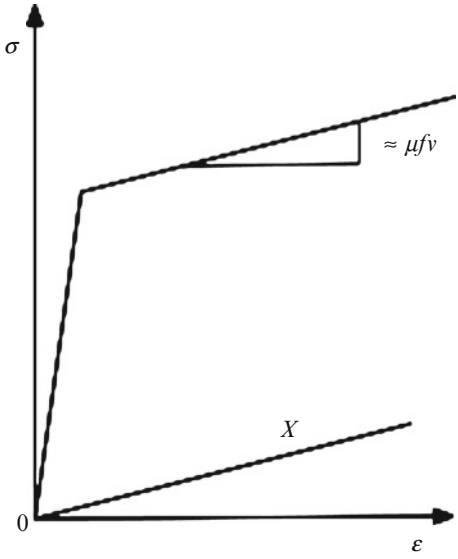


Fig. 3.103 Linear kinematic hardening in an alloy containing precipitates surrounded by dislocation loops

is high enough. As well as some recombinations of interstitials and vacancies, clusterings occur, resulting in di-, tri-, etc. vacancies and finally either micro-voids or loops of dislocations when vacancy discs collapse (Fig. 3.33). Tetrahedra of stacking faults can also be produced. The size of these irradiation products increases when the irradiation temperature is raised.

Dislocation loops appear at irradiation temperatures less than 600°C. The dislocation density increases as dislocation loops are created and also as vacancies super-saturation produces climb of parts of the dislocations; the dislocation density reaches a saturation value of about 10^{14} m^{-2} ($6{\cdot}10^{14}$ for 316 austenitic stainless steel).

Table 3.9 Effect of neutron irradiation on the mechanical properties of various alloys

Alloy	Flux, neutrons/m^2	Irradiation temp, K	R_m MPa	$R_{p0.2}$ MPa	
Austenitic	0		576	235	65
Steel	$12{\cdot}10^{25}$	373	720	632	43
Z6CN18.9					
Mild steel	0		517	276	25
0.2% C	$2{\cdot}10^{23}$	352	676	634	6
	$1{\cdot}10^{24}$	352	800	752	4
	$2{\cdot}10^{23}$	566	703	524	9
	$2{\cdot}10^{23}$	677	579	386	14
Aluminium	0		124	65	28.8
	$1{\cdot}10^{24}$	339	257	177	22.4
Aluminium	0		310	265	17.5
6061 T6	$1{\cdot}10^{24}$	339	349	306	16.2
Zircaloy 2	0		276	155	13
	$1{\cdot}10^{24}$	411	310	279	4

At temperatures higher than 400°C and for doses higher than 20 dpa (displacement per atom), *microvoids* form, resulting in a swelling of the material. This is also linked to transmutations, which result in the formation of helium; these atoms migrate to form micro-bubbles.

Another effect of the super-saturation of vacancies is the acceleration of *precipitation* of foreign atoms in solid solution. In ferritic steels dissolved copper precipitation forms ordered Cu_3Fe precipitates. They contribute to irradiation hardening. In austenitic stainless steels the precipitation of $M_{23}C_6$ and M_6C decreases the concentrations of Ni and C, which destabilises the gamma phase; some ferrite can be formed. Near grain boundaries there is a decrease of Fe and Cr concentrations and an increase of Si, Ni and P concentrations. This is not without influence on embrittlement and on corrosion resistance.

These various irradiation products contribute to irradiation hardening (Table 3.9). Figure 3.104 shows stress-strain recordings at various temperatures and for various doses for a ferritic 9Cr1Mo steel (Deo et al. 2008). The hardening is accompanied by a *decrease of the strain hardening rate,* which results from contributions of vacancies to the by-passing of obstacles by climb. In general, *the yield strength and the hardness are proportional to the square root of the integrated dose,* at least for irradiation times that are not too long (Fig. 3.105).

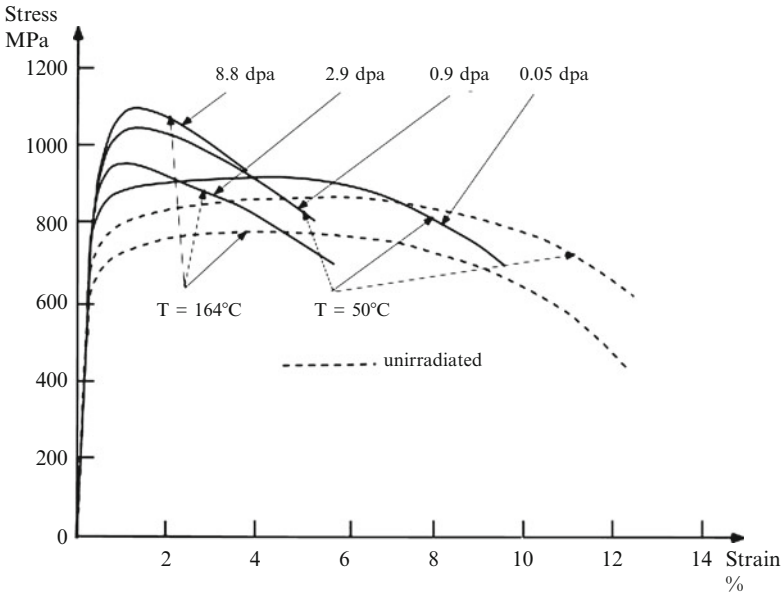
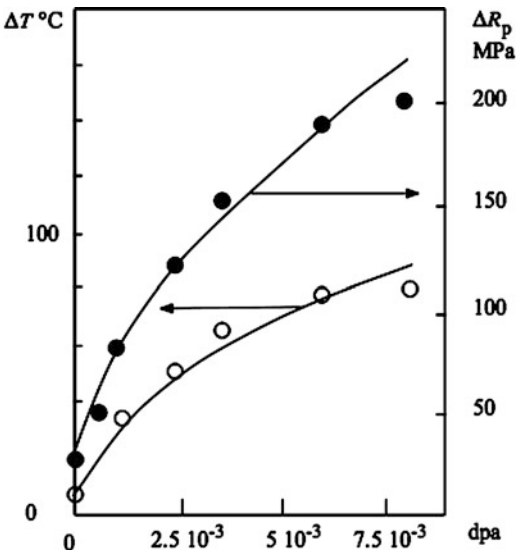


Fig. 3.104 Stress-strain recordings for unirradiated and irradiated ferritic 9Cr1Mo steel (From Deo et al. 2008)

Fig. 3.105 Variations of the yield strength and of the brittle/ductile transition temperature with irradiation dose for carbon steel



3.4.4 Volumetric Shear Deformation: Twinning, Martensitic Transformation, Transformation Plasticity

3.4.4.1 Introduction

So far we have been studying the deformations produced by dislocations glide; we now investigate other mechanisms, twinning of crystals and plasticity generated by phase transformations. They are characterised by the sudden deformation as a whole of part of a crystal. They are not entirely unrelated to dislocation slipping, since, as we shall see, they can be regarded as resulting from the activation of particular dislocation sources.

These volumetric transformations generate shear strains and possible volume changes. Stresses, whether externally applied or resulting from the transformations themselves, alter the free energy of the system and in turn react on the transformation.

We consider first the simple shearing of the crystal lattice resulting from twinning. In Sect. 3.3.9 we described twins and the deformation which they produce, and the important role they play in a number of industrial alloys. We must now understand how they are generated and the critical stress needed. We will then study martensitic transformations; they are closely associated with twinning but in addition to the shearing they involve a volume change and, especially, a free energy of chemical origin. More generally, any transformation is linked with plastic deformation.

3.4.4.2 Development of Twinning

In a single crystal we can imagine that mechanical twinning is generated by a shear strain traversing the cross-section of the sample and then progressively invading the whole body; the resulting deformation was calculated in Sect. 3.3.9 and it is given by (3.74).

Once a twin is nucleated, its growth is easy and quick; it stops when a grain boundary is met, or other obstacles, including other twins. Provided that the distortion remains purely elastic a change in sign of the stress will cause the twin to vanish; and the twinning is then said to be *reversible.* But the strain energy can be great enough to cause plastic deformation, which stabilises the twin and inhibits the reversal.

Another way in which the distortion can be accommodated is by forming other twins. These can belong to the same system, but with different orientations – they belong to other *variants*.

In this section the stress necessary to initiate a mechanical twin is determined first. Then the interactions between twins of different variants are examined in order to explain the anomalously large value of the work-hardening observed in the presence of mechanical twinning. Only FCC crystals are considered.

(a) *Twinning stress*

As explained in Sect. 3.3.9, an elementary small twin can be assimilated to a stacking fault: for example, in FCC structures the twin <112> {111} corresponds to a Shockley partial dislocation crossing a succession of planes, creating a stacking fault in each one (Müllner et al. 1994). Several mechanisms have been suggested to explain the formation of twins; for example the Venables (1964) model for FCC. This considers a tree with Burgers vector (1/2)[110], which crosses the primary slip plane (111). After the passage of many dislocations of this system, the tree acquires a large-size jog which, under the action of the applied stress, can be dissociated thus:

$$(1/2)\,[110] \rightarrow (1/3)\,[111] + (1/6)\left[11\bar{2}\right]$$

that is, into a Frank and a Shockley dislocation. The latter can then form a source, which generates the twin. For this to be possible the stress must be great enough to overcome the stacking fault energy γ_{f}, that is:

$$\tau_{\mathrm{w}} b_{[11\bar{2}]} = \gamma_{\mathrm{f}} \tag{3.143}$$

where τ_{w} includes both the applied shear stress τ and the stress resulting from the dislocation pile-up. A calculation by Venables yielded the following relation:

$$\left(1 - \frac{2m_{\mathrm{s}}}{3\beta} + \frac{(1-\nu)\,L{m_{\mathrm{s}}}^2}{1.84\mu b}\sigma_w\right)\sigma_w = \frac{\gamma_{\mathrm{f}}}{b} \tag{3.144}$$

where m_{s} is the Schmid factor, β a constant with value close to 1 and L the length of the dislocation pile-up.

This result shows that the critical twinning stress varies like $\sqrt{\gamma_{\mathrm{f}}}$, so that twinning becomes easier the lower the stacking fault energy. Figure 3.106 compares the theoretical prediction with experimental results, and shows that the relation is rather well supported by observations on various alloys.

(b) *Crossing of mechanical twins by perfect and partial dislocations*

The crossing of a primary twin by perfect or partial dislocations, which constitute embryos of secondary twins, is examined. Very few authors have addressed this question (see however the recent observations by Efstathiou and Sehitoglu (2010) on single crystals of Hadfield[21] steel). *In situ* TEM observations by Coujou (1987) and by Coujou et al. (1992) have shown that pre-existing micro-twins or stacking

[21]Hadfield steel contains about 1%C with 11–15% Mn. This austenitic steel is unique in that it combines high toughness and ductility with high work-hardening capacity and good resistance to wear. Sir Robert Hadfield invented this steel in 1882.

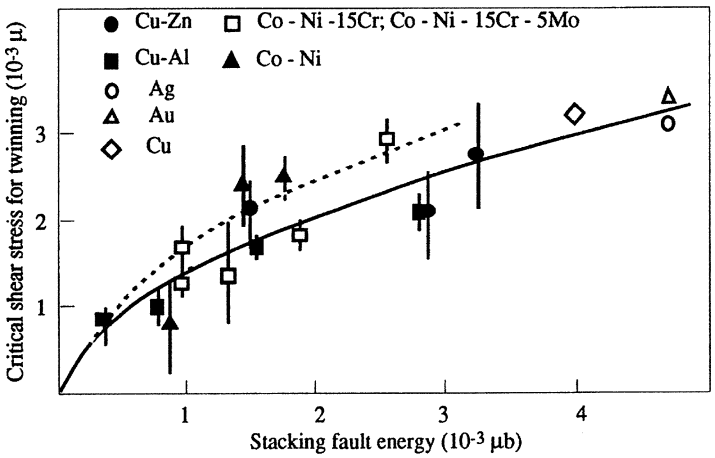


Fig. 3.106 Variation of twinning stress with stacking fault energy: *dotted curve* is Venables relation (3.144); *full curve* results found for CuZn and CuAl alloys and Ag, Au and Cu single crystals

faults can be sheared and can disappear when they are intersected by a secondary twin (Fig. 3.107). This suggests that pre-existing twins, which are special grain boundaries, are shearable obstacles. A number of researchers have examined in more detail the interaction between an annealing twin and perfect or partial dislocations (see *e.g.* Rémy 1978, 1981). Several situations can be distinguished from these observations.

Perfect dislocations

Let us consider two FCC crystals in twinning orientation relationship (Fig. 3.108). The Thompson tetrahedron (Sect. 3.3.4.1) of each crystal is shown in this figure. The situation where the perfect dislocation in the parent crystal is AB is trivial since this dislocation can easily cross-slip in the twin crystal. The situations where the leading dislocation does not belong to the twinning plane are more complex, as shown in Figs. 3.109a, b. The dislocation reactions at the interface can be written as:

$$\begin{aligned} & \mathrm{DC}_\beta \rightarrow \mathrm{CD}'_{\beta'} + 2\delta\mathrm{C} \\ \text{or} \quad & \frac{1}{2}[110] \rightarrow \frac{1}{2}[110]_\mathrm{T} + 2 \times \frac{1}{6}\left[11\bar{2}\right] \\ \text{and} \quad & \mathrm{DC}_\beta \rightarrow \mathrm{AD}'_{\beta'} + \mathrm{B}\delta \\ \text{or} \quad & \frac{1}{2}[110] \rightarrow \frac{1}{2}[101]_\mathrm{T} + \frac{1}{6}\left[\bar{2}\bar{1}\bar{1}\right] \end{aligned} \tag{3.145}$$

Both reactions are energetically unfavourable and lead to a change in the thickness of the primary twin by the propagation of partial dislocations along the twin boundary interface.

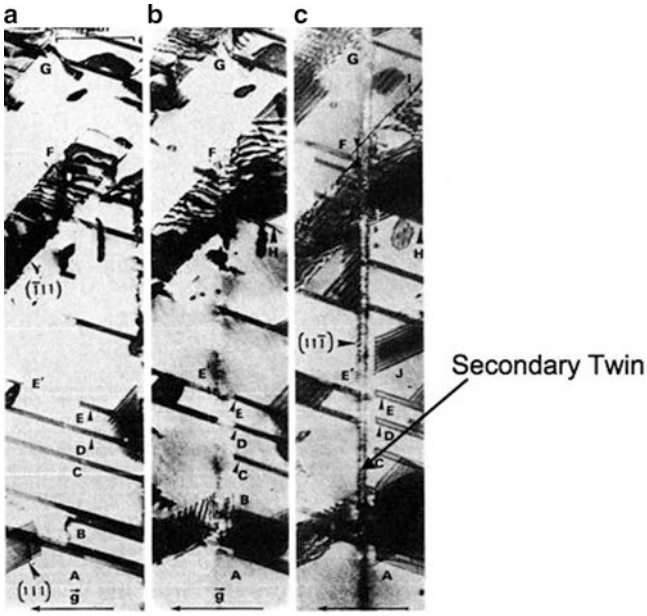


Fig. 3.107 Cu-6.5 at.% Si. Three *in situ* pictures illustrating the interactions between pre-existing stacking faults and micro-twins with another micro-twin nucleated on another variant in the *lower part* of the picture (**a**). This secondary micro-twin propagates from the lower to the *upper part* and shears the stacking faults A, B, C, D, E and a secondary twin in F. Notice that the pre-existing faults and micro-twins can disappear at the intersection with the secondary micro-twin (Coujou 1987)

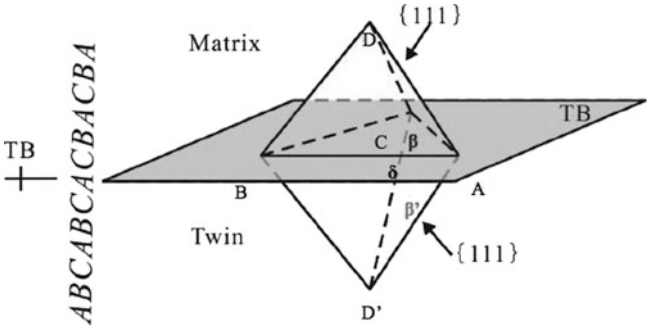


Fig. 3.108 Two FCC crystals in twinning orientation relationship. The Thompson tetrahedron of each crystal is shown. The twin boundary (TB) is indicated

Partial dislocations

Two typical reactions of partial dislocations impinging a twin boundary are shown in Figs. 3.110a, b. They can be written as:

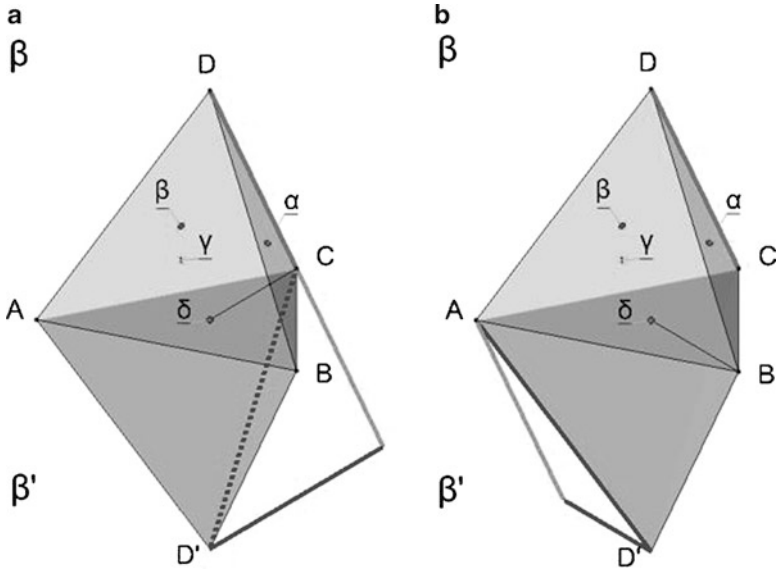


Fig. 3.109 Reaction with a perfect dislocation at the intersection with a twin boundary. (**a**) $DC_\beta \rightarrow CD_\beta + 2\delta C$; (**b**) $DC_\beta \rightarrow AD_\beta + B\delta$

$$
\begin{aligned}
&3D\beta_\beta \rightarrow AD'_{\beta'} + CD'_{\beta'} + 2B\delta \\
\text{or}\quad &3 \times \frac{1}{6}[211] \rightarrow \frac{1}{2}[101]_T + \frac{1}{2}[110]_T + 2 \times \frac{1}{6}[211] \\
\text{and}\quad &3D\beta_\beta \rightarrow CD'_{\beta'} + DC_\beta + B\delta \\
\text{or}\quad &3 \times \frac{1}{6}[211] \rightarrow \frac{1}{2}[110]_T + \frac{1}{2}[110] + \frac{1}{6}\left[2\bar{1}\bar{1}\right]
\end{aligned}
\tag{3.146}
$$

Both reactions are also energetically unfavourable and lead also to a change of the thickness of the primary twin. The applied stress must be sufficient to overcome the obstacle formed by a twin boundary. The crossing of primary twins by secondary twins can produce an extra-hardening effect similar to the Hall-Petch effect (see Sect. 3.4.2.3(a)). Rémy (1978, 1981) has suggested that the flow stress, R_p, can be written as:

$$R_p = R_{pM} + k\frac{1}{(2t)^r}\left(\frac{f}{1-f}\right)^r \tag{3.147}$$

where R_{pM} is the matrix flow stress, k is a constant similar to Hall-Petch constant, t is the thickness of the primary twins and f is the volume fraction of twins, $r = +1$ or $+½$ (the value of $+½$ is similar to the value proposed by Petch, see Sect. 3.4.2.3(a)). More recently, Bouaziz (2001; Bouaziz and Guelton 2001) has proposed a similar expression relating the flow stress to the volume fraction of mechanical twins. This expression uses an evolution law for the dislocation

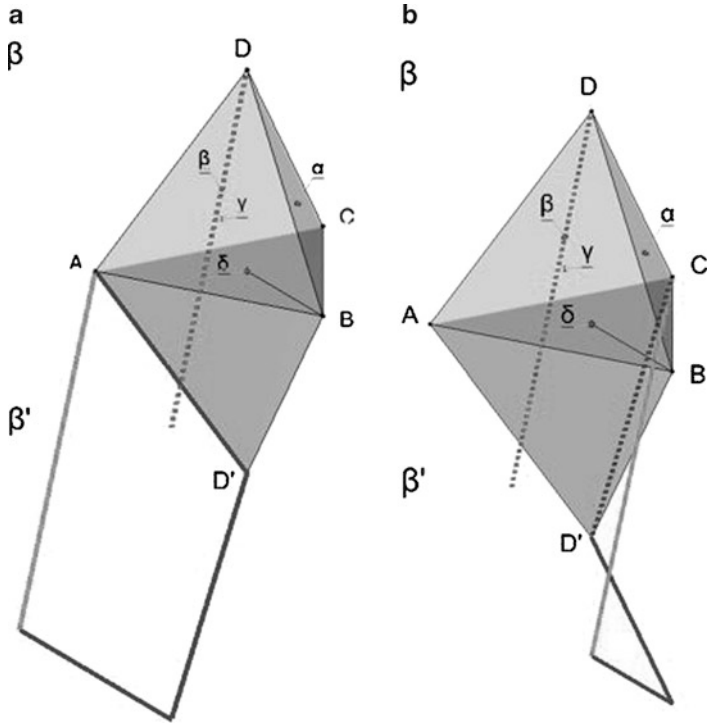


Fig. 3.110 Reaction with a partial dislocation at the intersection with a twin boundary. (**a**) $3D\beta_\beta \rightarrow AD_\beta + CD_\beta + 2B\delta$; (**b**) $3D\beta_\beta \rightarrow CD_\beta + DC_\beta + B\delta$

density ρ_D resulting from the competition between accumulation and annihilation of dislocations due to dynamic recovery (see Sect. 3.4.2.2(h)):

$$\frac{d\rho}{d\varepsilon} = M\left(\frac{1}{bL} + \frac{k}{b}\sqrt{\rho_D} - fr\right)$$

$$\text{and} \quad R_p = \alpha M \mu b \sqrt{\rho_D} \tag{3.148}$$

with M, the average Taylor factor, $L \approx t$ (mean distance between twins), and α a numerical constant of the order of 0.40. The stress-strain curve is obtained by integrating the first equation using a law for the variation of the volume fraction f with strain. It should be noticed that both expressions show the importance of mechanical twinning since the flow stress increases when t decreases, *i.e.*, when the volume fraction of mechanical twins increases.

The technical importance of mechanical twinning for the development of new steels (TWIP (twinning induced plasticity) steel) is highlighted later in Sect. 3.4.5.4.

Further considerations on deformation twinning, including BCC and HCP crystals, can be found in a detailed review by Christian and Mahajan (1995).

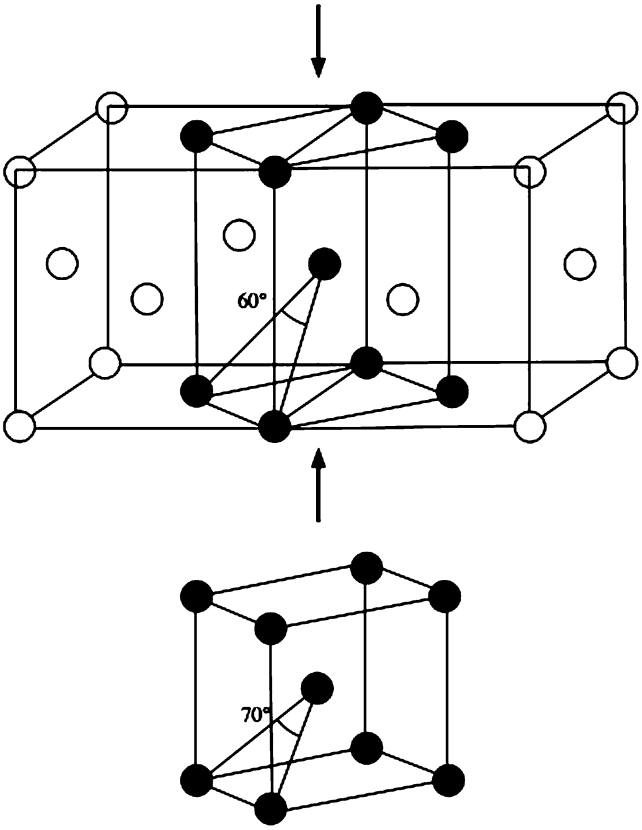


Fig. 3.111 Bain's transformation from FCC austenitic unit cell to body centred tetragonal unit cell of martensite

3.4.4.3 Martensitic Transformation

(a) *Deformation associated with martensitic transformation*
The geometry of the martensitic transformation is almost the same as that for twinning, with the addition of a volume change due to an elongation normal to the habit plane.

An important example is that of the martensitic transformation in austenitic steels. The deformation can be visualised by the Bain's[22] transformation (Fig. 3.111) (Bain 1924). The FCC unit cell contains the BCC unit cell and to obtain the body centred tetragonal unit cell of the martensite a further compression is needed. This is achieved by the development of twins (Fig. 3.112) (Bogers and Burgers 1964).

[22]Edgard C. Bain (1891–1971) was an American metallurgist.

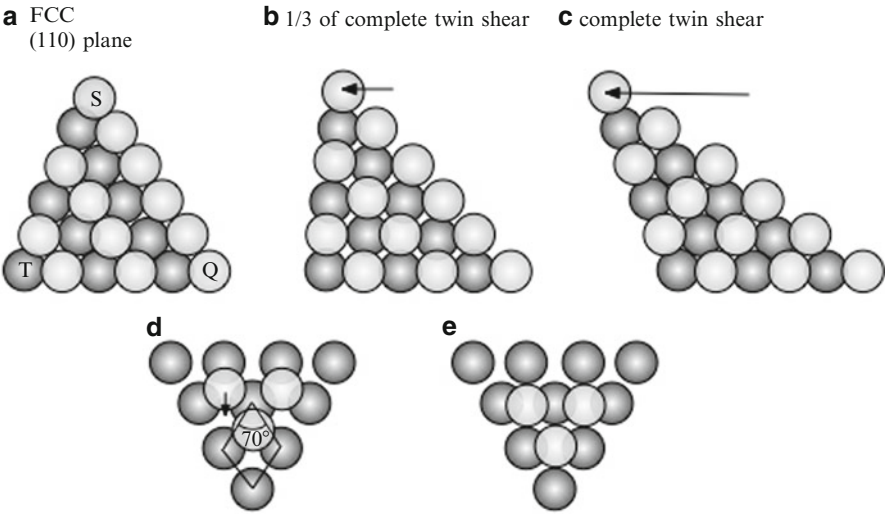


Fig. 3.112 Bogers-Burgers sketch of martensitic transformation: consider a tetrahedron SPQV. (**a**) STQ {110} plane of FCC lattice; (**b**) 1/3 of complete twin shear in QT <112> direction; the atomic rows SP and SV of the former {111} plane are now oriented at 70° from each other; (**c**) complete twin shear; (**d**) SPV plane in case b showing the 70° orientation of SP and SV; (**e**) final shear to achieve the transformation into BCC lattice

(b) *Chemical stress*

The essential change with respect to pure twinning is the need to take account of the free energy of chemical origin, which usually drives the transformation. At equilibrium, in the absence of applied stress the Gibbs free energy G is given by:

$$\Delta G_0 = \Delta H_0 - T_0 \Delta S_0 = 0 \tag{3.149}$$

where ΔH_0, ΔS_0 are the enthalpy and entropy respectively of the transformation and T_0 the equilibrium temperature. At a temperature $T < T_0$ the free energy of the austenite → martensite transformation is:

$$\Delta G_{\mathrm{A}\to\mathrm{M}} = \Delta H_0 - T\Delta S_0 = \Delta S_0 (T_0 - T) \tag{3.150}$$

Numerical data for ferrous alloys are given in Annex 2.

A critical value ΔG_{m} must be reached for the transformation to start at M_{s}, the martensite temperature. On heating the reverse happens. $\Delta G_{\mathrm{A}\to\mathrm{M}} = \Delta G_{\mathrm{m}}$ is given by:

$$\Delta G_{\mathrm{m}} = \Delta S_0 (T_0 - M_{\mathrm{S}}) \tag{3.151}$$

The applied stress changes the Gibbs energy and we can write, formally:

$$\Delta G_{\mathrm{m}} = \Delta S_0 (T_0 - T) + \sigma \Delta \varepsilon \tag{3.152}$$

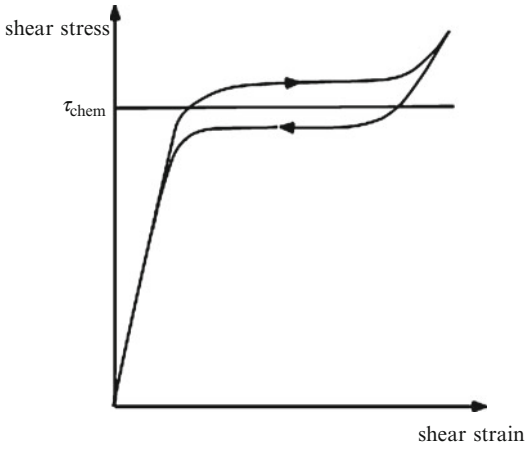


Fig. 3.113 Pseudo-elasticity

In a martensitic transformation the deformation energy comes mainly from shear:

$$\sigma \Delta \varepsilon = \tau \frac{s}{h} \tag{3.153}$$

where s is the shear displacement and h is the thickness of a martensitic plate.

Thus a shear stress favours the martensitic transformation and raises the point M_s; a deformation can destabilise the austenite and give *strain-induced martensite*. It is convenient to associate the chemical free enthalpy with a chemical stress τ_{chem} such that:

$$\tau_{chem} \frac{s}{h} = \Delta S_0 \, (T_0 - T) \tag{3.154}$$

(c) *Pseudo-elasticity*

The formation of martensite can bring about a large deformation at constant stress; referred to as *transformation plasticity*. This is illustrated in Fig. 3.113. If the martensite is completely reversible the inverse transform occurs on unloading at a stress $\tau < \tau_{chem}$ such that:

$$\tau_{chem} - \tau = \frac{\Delta G_m}{s/h} \tag{3.155}$$

The phenomenon is called *pseudo-elasticity*, and is seen, for example, in Cu-Al-Ni and NiTi alloys. Two martensitic transformations have even been observed, in CuZn, CuZnAl, CuAlNi and AuAgCd alloys. An analogous phenomenon, without the martensitic transformation, occurs in AuCd and InTh alloys, called *rubbery elasticity*; this is attributed to the reversible movement of the twin boundaries in the martensite.

(d) *Shape memory alloys*

In cuprous alloys especially a thermo-elastic, that is reversible, martensite forms at a temperature below M_s, consisting of self-accommodating variants with different orientations. The deformation of this is brought about by movements of the boundaries between the variants, so that one grows at the expense of others, possibly ending with a single crystal. If this crystal is now raised to a temperature above A_f the martensite vanishes and the sample returns to its original shape, so the alloy has been given a *shape memory*. Applications have been found for this, for example in the NiTi alloy Nitinol, for which M_s and M_f are on opposite sides of ambient temperature.

For certain alloys, for example CuZnAl, the unstressed martensitic transformation is accompanied by a change of shape resulting from the formation of a single variant which vanishes on reheating: these are said to have *reversible shape memory*. They can be "educated" by cycling the deformation, either in superelasticity or in conditions of non-reversible memory; the effect of this is to favour one variant of the transformation and gradually replace the imposed deformation by the transformation deformation.

3.4.4.4 Transformation Plasticity

In general, a transformation that is accompanied by a change in volume brings about a plastic deformation of the softest phase. Zirconium, for example, is transformed on heating to a β phase of lower resistance that occupies a smaller volume. Therefore, as the transformation front moves through the material, the high-temperature phase must contract: the outer radius R_{20} of a zirconium tube decreases (Fig. 3.114). With the inverse transformation the dilatation of the α phase that appears on cooling produces a decrease of the inner radius R_{10}. Back to room temperature, both radii become smaller. Thus if the tube is taken repeatedly through the transition temperature its diameter will be steadily shortened. This effect is called *Transformation induced plasticity* (TRIP).

Usually there will be several fronts moving in different directions, so that the deformations will not simply add, and a residual stress field will be generated. But if, while this is taking place, the sample is subjected to an external stress, even if small, the deformation in the direction of that stress will be favoured.

There is no really satisfactory model for transformation plasticity. Greenwood and Johnson (1965) have suggested one that gives a deformation proportional to the applied stress and to the volume change, but the basis for this is doubtful. More recently their relation has been generalised by Leblond et al. (1989), to give

$$\mathrm{d}\varepsilon_{ij}^{\mathrm{TP}} = \frac{5}{2}\frac{\Delta V}{V}\frac{s_{ij}}{R_{\mathrm{p}}}(1-m)\,\mathrm{d}m \tag{3.156}$$

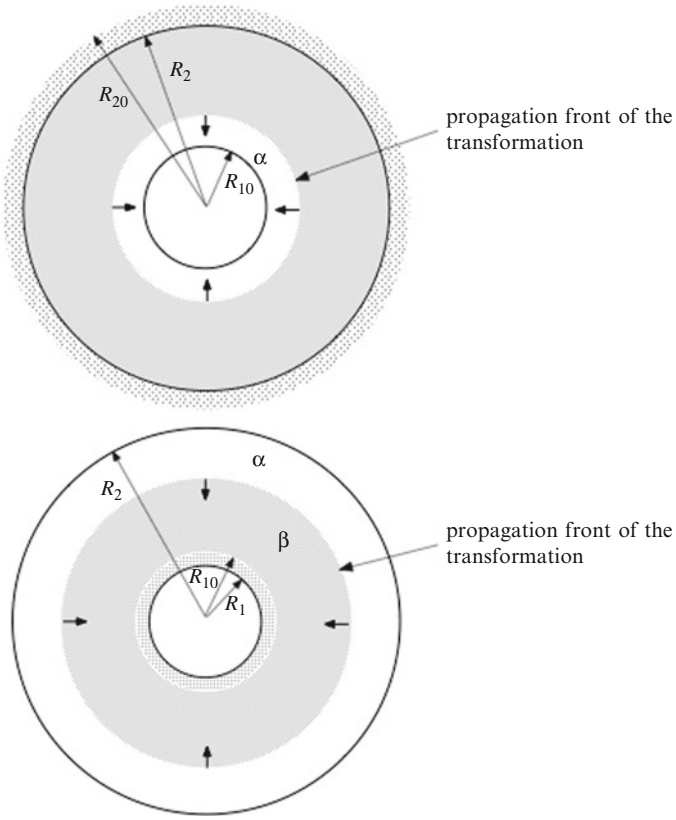


Fig. 3.114 Transformation induced plasticity: The high temperature β phase only is deformed plastically to accommodate a change in volume

where $d\varepsilon_{ij}^{PT}$ is the transformation deformation, $\Delta V/V$ is the relative change of volume, R_p is the yield strength of the mother phase (austenite), m the fraction transformed and s_{ij} is the deviatoric stress tensor.

This relation holds well when the origin of the transform plasticity is the volume change associated with the change of phase, and also, but not only, for structural transformations controlled by diffusion. Thus we may say that source of measured macroscopic transformation plasticity is the *orientation* of the plasticity that starts locally in the mother phase in the neighbourhood of the transformed zones.

For phase transformations that take place without diffusion – martensitic transformations – the essential origin of the plasticity is the orientation of transformed zones in response to the applied stress; this is the origin of the *shape memory effect* which is sometimes referred to as the "Magee effect" (Magee 1966). Although relations of the type (3.156) are used in such case, it is doubtful if this use is justified: this is so because there are only very few experimental verifications of relations of

this type for multiaxial loading. However recent studies on TRIP effect in a Cr-Ni-Mo steel have been made under multiaxial (tension + torsion) loading with radial and non-radial paths (Videau et al. 1996). A new model based on the formulation of internal variables (see Sect. 3.5) has been proposed in which it was assumed that the "Greenwood –Johnson" effect is proportional to $(s_{ij} - \alpha_{ij})$ instead of simply s_{ij}, where α_{ij} is a backstress tensor associated with kinematic hardening (Fisher et al. 2000).

3.4.5 An Application of Hardening: Strengthening of Steels

So far we have been considering different mechanisms of hardening, and in Table 3.10 we give some details for various alloys. In the next paragraph we indicate how the principles can be applied in the case of steels.

3.4.5.1 Micro-Alloyed Steels

The development of micro-alloyed steels, to replace mild steels, was made possible largely by a better understanding of the phenomenon of hardening. In these alloys the reduced grain size (refer to Sect. 3.4.2.3) is combined with the presence of precipitates (refer to Sect. 3.4.3.2) of carbides, produced by adding small amounts (0.1%) of vanadium and niobium. The grain size is lowered by starting with austenite of small grain size and lowering the $\gamma \rightarrow \alpha$ transformation temperature as much as possible. Controlled rolling, which means particular precise conditions of temperature and deformation, reduces the grain size in the austenite by inhibiting recrystallisation through precipitation of carbides and carbo-nitrides of Nb and V. Moreover, these precipitates themselves contribute to the hardening. By this means yield strengths of the order of 600 MPa can be achieved (Fig. 3.115).

3.4.5.2 Quenched and Tempered Steels

Steels are greatly hardened by the martensite transformation (refer to Annex 2) that results from quenching; a number of factors combine to produce this effect:

– First, the martensite platelets are very small, the more so the smaller the grain size of the original austenite; the range is from a few microns to a few tens. Further, they display a large amount of twinning and the twin boundaries play the same role in the movement of dislocations as do grain boundaries (refer to Sect. 3.4.2.3).
– Second, the heavy super-saturation with carbon results in a large solid solution hardening (refer to Sect. 3.4.3.1(c))

Table 3.10 Hardening mechanisms for various alloys

Alloy	Mechanism	Yield strength (MPa)
Aluminium		
AlCuMg heat-treated, work-hardened 2024 T3	Work-hardening, precipitation	345
AlZnMg annealed 7075 O	Solid solution hardening	100
Heat-treated 7075 T6	Precipitation	500
Copper		
Electrolytic Cu annealed (0.04%O)		70
OFHC Cu, work hardened	Work-hardening	275
CuZn α brass, annealed	Solid solution hardening	240
CuBe heat treated	Precipitation	965
Nickel		
NiCrFe (Inconel) work hardened	Work-hardening, solid solution hardening	1,035
NiMoFe (Hastelloy) heat treated	Solid solution hardening	275
NiCoCrMoTiAl (Udimet 700) heat treated	Precipitation	875 at 650°C
Steels		
Mild steel (0.01%C) annealed	Solid solution hardening	170
Steel (0.2%C) annealed	Solid solution hardening	500
FeNiMoMnCrC heat treated	Solid solution hardening	1,500
Maraging	Precipitation, grain refinement	2,000
Dual phase (DP) steel	Martensite hardening	600–1,000
TRIP steel	Martensite hardening	600–1,000
TWIP steel	Twin hardening	1,800 (UTS)

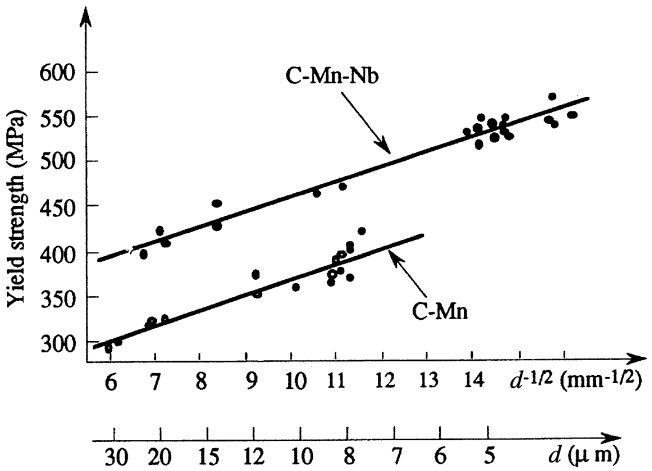


Fig. 3.115 Variation of the yield strength of micro-alloyed steels with the grain size

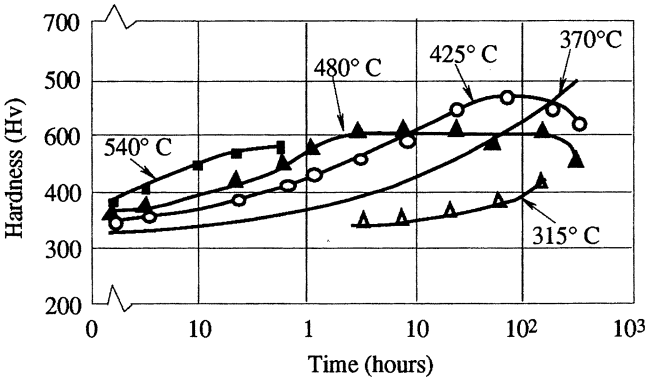


Fig. 3.116 Evolution of hardness of a maraging steel (18% Ni) as a function of the tempering time at various temperatures

– Third, during tempering the carbon atoms rearrange themselves and finally precipitate in the form of carbides, the precipitation contributing to the hardness (see Sect. 3.4.3.2).
– Fourth and last, the number of dislocations increases greatly during the transformation, this again contributing greatly to the hardness (refer to Sect. 3.4.2.2). Altogether, yield strengths of the order of 1,600 MPa can be achieved.

3.4.5.3 Maraging Steels

These steels contain very little carbon, and enough nickel, about 18%, to stabilise the austenite. Cooling below M_S gives a martensite with fairly low hardness that can be deformed plastically to multiply the dislocations; precipitating the intermetallic compounds Fe_2Mo and Ni_3Ti at around 490°C then pins the dislocations in the martensitic fine structure. The combination of small grain size (refer to Sect. 3.4.2.3) and abundant and very fine precipitates (refer to Sect. 3.4.3.2) results in a hardness that can take the yield strength up to 2,400 or even 3,000 MPa (Fig. 3.116).

Very recently stainless maraging steels have been developed in particular for aeronautical industry and the fabrication of a number of components for landing gears. Traditionally these components were protected against corrosion by electrolytic deposition of cadmium. As the new European community regulations on chemicals (*REACH: Registration, Evaluation, Authorisation and Restriction of Chemical Substances*) have prohibited the use of cadmium, the steelmakers have developed new maraging steels containing at least 12% chromium. These high strength martensitic stainless steels can reach a yield strength larger than 1,500 MPa.

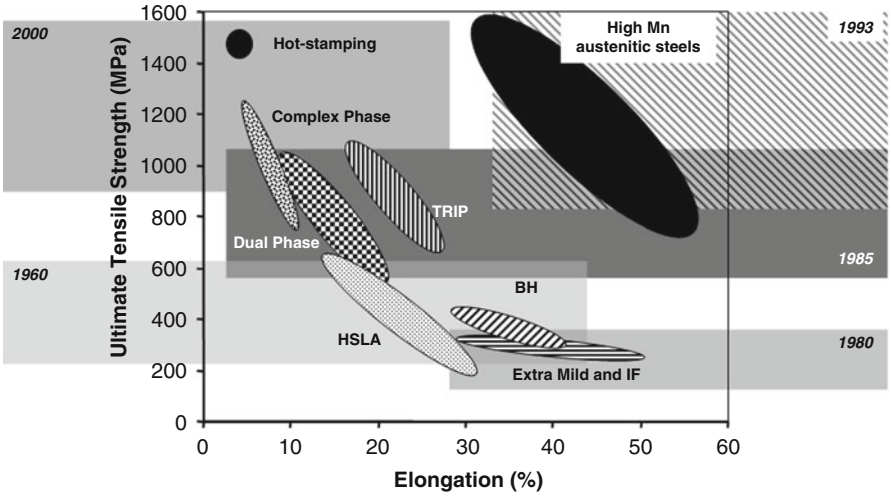


Fig. 3.117 Tensile strength (UTS) as a function of elongation in modern high strength steel sheets for automotive applications. The years correspond to the approximate beginning of industrial development by steelmakers (Bouaziz et al. 2011)

3.4.5.4 High Strength Steels for Automotive Applications

In this paragraph a brief illustration of the "material tetrahedron" shown in Fig. 1.1 of Chap. 1 is given for steels used in cars. In the automotive industry many studies have been devoted to the development of new steels for reducing the fuel consumption and improving the crashworthiness of vehicles. The situation over the last past two decades is summarised in Fig. 3.117, where the years indicate those corresponding to the real development of these new steels by steelmakers. This figure illustrates the compromise between the strength and the ductility, which must be found in the design of a car body in white. The high strength low alloy (HSLA) steels correspond to the family of microalloyed steels obtained by thermomechanical controlled processing (TMCP), as indicated previously in Sect. 3.4.5.1. The TMCP principle is shown in Fig. 3.118 where the phenomena of static and dynamic recrystallisation (see Sect. 3.4.2.2(j)) are indicated. The position of the last rolling pass, before phase transformation takes place during accelerated cooling, is of critical importance.

In Fig. 3.117 dual phase (DP) steels are new materials containing a mixture of martensite islands in a ferrite + bainite matrix. The materials can be produced in several ways. One of them includes a further heat treatment after rolling. The material is heated between AC_1 ($\alpha \rightarrow \gamma$ start) and AC_3 ($\alpha \rightarrow \gamma$ finish) (see Annex 2). The preferential partitioning of carbon atoms in the austenite phase in the dual phase domain allows for the formation of martensite (+ eventually residual

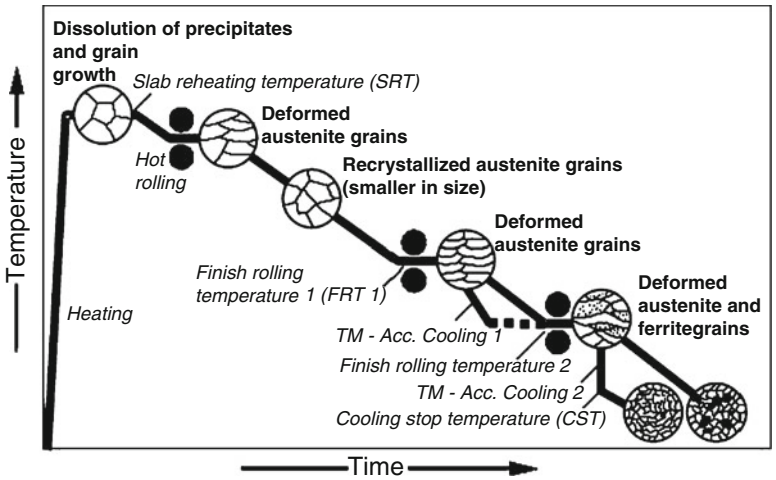


Fig. 3.118 Thermo-mechanical controlled processing (TMCP) of high strength steels

austenite) during controlled cooling. A micrograph illustrating the microstructure of a DP steel was shown in Chap. 1 (Fig. 1.60). These DP steels exhibit much better properties than those of mild steels. However their performance in terms of weldability and sensibility to shear cutting during sheet preparation must be controlled. Moreover the composition and the microstructure of these steels must be such that bath (also called Hot Dip) (or electrolytic) galvanisation can be easily made. During painting preparation, the car body in white is heated up to about 180°C. This temperature is sufficient to harden the parts of the components, which have been plastically deformed. The mechanism responsible for bake hardening[23] (BH) is the formation of Cottrell atmospheres (Sect. 3.4.3.1(d)).

Figure 3.117 shows two other families of high strength steels, which are the transformation induced plasticity (TRIP) steels and the twinning induced plasticity (TWIP High Mn Austenitic steels) steels. A brief description of the strengthening mechanisms of these new steels is given below.

The mechanical properties of TRIP steel is largely explained by the TRIP effect which was briefly introduced in Sect. 3.4.4.4. These materials contain a certain amount of Mn + Si or Mn + Al (with a sum close to 3–4%). They have been given a special thermo-mechanical treatment to develop a mixture of austenite particles in a ferrite + bainite matrix. The austenite particles must have a composition such that they transform to martensite during deformation. This phase transformation strongly enhances the work-hardening rate of these steels, which produces an increase of the

[23]Body in white (BIW) refers to the stage, in automobile manufacturing, in which the car body sheet metal components have been welded together, but before moving parts (doors, hoods and deck lids as well as fenders) the motor and chassis assemblies have been added before painting.

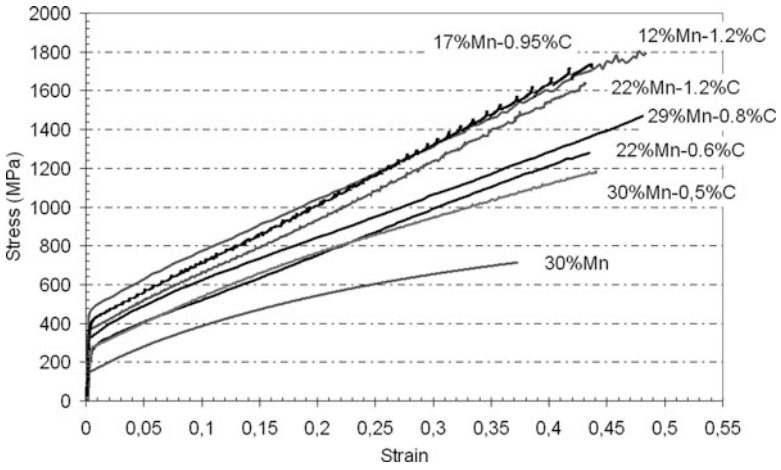


Fig. 3.119 Tensile stress-strain curves of TWIP steels with a wide range of carbon and manganese contents (Bouaziz et al. 2011)

tensile elongation (see Considère relation, Chap. 1, Eq. 1.8). These TRIP steels must also be welded by resistance spot welding, and must be easily galvanised to protect them against corrosion. Bath (or Hot Dip) galvanisation may be a problem in high Si steel. This is why for these steels electrolytic galvanisation should be preferred, but this also depends on the price of electricity!

TWIP steels have been introduced more recently although the principle of hardening mechanism by mechanical twinning in high manganese steels was known since a long time (see *e.g.* the studies devoted to Hadfield steels and Fe-18Mn-5Cr steels). These materials contain a high Mn content (15–25%) and a relatively high carbon content (0.3–0.6%). Figure 3.117 indicates that these new steels can satisfy a better compromise between strength and elongation. These steels exhibit an extremely large value of the work-hardening rate as illustrated in Fig. 3.119. These steels have an initial FCC structure. The value of the work-hardening rate observed in Fig. 3.119 is of the order of $\mu/25$ and is thus much larger than the value of $\mu/200$ measured in more conventional crystals which deform by dislocation glide (see Sect. 3.4). The explanation for this anomalous (but beneficial!) work-hardening behaviour lies in the propensity of these FCC alloys to be deformed by the propagation of mechanical microtwins, as observed in Fig. 3.120. In some circumstances a strain-induced phase transformation γ FCC $\rightarrow$ ε HCP can occur. Several models have been proposed to account for the impressive work-hardening rate of TWIP steels. In a number of these models it is assumed that a pre-existing twin acts as an impenetrable grain boundary, which is not reflecting the true situation as underlined previously in Sect. 3.4.4.2. Moreover there is very little quantitative information on the volume fraction occupied by the mechanical microtwins, which complicates the situation. Therefore the detailed explanation of the work-hardening

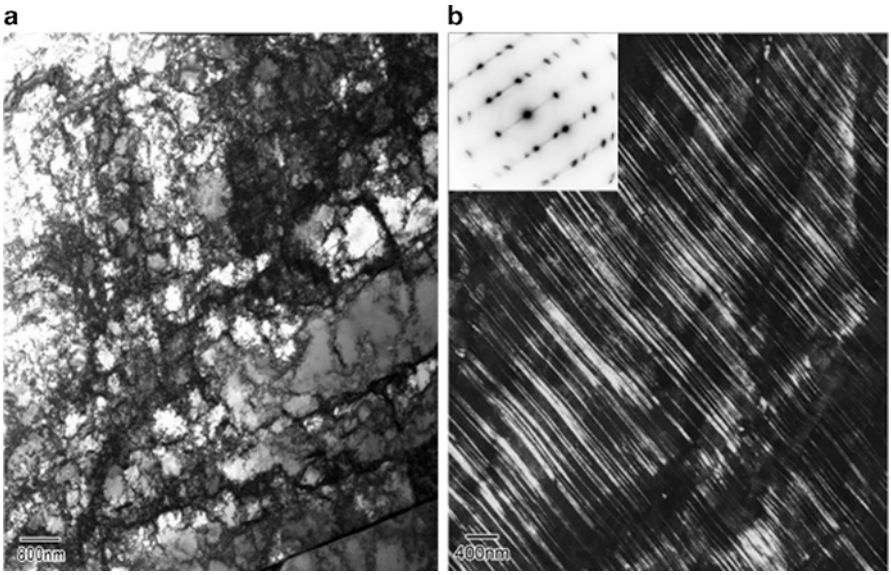


Fig. 3.120 (**a**) Bright field image of Fe-30Mn alloy deformed in tension (20%) showing well developed dislocation cell structure; (**b**) Dark field image of Fe-26Mn-0.6 C steel deformed in tension (50%) showing extensive mechanical twinning (Bouaziz et al. 2011)

of these materials is left to the experts. It should be added that the welding capability of these materials can also be a problem. Moreover the development of these high Mn steels depends largely on the price of manganese.

3.4.6 Fibres Reinforcement

3.4.6.1 Introduction

The number of composite materials incorporating fibres keeps increasing. The number of combinations of fibres and matrices is unlimited. They afford the possibility to design materials for specific needs. What strikes one first is the mechanical strength of certain fibres, and especially the ratio of this to their density: glass fibres, fibres of organic compounds (Kevlar), carbon fibres including nanotubes and also whiskers (Al_2O_3, SiC) have many attractions when there is a need to combine high strength with light weight. The problem is how to transfer the load to the fibres, and the only way to do this is through the intermediary of a matrix.

Table 3.11 A classification of some fibres reinforced composites

Fibre	Brittle	Ductile
Matrix		
Brittle	C-C, C-SiC, SiC-SiC_w	Al_2O_3-Ni
	SiC-Al_2O_3, SiC-BN	
	Glass-epoxy	
	Carbon-epoxy	Reinforced concrete (steel, cast iron)
	Cement-nanotubes	
Ductile	Al-C	–
	Al-SiC	

Some materials, polymers for example, combine the desirable properties of low density, good resistance to environmental conditions, ease of fabrication and modest cost with poor mechanical strength, and it is tempting to seek to reinforce these with fibres. As matrices they can be classified as ductile, for which the need is to increase the strength, or brittle, for which the need is to improve the fracture elongation, typically fibre-reinforced concrete. Table 3.11 gives some examples.

Clearly, incorporating fibres makes a material anisotropic. Even with short fibres it is difficult to ensure that the manufacturing process does not introduce a preferential orientation of the reinforcement. Users of composites aim to take as great advantage as possible of the anisotropies by orienting the fibres in the direction of the maximum principal stresses, as nature does in such materials as bone and wood. For the many applications in which the loading can be very varied the solution is to combine a number of unidirectional composites so as to give a product in which the properties are more or less homogeneous.

In the following, we give some elementary conditions in designing fibres reinforced composites (FRC) limited to ductile matrix composites.

3.4.6.2 Critical Volume Fraction

Fibres with a high elastic modulus can improve the elastic properties of a composite. A matrix with a low yield strength can be reinforced with high-strength fibres, such as carbon, boron or Kevlar. If the material is loaded in the direction of the fibres the overall stress can be expressed as the mean of the separate stresses on the fibres σ_f, and on the matrix σ_m:

$$\sigma = f\sigma_f + (1 - f)\sigma_m \tag{3.157}$$

where f is the volume fraction of fibres.

In general, the fibres are brittle with a fracture strength $R_m{}^f$ and the matrix is ductile and breaks by plastic instability at the tensile strength $R_m{}^m$; thus the maximum stress that the composite can support is the maximum of either

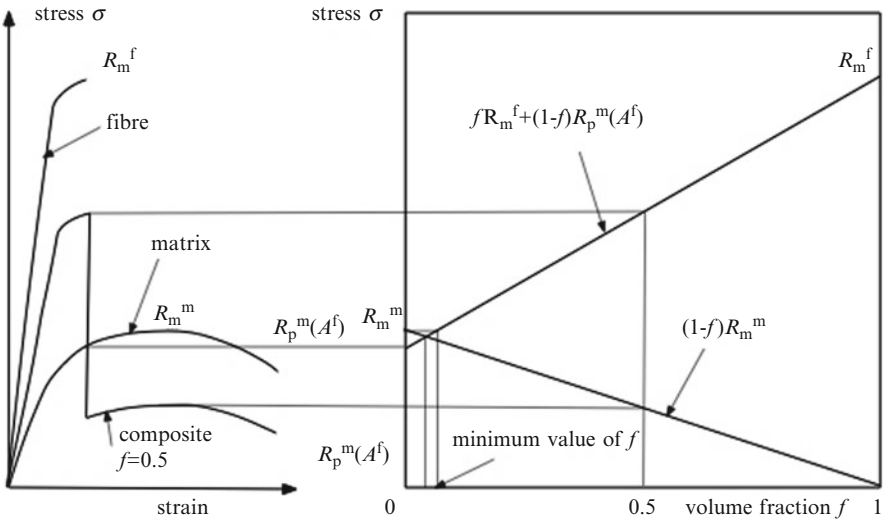


Fig. 3.121 Reinforcement as a function of the volume fraction f of fibres. On the *left* is shown the stress-strain curves of the fibres and of the matrix, and, from relation (3.142), that of the composite

$fR_m^f + (1-f)R_p^m(A^f)$, $R_p^m(A^f)$ being the proof strength of the matrix when the fibres break for an elongation A^f, or of $(1-f)\ R_m^m$. This shows that there is a minimum volume fraction of fibres below which there is no improvement with respect to the tensile strength of the matrix without any reinforcement (Fig. 3.121); it is given by:

$$\frac{f}{1-f} \geq \frac{R_m^m - R_p^m\left(A^f\right)}{R_m^f} \tag{3.158}$$

Thus the critical volumetric fraction increases with the ratio R_m^m/R_m^f and with the work-hardening coefficient of the matrix.

3.4.6.3 Load Transfer

The main purpose of a ductile matrix is to transfer the load to the fibres. Using the *shear lag model* we study the problem first for the case of elastic behaviour, and for simplicity assume that the fibre and the matrix have the same Poisson ratio, that the adhesion between the two is perfect and the load is parallel to the fibres. The geometry is as in Fig. 3.122, with origin at the end of the fibre; u_f is the displacement at a point of abscissa x, u_0 is what the displacement would have been in the absence of the fibre.

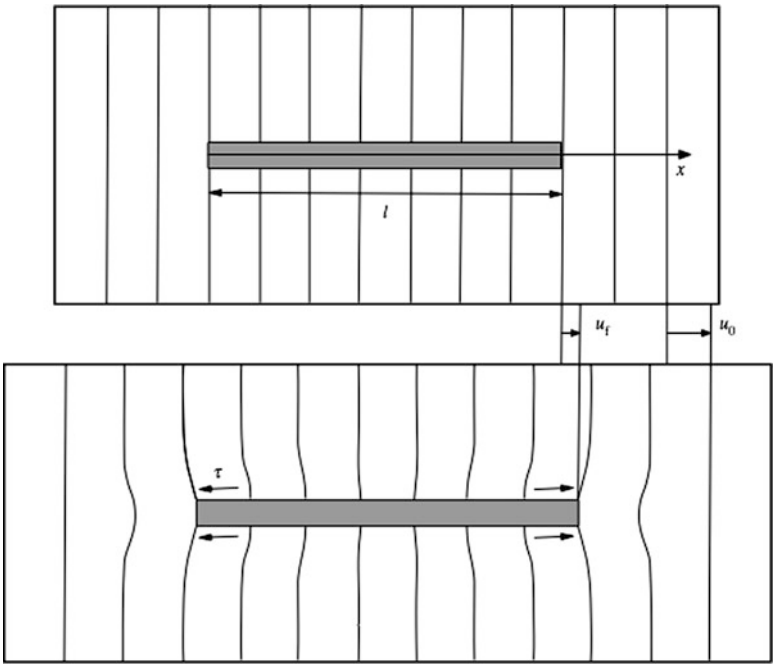


Fig. 3.122 Deformation of a matrix around a fibre of length l (The tensile axis is horizontal)

Consider a slice of length dx. The variation dP/dx of the load in the fibre must be balanced by the shear stress exerted by the matrix along the length dx, which, since the behaviour is elastic, is proportional to $u_f - u_0$. We can write:

$$\frac{dP}{dx} = H\,(u_f - u_0) \tag{3.159}$$

where the constant H, whose value we have to determine, is a function of the geometry and of the elastic moduli of the fibre and the matrix.

Differentiating:

$$\frac{d^2P}{dx^2} = H\left(\frac{du_f}{dx} - \frac{du_0}{dx}\right) = H\left(\frac{P}{S_f E_f} - \varepsilon\right) \tag{3.160}$$

where S_f, E_f are the cross-section and the Young's modulus respectively for the fibre.

Integrating this with the conditions $P = 0$ at $x = 0$ and $x = l$, where l is the length of the fibre, we have for the stress $\sigma_f = P/S_f$ in the fibre:

$$\sigma_f = E_f \varepsilon \left[1 - \frac{\mathrm{ch}\beta\,(l/2 - x)}{\mathrm{ch}\beta\,(l/2)}\right] \tag{3.161}$$

where $\beta = (H/S_f E_f)^{1/2}$

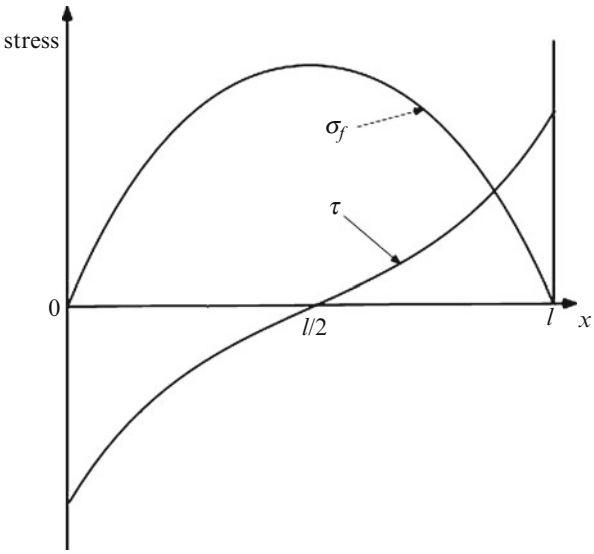


Fig. 3.123 Elastic tensile stress in the fibre and elastic shear stress at the interface

The shear stress τ at the interface is found by writing the equilibrium equation for the slice dx:

$$\frac{\mathrm{d}P}{\mathrm{d}x} = -2\pi r_\mathrm{f}\tau \tag{3.162}$$

where r_f is the radius of the fibre; together with Eq. 3.161 this gives:

$$\tau = -E_\mathrm{f}\varepsilon r_\mathrm{f}\beta\frac{\mathrm{sh}\beta\,(1/2 - x)}{2\mathrm{ch}\beta\,(l/2)} \tag{3.163}$$

Thus the shear stress τ at the surface of the fibre is maximum at the ends of the fibre and zero in the centre, whilst the reverse is true for tensile stress in the fibre σ_f (Fig. 3.123). This shows why the model is called the shear lag model: the matrix is like a lag pulling the fibre by shear on the interface.

We can now determine H. From (3.159) we get

$$H = -\frac{2\pi r_\mathrm{f}\tau}{u_\mathrm{f} - u_0} \tag{3.164}$$

If r is the radial distance from the axis of the fibre the equilibrium of the slice dx is:

$$2\pi r\tau(r) = \text{const.} = 2\pi r_f\tau \tag{3.165}$$

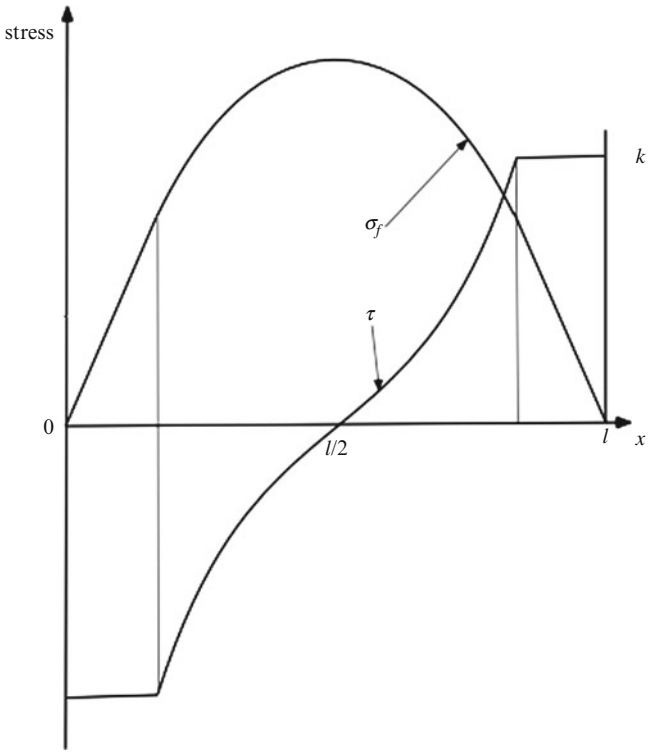


Fig. 3.124 Tensile stress in the fibre and shear stress at the interface in case of plasticisation

from which we find for the displacement u in the matrix:

$$\frac{\mathrm{d}u}{\mathrm{d}r} = \gamma = \frac{\tau(r)}{\mu_\mathrm{m}} = \frac{\tau}{\mu_\mathrm{m}} \frac{r_\mathrm{f}}{r} \tag{3.166}$$

Integration of this gives $(u_\mathrm{f} - u_0)$, and finally

$$H = \frac{2\pi\mu_m}{\log(R/r_\mathrm{f})} \tag{3.167}$$

where the distance between the fibres is $2R$. From this and (3.161) it follows that the parameter β, which determines the increase in load in the fibre, is proportional to $(\mu_\mathrm{m}/E_\mathrm{f})^{1/2}$.

If the deformation ε increases the stress τ at the interface increases up to either the elastic shear strength k_m of the matrix or the resistance of the interface. We now consider how the load transfer occurs with the matrix becoming plastic (Fig. 3.124).

The elastic shear strength of the matrix is reached first at the extremities of the fibre, and the plasticisation then extends towards the middle; when it is complete we have:

$$\sigma_{\mathrm{f}} \pi r_{\mathrm{f}}^2 = 2\pi r_{\mathrm{f}} k_{\mathrm{m}} (l/2) \tag{3.168}$$

In the fibre the stress increases linearly up to the maximum given by the above formula; this is proportional to the length l of the fibre and the longest fibres break. The critical length l_c is:

$$l_{\mathrm{c}} = \frac{r_{\mathrm{f}} R_{\mathrm{m}}^{\mathrm{f}}}{k_{\mathrm{m}}} = d_{\mathrm{f}} \frac{R_{\mathrm{m}}^{\mathrm{f}}}{R_{\mathrm{p}}^{\mathrm{m}}} \tag{3.169}$$

where d_{f} is the diameter of the fibres, $R_{\mathrm{m}}{}^{\mathrm{f}}$ their tensile strength and $R_{\mathrm{p}}{}^{\mathrm{m}}$ the yield strength of the matrix. If the fibres are shorter than this the composite cannot load them to their full capacity.

3.5 Macroscopic Formulation of Plastic Behaviour

3.5.1 From Microscopic to Macroscopic Plasticity

3.5.1.1 Outline

The results we have obtained so far have concerned plastic behaviour at the microscopic scale; the question now arises of how to derive from these a description of the behaviour on the macroscopic scale which will take into account the basic physics of the structural mechanisms of plastic flow in such a way as to enable us to use the tools of continuum mechanics to calculate the properties of large-scale components and structures. Such a transition from the microscopic to the macroscopic in plasticity has been the object of worldwide research for more than 50 years, and it has to be admitted that the problem is still far from having been solved. We shall not, therefore, attempt to treat the subject in depth; the more limited aim of this section is first to describe the main problems that arise and to indicate the nature of the efforts being made to solve them, and then to give the basic physics underlying various aspects of macroscopic plasticity. At the end we give just a sketch of the aim of current research to develop a more deductive treatment of the transition from micro- to macro-plasticity.[24]

[24]This section does not consider plastic deformation related to twinning or to phase transformations, which was studied in Sect. 3.4.4.

As we have seen in Sect. 3.2, the requirements of a macroscopic description of plastic behaviour can be summarised in the form of three types of problem and question

1. The need to define, for a material volume element, a plastic yield criterion f. For what states of stress, specified by the tensor $\underline{\underline{\sigma}}$, will the response be purely elastic? For what states will there be a possibility of plastic flow? This is the problem of finding a scalar function $f(\underline{\underline{\sigma}})$, the *yield function*, such that, by convention, the response remains elastic (or even rigid) as long as $f(\underline{\underline{\sigma}}) < 0$ and the possibility of plastic flow arises only when $f(\underline{\underline{\sigma}}) = 0$.
2. The need for a description of work-hardening and its evolution. For any particular volume element the plasticity criterion does not remain unchanged during plastic flow and the yield function does not depend solely on the state of stress. It is necessary to specify the dependence of the yield function on the internal state of the material. Which work-hardening parameters Y_m should be included? How do these parameters themselves change during plastic flow? What is the form of this extended yield function $f(\underline{\underline{\sigma}}, Y_m)$? What are the laws that express the evolution of the parameters Y_m, that is, the form of g where $dY_m/dt = g_m(\underline{\underline{\sigma}}, Y_n)$?
3. The need for a formulation of the laws of plastic flow, enabling the deformation in response to any stress history or loading path to be determined.

The answers must apply to a volume element whose characteristic dimensions, if it is to be considered as representative of a continuous medium, must be large compared to those of any internal discontinuities. This means that for a crystalline material we must be thinking in terms of a polycrystalline aggregate containing a large number of grains. If so, can we still predict the behaviour of the aggregate as a whole from what we know of the crystal defects? Can we argue in terms of sets of dislocations? Shall we need to describe all the grain boundaries? All the grains?

3.5.1.2 Single Crystals

To break the difficulties into parts of manageable size we start by considering a single crystal and attempt, with the help of the elementary mechanisms of plasticity, to throw some light on the way in which these questions of criteria, work-hardening and flow laws arise.

(a) *The criterion.* We already know that plastic flow is brought about by the multiplication of dislocations and the collective irreversible movement of a sufficient number of these. Whatever mechanism is assumed (multiplication, crossing of obstacles, ...), for this to occur there must be a great enough force exerted on the dislocations.

Consider a dislocation with Burgers vector $\underline{b}$ that remains in its slip plane P, the unit normal to which is $\underline{n}$. Let $\underline{l}$ be the unit tangent vector to the dislocation at the point M (Fig. 3.38). The local stress field $\underline{\underline{\sigma}}$ produces a force $\underline{f}$ per unit length acting on the dislocation, given by the Peach-Köhler relation (3.29): $\underline{f} = (\underline{b}.\underline{\underline{\sigma}}) \wedge \underline{l}$.

The effect of this force on the dislocation is determined by its component f_r in the plane P; if $\underline{\nu}$ is the unit normal to $\underline{l}$ in P, so $\underline{l} \wedge \underline{\nu} = \underline{n}$; then since $\underline{f}$ is normal to $\underline{l}$

$$f_r = \underline{f}.\underline{\nu} = ((\underline{b}.\underline{\underline{\sigma}}) \wedge \underline{l}).\underline{\nu} = (\underline{b}.\underline{\underline{\sigma}}).(\underline{l} \wedge \underline{\nu}) \tag{3.170}$$

Thus

$$f_{\mathrm{r}} = (\underline{b}.\underline{\underline{\sigma}}).(\underline{l} \wedge \underline{\nu}) = \underline{b}.\underline{\underline{\sigma}}.\underline{n} \tag{3.171}$$

If $\underline{m}$ is the unit vector in the direction of $\underline{b}$ (so $\underline{b} = b\underline{m}$), that is, the unit vector in the direction of slip, we have

$$f_{\mathrm{r}} = b\tau_{\mathrm{r}}, \quad \tau_{\mathrm{r}} = \underline{m}.\underline{\underline{\sigma}}.\underline{n} \tag{3.172}$$

where τ_{r} is the resolved shear stress in the direction of slip $\underline{m}$. Thus with $\underline{b}$ given for a family of dislocations we can assume the form of a plasticity criterion expressed as a criterion for a critical resolved shear stress (CRSS):

$$\mathrm{f}(\underline{\underline{\sigma}}) = \left|\underline{m}.\underline{\underline{\sigma}}.\underline{n}\right| - \tau_{\mathrm{c}} \tag{3.173}$$

This is valid for the family of dislocations considered. For the same family we should also consider the possibility of collective slip in the opposite direction, $-\underline{m}$, for which we should need to introduce another and possibly different critical resolved shear stress $\tau_{\mathrm{c}}{}'$. In what follows, it is more convenient to consider the systems $(\underline{n}, \underline{m})$ and $(\underline{n}, -\underline{m})$ as two different systems. More generally, in considering the set of possibilities for plastic slip in a given volume element of a single crystal, we must take account of the complete set of families of dislocations that could slip in their own plane. It is known (see Sect. 3.3) that in general this set is discrete and that very often the number of slip systems is small, that is, the number of planes of normal $\underline{n}^g$ and directions of slip $\underline{m}^g$ in these planes: these *easy-glide systems* are determined by crystallographic and chemical bonding considerations and can be activated at low temperature.

With each easy glide system (g) we have to associate its own criterion $\mathrm{f}^g(\underline{\underline{\sigma}})$ and its own critical shear τ_{c}^g. The crystal cannot experience any plastic activity unless the stress state $\underline{\underline{\sigma}}$ is such as to annul one of the yield functions (or several simultaneously), the rest remaining negative: the condition is

$$\sup_{g \in G} \left(\left|\underline{m}^g.\underline{\underline{\sigma}}.\underline{n}^g\right| - \tau_{\mathrm{c}}^g \right) = 0 \tag{3.174}$$

where G is the set of easy-glide systems. Thus in the stress space the corresponding yield surface will be a (hyper-) polyhedron.

In this we have generalised the Schmid law (Sect. 3.3.2.1) to provide a 3-dimensional criterion for single crystals; the *Schmid factors* of the original law are relevant only for simple tension and compression. We must emphasise again, however, that the law does not apply universally and that even in the case of low-temperature plasticity it can be quite inappropriate for certain materials, in particular those in which the dissociation of the dislocations gives a complex core configuration such that the normal stresses can play an important role, as in BCC crystals (Sect. 3.3.5.1).

In the above expressions the critical shears τ_c^g act as work-hardening parameters; so we must consider how they change during plastic flow. In attempting to answer this question we are led to define other parameters, with a more explicit physical content, on which they depend.

(b) *Work-hardening*. The vital influence on work-hardening of pure metals of the interactions between dislocations makes it natural to attempt to express the critical shears τ_c^g as a function of the dislocation density. The study of the interactions between dislocations (Sect. 3.4.2.2 and Eq. 3.84), gives the relation $\tau_c = \alpha\mu b\sqrt{\rho}_D$ for isotropic work-hardening, where ρ_D is the dislocation density and b the modulus of the Burgers vector. A more detailed analysis, taking account of short-range interactions and the different forms these can take, according to the pair of dislocations considered, gives anisotropic expressions such as

$$\tau_c^g = \tau_0 + \mu b(\Sigma_l a^{gl}\rho^l)^{1/2} \tag{3.175}$$

where the matrix a^{gl} gives the action of the family (l) of dislocations, of density ρ^l, on the work-hardening of the system (g). Other expressions could be derived, representing more complex effects that account for a kinematic part of the work-hardening. However, in attempting to derive sufficiently general relations we come up against many difficulties, including that of finding an expression for the evolution of the densities ρ^l (see below).

(c) *Continuous distributions of dislocations.* In the last 1950s, it was hoped to bridge the gap between the microscopic and macroscopic levels by developing a theory of continuous distributions of dislocations that enabled reliable calculations of the associated stress fields (Kondo 1952; Kröner 1958; Bilby 1960).

Consider a surface element ΔS, with outward normal $\underline{n}$, crossed by a number of dislocations with the resultant Burgers vector $\Delta\underline{b}$. Defining a second-order tensor $\underline{\underline{\alpha}}$ for the dislocation density by

$$\Delta\underline{b} = \underline{\underline{\alpha}}.\underline{n}\Delta S \tag{3.176}$$

and taking the limit as $\Delta S \to 0$, we arrive at the concept of a continuous distribution $\underline{\underline{\alpha}}(\underline{x})$. In a given system of axes the diagonal terms ($\alpha_{11}, \alpha_{22}, \alpha_{33}$) are the densities of screw dislocations (Burgers vector parallel to an axis normal to the surface element); the off-diagonal terms are the densities of edge dislocations *(e.g.* α_{12} relates to those that are parallel to $\underline{x}_2$ and have Burgers vectors parallel to $\underline{x}_1$).

The resultant Burgers vector for the dislocations crossing the surface S is

$$b_i = \int_S \alpha_{ij} n_j \mathrm{d}S \tag{3.177}$$

This vector will vanish if S is a closed surface; in this case therefore

$$0 = \int_S \alpha_{ij} n_j \mathrm{d}S = \int_S \alpha_{ij,j} \mathrm{d}S \Rightarrow \mathrm{div}\,\underline{\underline{\alpha}} = 0 \tag{3.178}$$

The condition $\mathrm{div}\underline{\underline{\alpha}} = 0$ is a generalisation to this continuum theory of the *law of nodes* given in Sect. 3.3.3.1 for discrete dislocations. We can use this to calculate the internal stress field associated with a given distribution $\underline{\underline{\alpha}}\,(\underline{x})$; one way to do this, drawing on the methods of continuum mechanics, is to associate with a dislocation density $\underline{\underline{\alpha}}(\underline{x})$ an equivalent distribution of plastic strains $\underline{\underline{\varepsilon}}^{\mathrm{p}}(\underline{x})$ and plastic rotations $\underline{\underline{\omega}}^{\mathrm{p}}(\underline{x})$, with the plastic distortions $\underline{\underline{\beta}}^{\mathrm{p}}(\underline{x})$ given by $\underline{\underline{\beta}}^{\mathrm{p}} = \underline{\underline{\varepsilon}}^{\mathrm{p}} + \underline{\underline{\omega}}^{\mathrm{p}}$.

This can be done by a similar procedure to that of the Burgers circuit (Sect. 3.3.3.1, Fig. 3.19). Let C be the edge of a surface S, which is crossed by dislocations of continuous density $\underline{\underline{\alpha}}$. As in the circuit procedure, we integrate along C the differential form $\mathrm{d}u_i{}^{\mathrm{p}} = \beta_{ij}^{\mathrm{p}}\ \mathrm{d}x_j$, which corresponds to the movements associated with the dislocations passing through S that have crossed the line C; this gives the opposite of the resultant Burgers vector. Thus

$$b_i = \int_S \alpha_{ij} n_j \mathrm{d}S = -\int_C \beta_{ij}^{\mathrm{p}} \mathrm{d}x_j = -\in_{jkl} \int_S \beta_{il,k}^{\mathrm{p}} n_j \mathrm{d}S \tag{3.179}$$

using Gauss's theorem; $\in_{jkl} = 0$ if any two suffixes are equal, otherwise $= \pm 1$ according as (j,k,l) are an even or an odd permutation of (1,2,3). Since the result holds for any surface S we must have

$$\alpha_{ij} = -\in_{jkl} \beta_{il,k}^{\mathrm{p}} \tag{3.180}$$

or in intrinsic notation $\underline{\underline{\alpha}} = \mathrm{curl}_{\mathrm{r}}\underline{\underline{\beta}}^{\mathrm{p}}$ (where the index r stands for "right").

Thus the dislocations $\underline{\underline{\alpha}}(\underline{x})$ have the same stress field as the incompatible plastic distribution $\underline{\underline{\beta}}^{\mathrm{p}}(\underline{x})$ (if $\underline{\underline{\beta}}^{\mathrm{p}}(\underline{x})$ were compatible its curl, and therefore $\underline{\underline{\alpha}}(\underline{x})$, would vanish). Note that using the same process a distribution of *quasi-dislocations* can be associated with any field of *stress-free strains* (or *eigenstrains*, *e.g.* of thermal or transformational origin), to have the same static effects: this is a concept that is identified with, and generalises, that of *geometrically necessary dislocations*, which we developed in Sect. 3.3.3.5.

Given an incompatible distribution $\underline{\underline{\beta}}^{\mathrm{p}}(\underline{x})$ which, if applied to a continuous medium would destroy its cohesion and continuity everywhere, we can envisage restoring the continuous geometry by distributing through the medium a continuous field of dislocations with density

$$\underline{\underline{\alpha}}' = -\underline{\underline{\alpha}} = \mathrm{curl}_{\mathrm{r}}\underline{\underline{\beta}}^{\mathrm{p}} \tag{3.181}$$

To determine the internal stress field associated with $\underline{\underline{\alpha}}$ we have to calculate the elastic field $\underline{\underline{\beta}}^{\mathrm{e}}$ that will restore the compatibility of the total transformation with gradient $\underline{\underline{\beta}} = \underline{\underline{\beta}}^{\mathrm{e}} + \underline{\underline{\beta}}^{\mathrm{p}}$, such that $\mathrm{curl}_{\mathrm{r}}\underline{\underline{\beta}} = 0$. Several computational methods for solving this elastic problem enable the associated stress field to be found; they can be applied to discrete dislocations too (reproducing results already known) by representing the densities $\underline{\underline{\alpha}}$ as sets of Dirac functions along the dislocation lines.

We might consider starting our attempt to formulate the laws of work-hardening by expressing the fact that in order to move a particular dislocation the stresses due to all the others must be overcome; but we should meet again the same difficulty as before, that of finding the laws governing the evolution of the work-hardening parameters. This is one of the reasons why the present development of the theory of continuously-distributed dislocations applies mainly to static conditions; attempts to describe the way in which the densities of dislocations evolve, and their effects on work-hardening, have so far not made significant progress – not to mention the fact that the density $\underline{\underline{\alpha}}$ gives a very poor representation for dislocation dipoles or small dislocation loops, which can play an important role.

(d) *Plastic flow*

(d1) *Kinematics*. It is clear that comparable difficulties will be met in attempting to formulate the laws of plastic flow: that is, to find the relation between the plastic strain rate $\underline{\underline{\dot{\varepsilon}}}^{\mathrm{p}}$ and the stress rate $\underline{\underline{\dot{\sigma}}}$. One could start with Orowan's analysis (Sect. 3.3.6.2) and formula (3.24) $\dot{\gamma} = \rho_{\mathrm{M}} b v$, giving the plastic slip rate as a function of the scalar density ρ_{M} of the mobile dislocations and their average velocity v. This scalar relation has to be replaced by a 3D relation involving the velocities relative to two crystallographic planes of the same family; if these are a distance D apart, have normal $\underline{n}$ and are slipping in a direction $\underline{m}$, then with the origin O in one plane and the point $M(x_i)$ in the other (Fig. 3.125) we have

$$\underline{V}(M) = \rho_{\mathrm{M}} D v \underline{b} \tag{3.182}$$

where $\underline{b} = b\underline{m}$ and $D = \underline{OM}.\underline{n} = x_k n_k$.

From this

$$V_i = \dot{\gamma} x_k n_k m_i \tag{3.183}$$

and if γ can be assumed uniform in the volume being considered

$$\dot{\varepsilon}^{p}_{ij} = \tfrac{1}{2}(V_{i,j} + V_{j,i}) = \tfrac{1}{2}\dot{\gamma}(m_i n_j + m_j n_i) \tag{3.184}$$

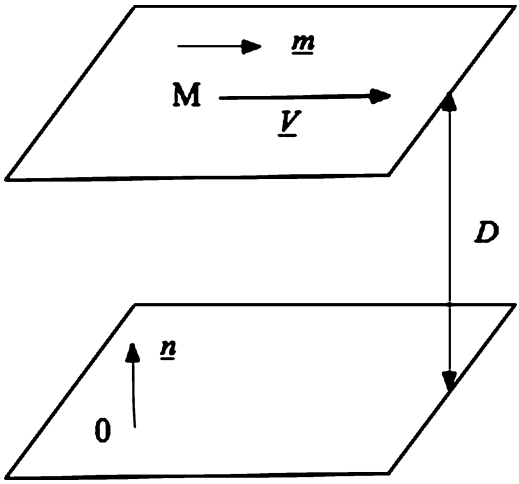


Fig. 3.125 Kinematics of plastic glide

or in intrinsic notation

$$\underline{\underline{\dot{\varepsilon}}}^{\mathrm{p}} = \dot{\gamma}(\underline{m} \otimes \underline{n})^{\mathrm{S}} \tag{3.185}$$

where $\underline{\underline{T}}^{\mathrm{S}}$ denotes the symmetric part of $\underline{\underline{T}}$ (see the microscopic equivalent in Eq. 3.32b). If several slip systems are active simultaneously then

$$\underline{\underline{\dot{\varepsilon}}}^{\mathrm{p}} = \sum_g \dot{\gamma}^g (\underline{m}^g \otimes \underline{n}^g)^{\mathrm{S}} \tag{3.186}$$

and if $\underline{\underline{T}}^{\mathrm{A}}$ is the anti-symmetric part of $\underline{\underline{T}}$, the plastic spin $\underline{\underline{\dot{\omega}}}^{\mathrm{p}}$ is

$$\underline{\underline{\dot{\omega}}}^{\mathrm{p}} = \sum_g \dot{\gamma}^g (\underline{m}^g \otimes \underline{n}^g)^{\mathrm{A}} \tag{3.187}$$

This is a purely kinematic description. We need to go further and find what part ρ_{M} of the total dislocation density ρ can be moved by the existing load, and how this depends on $\underline{\underline{\sigma}}$; and also find how the average velocity v of the dislocations varies. The questions that arise here have been answered in many special cases, but answers general enough to help us to go from the scale of defects in crystals to that of overall behaviour of a single crystal, over a large field of applications, have still to be found.

Nevertheless, let us mention here a promising approach (Mecking and Kocks 1981; Teodosiu et al .1993) which associates with the description of work-hardening given by (3.175) a plastic flow equation through the prescription of the evolution of the dislocation densities on every slip system, such as

$$\dot{\rho}^{(l)} = \frac{1}{b}\left(\frac{1}{L^{(l)}} - G_c \rho^{(l)}\right)\left|\dot{\gamma}^{(l)}\right| \tag{3.188}$$

where G_c is a parameter proportional to the characteristic length associated with the annihilation process of dislocation dipoles and $L^{(l)}$ the mean free path of system (l) (see Eq. 3.96). The latter can be expressed as

$$L^{(l)} = K\left(\sum_{g \neq l} \rho^{(g)}\right)^{-1/2} \tag{3.189}$$

where K is a material parameter. This approach is expected to be combined with Discrete Dislocations Dynamics (DDD) simulations in order to evaluate the concerned material parameters and then with upscaling models from single crystal to polycrystal (see Sect. 3.5.4) so as to develop a full multiscale analysis of polycrystal plasticity.

(d2) *Work-hardening matrix*. It is because of the difficulty to derive general enough predictive treatments based on the scale of dislocation densities that recourse is often made to a semi-phenomenological description which avoids the difficulties by choosing the slips $\dot{\gamma}^g$ themselves as work-hardening parameters. With this choice the laws of work-hardening and of flow coincide since we can write, using the work-hardening matrix h^{gl}, $\dot{\tau}^g = \dot{\tau}_c^g$ on an active system. Then

$$\dot{\tau}_c^g = \sum_l h^{gl}\dot{\gamma}^l \tag{3.190}$$

$$\underline{\underline{\dot{\varepsilon}}}^p = \sum_g \dot{\gamma}^g (\underline{m}^g \otimes \underline{n}^g)^S = \sum_g \dot{\gamma}^g \underline{\underline{R}}^g \tag{3.191}$$

$$\dot{\tau}^g = \underline{m}^g . \underline{\underline{\dot{\sigma}}} . \underline{n}^g = (\underline{m}^g \otimes \underline{n}^g)^S : \underline{\underline{\dot{\sigma}}} = \underline{\underline{R}}^g : \underline{\underline{\dot{\sigma}}} \tag{3.192}$$

Combining these, writing $[k] = [h]^{-1}$ and ignoring any changes that may occur in $\underline{m}^g$ and $\underline{n}^g$ in the course of the flow, we get the plastic flow equation

$$\underline{\underline{\dot{\varepsilon}}}^p = \sum_{g,l} k^{gl}\left(\underline{\underline{R}}^l : \underline{\underline{\dot{\sigma}}}\right)\underline{\underline{R}}^g = \sum_{g,l}\left(k^{gl}\underline{\underline{R}}^g \otimes \underline{\underline{R}}^l\right) : \underline{\underline{\dot{\sigma}}} \tag{3.193}$$

Some comments are needed here:

1. *The work-hardening matrix* h^{gl}. This expresses the hardening produced in the system (g) by the plastic activity on the system (l); it therefore expresses anisotropy in the interactions of slip systems, in particular the distinction between *self-hardening* (the diagonal terms) and *latent hardening* (the off-diagonal terms).

This matrix will change as the flow proceeds; this change can be investigated experimentally in single crystals (Franciosi et al .1980): a primary specimen is pre-deformed by simple slip (g), and from this are cut several secondary samples at different orientations. Each of these latter is then deformed again according to some system (l), for which the new critical resolved shear stress will depend on the short-range interactions between (g) and (l). In general, the off-diagonal terms are found to be greater than the diagonal, with the strongest interactions corresponding, in the case of FCC crystals, to pairs of systems for which the dislocations can form Lomer-Cottrell locks (Sect. 3.3.8.1) or, to a lesser degree, glissile junctions. DDD simulations have shown recently that collinear interactions are very strong too (Madec et al. 2003). The dissociation, favoured by a low stacking fault energy, increases these interactions.

The weak point of the model is the choice of the work-hardening parameters: with this it is difficult to express the fact that micro-slips produced by loads that are too small, according to the Schmid criterion, to initiate plastic activity can have a significant effect on the hardening (positive or negative). Further, the inversion of the work-hardening law and of the work-hardening matrix can raise problems, particularly in the case of non-uniqueness of the combination of active systems; we then have to investigate whether the matrix is positive definite or not, which is difficult to check experimentally.

2. *The vectors* $\underline{m}^g$, $\underline{n}^g$. In general, it is only for very small deformations that these vectors can be regarded as remaining fixed; they are defined crystallographically and are susceptible to a rotation of the crystal lattice. Such a lattice rotation (Sect. 3.3.2), which must not be confused with that associated with plastic slip, does nevertheless accompany this slip and the resultant total rotation enters the compatibility equations, curl $\underline{\underline{\beta}} = 0$. Thus in a tensile test on a single crystal in which the machine's jaws are constrained to remain aligned the conditions imposed on the total rotation determine the lattice rotation by difference from the plastic rotation.

If the total spin is zero the lattice and plastic spins must be equal and of opposite signs ($\underline{\underline{\dot{\omega}}}^{\mathrm{L}} = -\underline{\underline{\dot{\omega}}}^{\mathrm{p}}$). The vectors $\underline{m}^g$, $\underline{n}^g$ therefore vary with time and the rate of change of the resolved shear stress $\dot{\tau}^g$ depends on them as well as on $\underline{\underline{\sigma}}$. This effect, called *geometrical work-hardening*, is especially important in crystal plasticity at finite strains as it is responsible for the formation of crystallographic textures.

3.5.1.3 Polycrystals

Polycrystal plasticity presents new levels of complexity in addition to the difficulties just described:

- The single crystal grains are separated by boundaries (Sect. 3.3.8.3), which constitute complex systems, rarely reducible to simple plane arrangements of dislocations. Their influence on the plastic behaviour of the polycrystal is exerted at several levels:

  - at the microscopic level it is determined by the nature of the interactions between dislocations of the lattice and defects constituting the boundary. In addition to the image-force effects are those of the obstacle to the propagation of the lattice dislocations; these last are responsible in particular for the pile-up of dislocations and the consequential stress concentrations which in turn can activate sources of dislocations in neighbouring grains, resulting in the transmission of plastic flow from grain to grain (see Sect. 3.4.2.3, Eq. 3.104: interpretation of the Hall-Petch law concerning the influence of average grain size on the flow stress of a polycrystal). Depending on its structure the boundary can also act in a more complex manner, equally as a source or as a sink of dislocations, all the more so because it is generally a favoured site for the segregation of impurities, deposition of precipitates, formation of cavities, etc.
  - at an intermediate level the boundary is responsible for the formation of a zone in its neighbourhood, rich in defects and on average more work-hardened than the interior of the grains. This results in a heterogeneity in which the grain is divided into a number of sub-domains of greater or lesser hardness, giving the grain itself a complex structure and an average behaviour which, because of the non-linearity of plasticity, is very different from that of an isolated single crystal.
  - finally, at the macroscopic level the grain boundary forms a frontier between two different crystal orientations: even if the plastic deformation of each grain is uniform (and therefore compatible) the overall plastic strain field is no longer compatible. The consequence is an internal stress field – possibly computable from the theory of continuous distributions of dislocations, in this case for a surface distribution localised on the boundary – which causes the local stresses to vary significantly from grain to grain, reproducing only in the mean the stresses applied to the polycrystalline element.

- In addition to the effects of the individual grain boundaries, the overall behaviour of the polycrystal is determined by the overall distribution of the grain boundaries (location, orientation, dimensions) which controls the shapes and sizes of the grains, and also by the crystal orientations of each grain and the way these change as the flow proceeds. It is not enough, for a quantitative analysis of these parameters and their influence, to know only their probabilities; their spatial distributions also must be known, whether deterministic or probabilistic. Thus for example the response of a given grain depends not only on its shape, dimensions and crystal orientation but also on the details of its neighbourhood, that is, the shape, dimensions and orientation of its near neighbours and even of more distant grains if they interact with it.

Faced with these difficulties, we have a choice of three methods:

1. Analyse all the stages in shifting from a single crystal to a polycrystal in as formally a deductive manner as possible. This approach, making the greatest

demands of the three on an understanding of the mechanisms and phenomena involved, is the objective of much current research.
2. Ignore the most detailed phenomena and retain only the parameters that are the most important for the description of the polycrystalline aggregate, treating the polycrystal as a composite material for which only an incomplete, statistical description is known. The problems encountered in this approach, and the methods that have been developed to solve them, are related to what we have given in Sect. 2.9 on the elastic properties of heterogeneous materials. We give a short illustration at the end of this chapter.
3. Take a more phenomenological point of view, using only some of the more general microstructural features of plasticity of crystalline materials as a basis for defining special classes of elastoplastic behaviours, possibly applying to other materials as well. This is the method we shall use here, developing a rheological formulation of elastoplastic behaviour under three headings, criteria, work-hardening and flow. After this we give some introductory indications to the problems of computing the mechanical response of macroscopic structures.

3.5.2 Criteria, Work-Hardening and Plastic Flow

3.5.2.1 Plasticity Criteria

The problem here is to derive explicit criteria for the plastic yielding of the material, relevant to some aspect of its properties and corresponding to a given state, that is, to given values of the work-hardening parameters. We deal first with the isotropic case and then with the anisotropic.

(a) *Isotropic criteria*

We assume an isotropic plastic behaviour such that the conditions for the starting or continuing the flow are independent of the orientation of the load with respect to the material. Before treating the general problem we consider the form the Schmid law, so far restricted to single crystals which are obviously anisotropic, might take when generalised to the isotropic case.

(a1) *Generalisation of Schmid law; Tresca criterion and offshoots.*

The Schmid law is anisotropic both in that it presupposes only a limited number of easy glide systems, so that the critical shear stress for all others could be considered to be infinite, and in that it admits the possibility of different systems having

different critical shear stresses. "Isotropising" the law means requiring all the critical shear stresses to have the same value τ_c and allowing *any* slip system, defined by the couple ($\underline{m}$, $\underline{n}$), to count as "easy glide" system, that is, to become plastically active when the resolved shear stress reaches this critical value. The first change, equalising the shear stresses, makes the law a criterion for the maximum critical resolved shear stress with the maximum to be found within the limited set of easy-glide systems; the second removes this last restriction and allows the search for the maximum to be made over all possible slip systems. For any given stress $\underline{\underline{\sigma}}$ the greatest resolved shear stress over a plane is that defined by the projection of the stress vector $\underline{T} = \underline{\underline{\sigma}}.\,\underline{n}$ on that plane, and not on a particular direction $\underline{m}$. If we are to take all possible planes into consideration we have to find the greatest shear stress. We thus arrive at a criterion for the *critical maximum shear stress*, called the Tresca[25] criterion.

The maximum shear stress being equal to half the difference between the extreme principal values (*cf.* Mohr's construction, which gives this in the form of the radius of the largest Mohr circle), the Tresca criterion can be written as

$$\mathrm{Sup}_i \sigma_i - \mathrm{Inf}_j \sigma_j \leq 2\tau_c \tag{3.194}$$

where σ_i, σ_j are the principal stresses of the tensor $\underline{\underline{\sigma}}$.

Before studying the implications of such a criterion it is as well to look into its physical meaning. The occurrence of the critical value τ_c in the Tresca criterion, when it occurs also in the Schmid law, might suggest that this is a strict statement of the necessary condition for plastic flow in an isotropic polycrystal (that is, one without crystallographic texture) consisting of crystalline grains, each of which satisfies the Schmid law when all their (easy glide) slip systems have the same critical shear stress τ_c. We might argue that a polycrystalline volume element having sufficiently small grains will have a quasi-continuous distribution of easy-glide systems in which all possible orientations have the same probability. For any given stress state the first systems to be activated would be those subjected to the maximum resolved shear stress, in practice to the maximum shear stress, and when this latter reaches the critical value τ_c the polycrystalline element as a whole would become active plastically.

This argument however is far from sound. On the one hand, before even the first grain shows plastic activity on the most loaded of its slip systems the stress state will not in general be the same from one grain to another. It only needs their elastic behaviour to be anisotropic for internal stresses to appear, with the result that the maximum shear for the overall stress state will have no local significance. On the other hand, even if it could be assumed that the stress states were identical for all

[25] Henri Tresca (1814–1885) was a French mechanical engineer.

the grains, plastic activity of a small fraction of the grains would be constrained by the elastic response of all the others and could not have any significant macroscopic effect: a large fraction of the grains would have to be already exhibiting plastic flow before the aggregate as a whole could be considered as being in the plastic regime. If plastic flow has started in different grains the stress state certainly cannot be uniform, for the reasons already given in Sect. 3.5.1.3; and it is not sufficient to know that the polycrystal is without texture to be able to analyse the evolution of the local stresses, which depends on all the features of the spatial distribution of the grains, their shape, dimensions, orientation, relative location, etc. Thus the question is very complex, and cannot be answered as readily as might be supposed; we shall return to it in Sect. 3.5.3.6 below, in connexion with the Taylor model.

It is advisable here not to give too precise a physical sense to the Tresca criterion, and to treat it as a phenomenological criterion, taking account of the isotropic nature of the material. To emphasise this we replace the quantity τ_c by k, where k has no significance except that of a yield strength in simple shear (since in this case $\sigma_1 = -\sigma_2$ and $\sigma_3 = 0$), to be found experimentally. Given this, if the Tresca criterion is satisfied then the yield strength can be predicted for any other stress state – clearly, for simple tension for example, $R_p = 2k$.

It should be noted in particular that Tresca's criterion does not depend on the principal intermediate stress nor, more generally, on the mean stress $\sigma_m = (1/3)\mathrm{Tr}\underline{\underline{\sigma}}$; thus according to this criterion two stress states that differ only by an isotropic state $(\lambda\underline{\underline{\delta}}, \forall\lambda)$ are equivalent. This is the same as saying that the criterion is unaffected by a hydrostatic pressure (positive or negative) and depends only on the deviatoric stress $\underline{\underline{s}} = \underline{\underline{\sigma}} - \sigma_m\underline{\underline{\delta}}$. This is not surprising, since it was derived as an extension of the Schmid law, which already has this property: a resolved shear obeys

$$\tau_r = \underline{m}.\underline{\underline{\sigma}}.\underline{n} = \underline{m}.(\underline{\underline{s}} + \sigma_m\underline{\underline{\delta}}).\underline{n} = \underline{m}.\underline{\underline{s}}.\underline{n} \text{ since } \underline{m}.\underline{n} = 0 \tag{3.195}$$

and then depends only on the deviatoric stress.

In the (σ,τ) plane of the Mohr[26] diagram the elastic – or, if rigid-plastic behaviour is to be considered, *non-plastic* – domain is a strip bounded by the lines $\tau = \pm k$ parallel to the σ-axis; plastic response is possible only when the Mohr circle touches these (Fig. 3.126).

This last statement has suggested, especially for rock and concrete mechanics, that the Tresca criterion might be extended to apply to materials that do not satisfy it as it stands, by considering, in the Mohr plane, an elastic domain bounded by the so-called "intrinsic curve" $|\tau| = g(\sigma)$, the envelope of the Mohr circles corresponding, for all possible stress states, to the plastic limit. In spite of the name such criteria are not in general "intrinsic", but they are often used for such materials as soils, rocks, concrete, etc. because they have the advantage of being able to express the property of sensitivity of the plastic regime to normal stresses, frequently observed in these

[26]Christian Otto Mohr (1835–1918) was a German civil engineer.

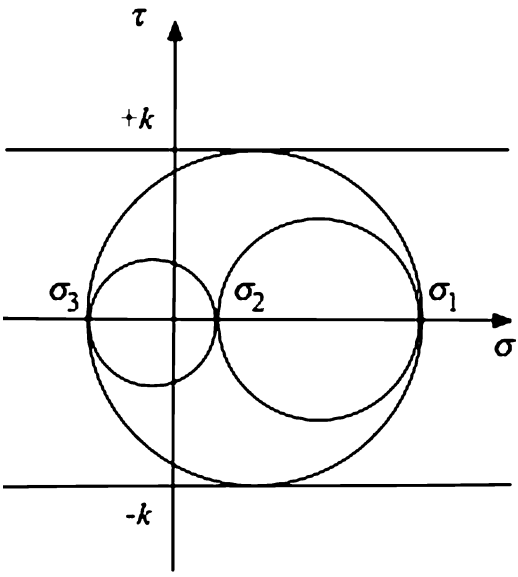


Fig. 3.126 The Tresca criterion

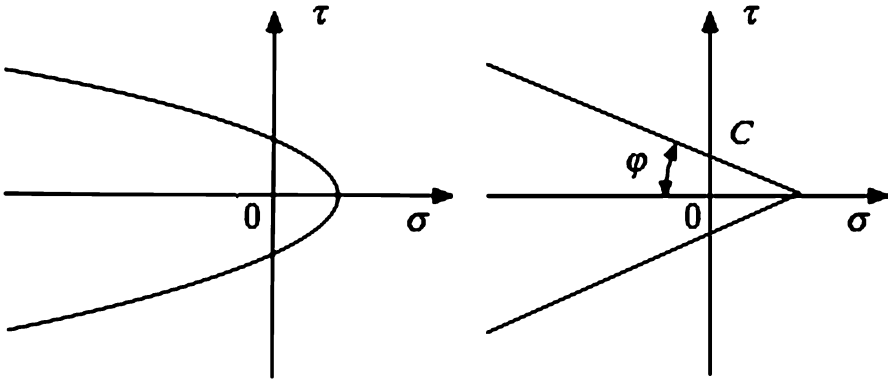


Fig. 3.127 Intrinsic curve and Coulomb's criterion

cracked or weakly-coherent materials. It is equally possible to express in this way the strong differences between the elastic limits (and the fracture stresses) in tension and compression, and, more generally, the great lack of symmetry between the effects of hydrostatic pressure and of symmetrical tri-axial expansion ($\underline{\underline{\sigma}} = \lambda \underline{\underline{\delta}}$, $\lambda > 0$). It follows that the intrinsic curves are concave towards negative normal stresses (Fig. 3.127).

A much-used member of this family of criteria is the *Coulomb*[27] *criterion* (Fig. 3.127), represented in the (σ,τ) plane by a region of non-plasticity (or even of non-fracture), bounded by two straight-line segments intersecting on the σ-axis and making angles $\pm\varphi$ with this. For certain materials the "cohesion" yield strength C, determined by the critical shear at zero normal stress, can be taken as negligible.

(a2) *General expression for isotropic criteria: von Mises criterion.*

It is easy to give a general expression for criteria for isotropic plasticity; the Tresca criterion is just one particular case. For a given state of work-hardening the yield function $f(\underline{\underline{\sigma}})$ depends only on the stress tensor $\underline{\underline{\sigma}}$. Instead of representing this symmetric tensor by its six components relative to one particular set of axes we can describe it by its three eigen-values σ_1, σ_2, σ_3 and the orientation of its eigen-vectors. The latter being orthogonal it can be defined by three angles, for example its Eulerian angles in some reference system, say ψ, θ, φ. The yield function is thus a function of the six variables, σ_1, σ_2, σ_3, ψ, θ, φ.

However, the criterion is isotropic, which means that the value of the yield function does not change if this set of eigen-vectors is rotated in any way with respect to the material. The yield function must therefore depend only on σ_1, σ_2, σ_3; so *in this case* the stress space can be reduced to one of three dimensions only. Further, the isotropy implies that f is a symmetrical function of the σ_i, none of the principal stresses having any privilege over the others. These are the roots of the characteristic equation

$$\left|\sigma_{ij} - \lambda\delta_{ij}\right| = 0 \tag{3.196}$$

a cubic in λ which we can write

$$-\lambda^3 + \mathrm{I}_1\lambda^2 - \mathrm{I}_2\lambda + \mathrm{I}_3 = 0 \tag{3.197}$$

where

$$\begin{aligned}
\mathrm{I}_1 &= \mathrm{Tr}(\underline{\underline{\sigma}}) = \sigma_{kk} = \sigma_1 + \sigma_2 + \sigma_3 \\
\mathrm{I}_2 &= 1/2\left[\left(\mathrm{Tr}(\underline{\underline{\sigma}})\right)^2 - \mathrm{Tr}(\underline{\underline{\sigma}}^2)\right] = 1/2\left(\sigma_{kk}^2 - \sigma_{ij}\sigma_{ij}\right) \\
&= 1/2\left[(\sigma_1 + \sigma_2 + \sigma_3)^2\right] - (\sigma_1^2 + \sigma_2^2 + \sigma_3^2) \\
\mathrm{I}_3 &= \det(\underline{\underline{\sigma}}) = \sigma_1\sigma_2\sigma_3
\end{aligned} \tag{3.198}$$

I_1, I_2, I_3 are invariants of $\underline{\underline{\sigma}}$: their values remain unchanged when the axes of reference are changed; they are symmetric functions of σ_1, σ_2, σ_3 of degrees 1,

[27] Charles-Augustin de Coulomb (1736–1806) was a French physicist.

2 and 3 respectively. Any invariant of $\underline{\underline{\sigma}}$ can be derived from them, or from the following equivalent forms:

$$\begin{aligned} \mathrm{I'}_1 &= \mathrm{Tr}(\underline{\underline{\sigma}}) = \sigma_{kk} \\ \mathrm{I'}_2 &= 1/2\,\mathrm{Tr}(\underline{\underline{\sigma}}^2) = 1/2\,\sigma_{ij}\sigma_{ij} \\ \mathrm{I'}_3 &= 1/3\,\mathrm{Tr}(\underline{\underline{\sigma}}^3) = 1/3\,\sigma_{ij}\sigma_{jk}\sigma_{ki} \end{aligned} \tag{3.199}$$

and any isotropic criterion is necessarily of the form $f(I_1, I_2, I_3) = 0$ or $f(I_1', I_2', I_3') = 0$: this is the general form for such a criterion.

If now we restrict ourselves to criteria *independent of hydrostatic pressure*, and therefore dependent only on the deviatoric stress $\underline{\underline{s}} = \underline{\underline{\sigma}} - \sigma_m \underline{\underline{\delta}}$, we can work through the argument again in the same way; we find that these depend only on the basic symmetrical invariants of the deviatoric stress, for example

$$\begin{aligned} \mathrm{J}_1 &= \mathrm{Tr}(\underline{\underline{s}}) = s_{kk} = 0 \\ \mathrm{J}_2 &= 1/2\,\mathrm{Tr}(\underline{\underline{s}}^2) = 1/2\,s_{ij}s_{ij} \\ \mathrm{J}_3 &= 1/3\,\mathrm{Tr}(\underline{\underline{s}}^3) = 1/3\,s_{ij}s_{jk}s_{ki} \end{aligned} \tag{3.200}$$

Thus these criteria depend only on two variables, J_2 and J_3. Further, the isotropy implies symmetry, *i.e.*, $f(-\underline{\underline{s}}) = f(\underline{\underline{s}})$, and as J_2 is an even function of $\underline{\underline{s}}$ whilst J_3 is odd, f can be written as

$$f\left(\mathrm{J}_2, \mathrm{J}_3{}^2\right) = 0 \tag{3.201}$$

Note: Since the Tresca criterion also is isotropic and independent of hydrostatic pressure, it also can be put in the form (3.201).

It is interesting to study, as an example of this family of criteria, those that are independent of the third invariant J_3. They can be put in the form $f(J_2) = 0$, or, what comes to the same thing,

$$\mathrm{J}_2 = k^2 (\text{since } \mathrm{J}_2 = {}^1\!/_2 s_{ij}s_{ij} \geq 0) \tag{3.202}$$

This is the *von Mises*[28] *criterion*, very much used because of its simplicity and regularity. It can be used in any of the following equivalent forms:

[28] Richard Edler von Mises (1883–1953) was an Austro-Hungarian scientist and engineer.

$$\begin{aligned} \mathrm{J}_2 &= (1/2)\, s_{ij}\, s_{ij} = (1/2)\left(s_1^2 + s_2^2 + s_3^2\right) = -\left(s_1 s_2 + s_2 s_3 + s_3 s_1\right) \\ &= (1/2)\left[\sigma_{ij}\sigma_{ij} - (1/3)\,\sigma_{kk}^2\right] = (1/6)\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \\ &= (1/6)\Big[(\sigma_{11} - \sigma_{22})^2 + (\sigma_{22} - \sigma_{33})^2 + (\sigma_{33} - \sigma_{11})^2 \\ &\qquad + 6(\sigma_{23}^2 + \sigma_{31}^2 + \sigma_{12}^2)\Big] = k^2 \end{aligned} \tag{3.203}$$

Here, like for the Tresca criterion, k has the meaning of the yield strength in simple shear ($\sigma_1 = -\sigma_2, \sigma_3 = 0$); the yield strength in simple tension is then $R_\mathrm{p} = k\sqrt{3}$. Although introduced in a purely phenomenological manner the von Mises criterion can be given a physical interpretation when applied to an isotropically elastic material. It is then easy to prove that J_2 is proportional to the elastic distortion energy, that is, the total elastic energy less the part arising from a pure volumetric dilatation θ; so one can say that in this case the criterion expresses the fact that plastic yield occurs when the elastic distortion associated with elastic shear strains reaches a critical value.

Less suggestively, the von Mises criterion can also be regarded as a criterion for "critical octahedral stress", the octahedral stress being the stress exerted on a plane normal to the tri-sector of the principal directions of the stresses. A more reasonable view is to consider the von Mises criterion as a typical representative of isotropic criteria that are independent of hydrostatic pressure, which has a particularly convenient and smooth analytic expression – in contrast to the Tresca criterion, for which the expression varies according to the nature of the intermediate principal stress.

It is reassuring to find that the Tresca and von Mises criteria give quite close predictions concerning plasticity. This can be seen geometrically by comparing their yield surfaces in the principal stress space (Fig. 3.128). As with all isotropic criteria that are independent of hydrostatic pressure these are cylinders with generators parallel to the tri-sector Δ, since a translation in this direction corresponds to the addition of such a pressure, which has no effect on the plastic yield conditions. The cylinder is circular for von Mises and a regular hexagonal prism for Tresca, and both have Δ as axis. To compare the two predictions we must adjust the cylinders so that both contain given experimental data, corresponding to a certain stress state. According to the data chosen, the von Mises surface can vary with respect to the Tresca from a cylinder inscribed in the hexagonal prism to one exscribed about it; the difference between the two never exceeds about 15%, which is acceptable in comparison with the experimental uncertainties. Though some micromechanical simulations have shown that the plastic yield of an isotropic FCC polycrystal evolves from Tresca's to von Mises's prediction when the plastic offset increases from zero, there is no strong physical reason to regard either of these criteria as "better" than the other; the important thing is that both are isotropic and independent of hydrostatic pressure, and therefore well suited to the study of isotropic crystalline materials that are deformed plastically by dislocation glide, and therefore without volume change.

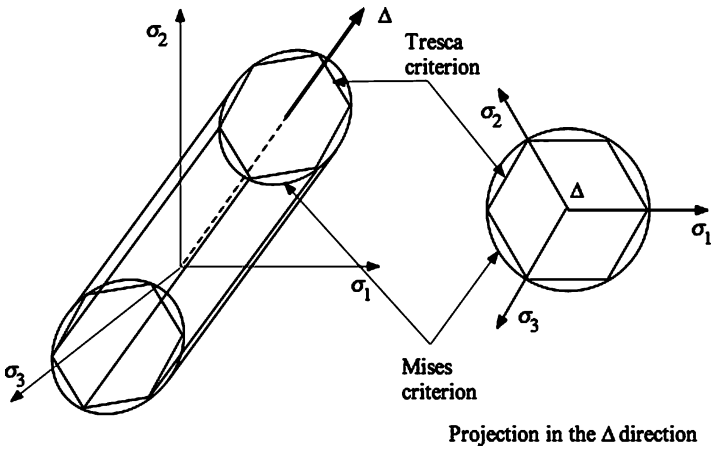


Fig. 3.128 Representations of the von Mises criterion (*circular cylinder*) and of the Tresca criterion (regular hexagonal prism) in the stress space

As an example, we can compare the two criteria for a tension-torsion test, with either a simple traction or a simple shear chosen for the adjustment. If the outcome tends to favour von Mises this is primarily because of its greater analytic regularity, which, as we shall see, has great advantages in describing plastic flow.

(b) *Anisotropic criteria.*

These are more complex and it is hardly possible to give a general form. Taking the typical form of an isotropic criterion as guide, we can, however, specify a whole family of anisotropic criteria that have two of the simplifying properties of the Mises criterion: independence of hydrostatic pressure (and therefore dependence only on the deviatoric stresses) and a quadratic function of these stresses. The isotropic expression of the form $\underline{\underline{s}} : \underline{\underline{s}}$ can then be "anisotropised" to $\underline{\underline{s}} : \mathbf{B} : \underline{\underline{s}}$, where $\mathbf{B}$ is a fourth-order tensor holding all the available information concerning the symmetries of the anisotropy that is being taken into account. For the most general anisotropy and a given state of work-hardening $\mathbf{B}$ provides 15 independent coefficients for fitting to experimental data: the initial $3^4 = 81$ components are reduced to 21 by the obvious symmetries $B_{ijkl} = B_{ijlk} = B_{jikl} = B_{klij}$ and then to 15 by $B_{jjkl} = 0$, expressing independence of hydrostatic pressure.

Particular symmetries of the anisotropy will reduce this number further, just as we saw in the case of the elastic moduli or compliances tensor in linear elasticity (see Sect. 2.3.2).

A particularly interesting case is that of orthotropic symmetry (see Sect. 2.3.2.4), illustrated by a study by Hill of plastic anisotropy in thin rolled sheets having orthotropic symmetries with respect to the rolling (RD), transverse (TD) and normal (ND) directions. With these as axes the expression $\underline{\underline{s}} : \mathbf{B} : \underline{\underline{s}}$ as a function of $\underline{\underline{\sigma}}$ has the simple form

$$\underline{\underline{s}}:\mathbf{B}:\underline{\underline{s}} = F(\sigma_{22}-\sigma_{33})^2 + G(\sigma_{33}-\sigma_{11})^2 + H(\sigma_{11}-\sigma_{22})^2 + 2L\sigma_{23}^2 + 2M\sigma_{31}^2 + 2N\sigma_{12}^2 \tag{3.204}$$

so that the criterion depends on only six coefficients. They can be determined experimentally by submitting thin or thick sheets to tests which are, in principle, similar to the "latent hardening tests" described in Sect. 3.5.1.2(d2) for single crystals (see Exercises in Volume III). The transverse isotropy with respect to ND (here the x_3 axis) would give $F = G$, $L = M$, $N = F + 2H$, reducing the number to 3. Finally, if there is complete isotropy then $F = H$, $L = N$, $N = 3F$ and we are left, apart from the factor $6F$, with the invariant J_2 of $\underline{\underline{s}}$ (which, as expected, corresponds to von Mises criterion for isotropy).

A limitation of this form of criterion is in its symmetry: $\underline{\underline{s}}:\mathbf{B}:\underline{\underline{s}}$ has the same value for $\underline{\underline{s}}$ and $-\underline{\underline{s}}$, implying, for example, that the yield strength for simple tension in any direction is equal and opposite to that for simple compression in the same direction, which can conflict with observation, particularly with composite materials. The restriction can be removed by using *Tsai's criterion*, which adds a linear form to the quadratic; thus for orthotropic symmetry (see Sect. 2.3.2.4), for example, the form is

$$F(\sigma_{22}-\sigma_{33})^2 + G(\sigma_{33}-\sigma_{11})^2 + H(\sigma_{11}-\sigma_{22})^2 + 2L{\sigma_{23}}^2 + 2M{\sigma_{31}}^2 + 2N{\sigma_{12}}^2 + P(\sigma_{11}-\sigma_{33}) + Q(\sigma_{22}-\sigma_{33}) = k^2 \tag{3.205}$$

Many other criteria have been devised, too many to review here. We must now consider how work-hardening and plastic flow are to be represented.

3.5.2.2 Laws of Work-Hardening and Plastic Flow

Since we have particular explicit expressions for yield functions $f(\underline{\underline{\sigma}})$ for some reference state we can derive for each one some forms for work-hardening, such as isotropic, kinematic or a combination of these. However, we still have to formulate the laws of work-hardening and of plastic flow. Significant progress can be made here if certain restrictions are imposed, guided by information obtained in the study of micro-plasticity in crystalline materials, and we shall develop comprehensive models of elastoplastic behaviour within such a restricted framework. Before embarking on this we can fill out some of the details in the introductory sketch we gave of these laws in Sect. 3.2.2.3.

We had arrived at the following expressions for a regular yield function:

$$\begin{cases} \mathrm{d}\varepsilon_{ij}^{\mathrm{p}} = A_{ijkl}(\sigma_{pq}, Y_m)\mathrm{d}\sigma_{kl} \\ \mathrm{d}Y_m = K_{mkl}(\sigma_{pq}, Y_n)\mathrm{d}\sigma_{kl} \end{cases} \text{if } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } \frac{\partial \mathrm{f}}{\partial \sigma_{kl}}\mathrm{d}\sigma_{kl} \geq 0$$
$$\begin{cases} \mathrm{d}\varepsilon_{ij}^{\mathrm{p}} = 0 \\ \mathrm{d}Y_m = 0 \end{cases} \text{if } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) < 0 \text{ or } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } \frac{\partial \mathrm{f}}{\partial \sigma_{kl}}\mathrm{d}\sigma_{kl} < 0 \tag{3.206}$$

together with the consistency condition ($\mathrm{df} = 0$):

$$\left(\frac{\partial \mathrm{f}}{\partial \sigma_{kl}} + \frac{\partial \mathrm{f}}{\partial Y_m}K_{mkl}\right)\mathrm{d}\sigma_{kl} = 0 \text{ if } f(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } \frac{\partial \mathrm{f}}{\partial \sigma_{kl}}\mathrm{d}\sigma_{kl} \geq 0 \tag{3.207}$$

It is interesting to consider the special case of a neutral loading, for which we must have

$$\begin{cases} A_{ijkl}(\sigma_{pq}, Y_m)\mathrm{d}\sigma_{kl} = 0 \\ K_{mkl}(\sigma_{pq}, Y_n)\mathrm{d}\sigma_{kl} = 0 \end{cases} \text{if } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } \frac{\partial \mathrm{f}}{\partial \sigma_{kl}}\mathrm{d}\sigma_{kl} = 0, \forall \mathrm{d}\sigma_{kl} \tag{3.208}$$

The simultaneous vanishing of $A_{ijkl}\mathrm{d}\sigma_{kl}$,$K_{mkl}\mathrm{d}\sigma_{kl}$ and $(\partial \mathrm{f}/\partial \sigma_{kl})\mathrm{d}\sigma_{kl}$ for all $\mathrm{d}\sigma_{kl}$ implies that $\partial \mathrm{f}/\partial \sigma_{kl}$ is a factor (from the point of view of tensor products) in both A_{ijkl} and K_{mkl}; therefore these two tensors can be written as

$$A_{ijkl} = \mathrm{h}_{ij}\,\partial \mathrm{f}/\partial \sigma_{kl} \quad K_{mkl} = \mathrm{k}_m \partial \mathrm{f}/\partial \sigma_{kl} \tag{3.209}$$

We thus have the following relations, valid only during the plastic flow:

$$\begin{cases} \mathrm{d}\varepsilon_{ij}^{\mathrm{p}} = \mathrm{h}_{ij}(\sigma_{pq}, Y_m)(\partial \mathrm{f}/\partial \sigma_{kl})\mathrm{d}\sigma_{kl} \\ \mathrm{d}Y_m = \mathrm{k}_m(\sigma_{pq}, Y_n)(\partial \mathrm{f}/\partial \sigma_{kl})\mathrm{d}\sigma_{kl} \end{cases} \text{if } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } (\partial \mathrm{f}/\partial \sigma_{kl})\mathrm{d}\sigma_{kl} \geq 0 \tag{3.210}$$

The consistency condition becomes

$$(1 + \mathrm{k}_m \partial \mathrm{f}/\partial Y_m)(\partial \mathrm{f}/\partial \sigma_{kl})\mathrm{d}\sigma_{kl} = 0, \forall \mathrm{d}\sigma_{kl}, (\partial \mathrm{f}/\partial \sigma_{kl})\,\mathrm{d}\sigma_{kl} \geq 0 \text{ and } \mathrm{f} = 0 \tag{3.211}$$

or more simply

$$1 + \mathrm{k}_m \partial \mathrm{f}/\partial Y_m = 0 \tag{3.212}$$

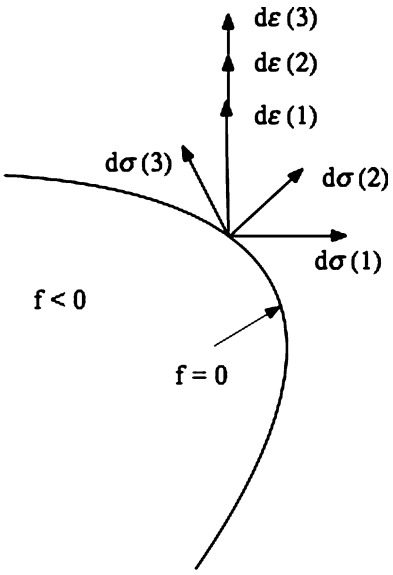


Fig. 3.129 Sketch showing that the direction of $d\varepsilon^{p}{}_{ij}$ is determined by σ_{ij} and not by $d\sigma_{ij}$

Thus in order to specify the work-hardening and flow functions, all that needs to be known, in addition to the yield function f and its gradient ∂f, is a second-order symmetric tensor function $h_{ij}(\underline{\underline{\sigma}}, \underline{Y})$ and a work-hardening function $k_m(\underline{\underline{\sigma}}, \underline{Y})$, which, for the sake of simplicity, we have taken as a vector function. Though this is still a very general form for the laws, we can derive from it an important statement concerning plastic behaviour and its dependence on the stresses.

Consider a limit loading such that $f(\underline{\underline{\sigma}}, \underline{Y}) = 0$. In this state plastic flow can be started, or continued if already started, by any stress increment $d\underline{\underline{\sigma}}$ such that $(\partial f/\partial \sigma_{kl})d\sigma_{kl} > 0$. The above equations show that the increment in plastic strain $d\varepsilon_{ij}^{p}$ will have the same direction whatever increment $d\underline{\underline{\sigma}}$ is chosen, since it is always proportional to $h_{ij}(\underline{\underline{\sigma}}, \underline{Y})$, the value of which is fixed by the stress state and the level of work-hardening reached *before* the incrementation. Thus the principal directions and the ratios of the principal values taken in pairs, or equally of the components in some reference system, of the plastic strain increment will be the same for all $d\underline{\underline{\sigma}}$, and only the norm of ε_{ij}^{p} will depend on the particular choice (see Fig. 3.129). This can be expressed by saying that plastic flow depends much more on the stress state that made it possible than on the stress increment that initiated it. The same is clearly true for the changes in the work-hardening parameters.

This property is made clear by a simple test. Suppose a thin tube subjected to a simple tensile test that takes it well into the plastic regime, where at each step the plastic deformation – and therefore the tensor h_{ij} – is in conformity with the load, that is, consists of axial elongation together with lateral contraction. While this is going on let an additional stress increment, obeying Eq. 3.210, be made that would correspond to a twisting of the tube: the immediate response will be, not the expected distortion by twisting, but an additional plastic *elongation* (accompanied by a lateral plastic contraction). Plastic distortion develops only if a strong enough

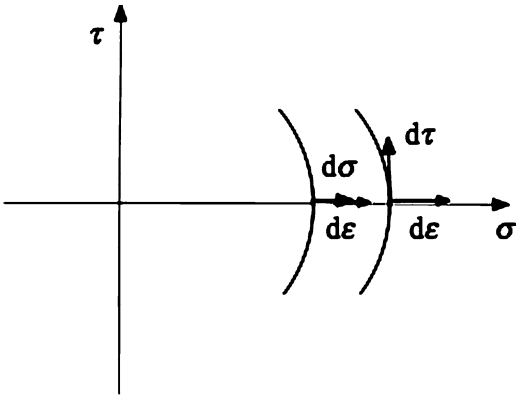


Fig. 3.130 A tube is first elongated by an axial stress increment; applying then a twist increment does not modify the direction of the strain increment, the deformation remaining an elongation

twist couple continues to be applied. Of course, the immediate *elastic* response to a suddenly-applied couple is a distortion: this shows the fundamental difference between elastic and plastic behaviour (Fig. 3.130).

This strong dependence of incremental plastic flow on the prevailing stress state rather than on the change in this is seen still more clearly in a particular case that we study next – that of the *standard material*.

3.5.2.3 Principle of Maximum Plastic Work: Standard Material

Consider the plastic deformation on a single slip system ($\underline{m}$, $\underline{n}$) of a crystal that obeys the Schmid law. We have already seen (Eq. 3.185) that with an appropriate orientation of the unit vectors $\underline{m}$, $\underline{n}$

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \mathrm{d}\gamma(\underline{m} \otimes \underline{n})^{\mathrm{S}} \quad \mathrm{d}\gamma > 0 \tag{3.213}$$

Moreover, in the plastic flow regime, the resolved shear stress $\tau_{\mathrm{r}} = \underline{m}.\underline{\underline{\sigma}}.\underline{n}$ is equal to the current critical resolved shear stress τ_{c}, and therefore

$$\underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \mathrm{d}\gamma\, \underline{\underline{\sigma}} : (\underline{m} \otimes \underline{n})^{\mathrm{S}} = \tau_{\mathrm{r}}\mathrm{d}\gamma = \tau_{\mathrm{c}}\mathrm{d}\gamma \tag{3.214}$$

Suppose now that a virtual stress $\underline{\underline{\sigma}}^*$ is applied while the plastic flow is proceeding; the virtual plastic work corresponding to the increment $\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}$ is given by

$$\underline{\underline{\sigma}}^* : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \mathrm{d}\gamma \underline{\underline{\sigma}}^* : (\underline{m} \otimes \underline{n})^{\mathrm{S}} = \tau_{\mathrm{r}}^* \mathrm{d}\gamma \tag{3.215}$$

If we take for $\underline{\underline{\sigma}}^*$ any stress state that does not violate the Schmid law for this slip system we have $\tau_{\mathrm{r}}^* \leq \tau_{\mathrm{c}}$, whence with $\mathrm{d}\gamma > 0$, $\tau_{\mathrm{r}}^* \mathrm{d}\gamma \leq \tau_{\mathrm{c}}\mathrm{d}\gamma$. It then follows that

$$\underline{\underline{\sigma}}^* : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} \leq \underline{\underline{\sigma}} : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} \quad \text{or} \quad (\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} \geq 0 \tag{3.216}$$

$\forall \underline{\underline{\sigma}}^*$ that does not violate the considered plasticity criterion.

If the crystal slips on two or more systems simultaneously and the Schmid law continues to hold, possibly with different critical resolved shear stresses in the different systems, the same reasoning applies and leads to the same result.

Finally, consider a polycrystalline volume element in which the Schmid law holds for each grain and in which the only plastic deformation mechanisms at work are intragranular slips such as those we have been considering so far – in particular, there is no intergranular sliding, which clearly would violate the Schmid law. Here again the same reasoning will hold and the same result be obtained.

Thus we have proved that for any (poly)crystalline material in which every grain obeys the Schmid law – expressed symbolically as $f(\underline{\underline{\sigma}}) \leq 0$ – and which deforms only by intragranular glide the following relation holds:

$$(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : d\underline{\underline{\varepsilon}}^p \geq 0, \quad \forall \underline{\underline{\sigma}}^*,\ f(\underline{\underline{\sigma}}^*) \leq 0 \text{ and } f(\underline{\underline{\sigma}}) = 0 \tag{3.217}$$

Note: The condition $f(\underline{\underline{\sigma}}) = 0$ can be removed since for $f(\underline{\underline{\sigma}}) < 0$, $d\underline{\underline{\varepsilon}}^p = 0$ and the inequality still holds in the broad sense.

The property just established can be expressed thus. For any crystalline material for which the Schmid law holds the stress state that initiates a given plastic flow is the one that develops the greatest amount of plastic work compared with all other virtual states that do not violate the plastic criterion.

The property can be given the status of a principle, the principle of maximum plastic work, or Hill's principle (Hill 1950). The class of materials to which it applies are called *standard materials* – that it does apply in any particular case should always be checked experimentally. There are two associated properties, the normality of the flow and the convexity of the criterion.

(a) *Normality rule.*

If we choose $\underline{\underline{\sigma}}^*$ to be indefinitely close to $\underline{\underline{\sigma}}$ and, like the latter, at the flow limit, the vector in stress space $(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*)$ will be tangent to the yield surface for $\underline{\underline{\sigma}}$, that is, normal to the gradient $\partial f/\partial \underline{\underline{\sigma}}$ of the yield function. Therefore

$$(\partial f/\partial \underline{\underline{\sigma}}) : (\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) = 0 \tag{3.218}$$

In the maximum-work relation (3.217) the roles of $\underline{\underline{\sigma}}$ and $\underline{\underline{\sigma}}^*$ can be interchanged, so we have both $(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : d\underline{\underline{\varepsilon}}^p \geq 0$ and $(\underline{\underline{\sigma}}^* - \underline{\underline{\sigma}}) : d\underline{\underline{\varepsilon}}^p \geq 0$; and therefore

$$(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : d\underline{\underline{\varepsilon}}^p = 0, \quad \forall \underline{\underline{\sigma}}^*,\ (\partial f/\partial \underline{\underline{\sigma}}) : (\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) = 0 \tag{3.219}$$

It follows that $d\underline{\underline{\varepsilon}}^p$ and $\partial f/\partial \underline{\underline{\sigma}}$ are collinear and therefore, if both are represented in the same space, $d\underline{\underline{\varepsilon}}^p$ also is normal to the yield surface; so we can write

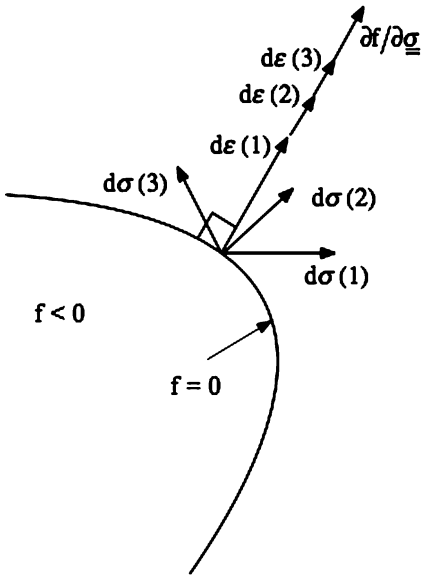


Fig. 3.131 Normality rule. The strain increments are normal to the yield surface

$$\mathrm{d}\underline{\underline{\varepsilon}}^{p} = (\partial \mathrm{f} / \partial \underline{\underline{\sigma}})\, \mathrm{d}\lambda,\ \mathrm{d}\lambda > 0 \tag{3.220}$$

where the increment $\mathrm{d}\lambda$ is a scalar, and moreover is positive (if we choose $\underline{\underline{\sigma}}^*$ inside the elastic domain and such that $(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*)$ and $(\partial \mathrm{f}/\partial \underline{\underline{\sigma}})$ are collinear and in the same sense, we have $(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} \geq 0$: thus $\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}$ is in the same sense as $(\partial \mathrm{f}/\partial \underline{\underline{\sigma}})$, that of the outward normal to the yield surface, since f is negative in the elastic domain and zero on the surface; so $\mathrm{d}\lambda$ is positive).

The *normality* of the flow (its direction normal to the yield surface) follows from that of $\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}$. The yield function plays the part of a *plastic potential* and determines, to within a positive scalar factor, the incremental plastic flow. This is a particular form of the general conclusion drawn in the last paragraph: that the *direction* of the plastic flow is determined by the stress and work-hardening states at that instant, and is now, in addition, constrained to be normal to the yield surface (Fig. 3.131).

(b) *Convexity rule.*

With $\underline{\underline{\sigma}}^*$ inside the elastic domain it follows from the maximum-work inequality together with the normality rule (with $\mathrm{d}\lambda$ positive) that

$$(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : (\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) \geq 0, \quad \forall \underline{\underline{\sigma}}^*,\ \mathrm{f}(\underline{\underline{\sigma}}^*) \leq 0 \tag{3.221}$$

This means that a situation such as that shown in Fig. 3.132 cannot arise: the yield surface must lie entirely to one side of any tangent plane, which in turn

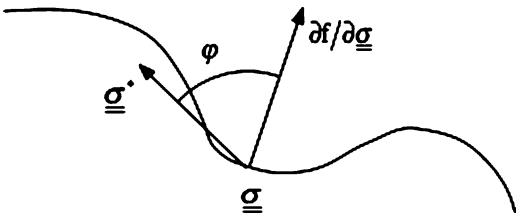


Fig. 3.132 A situation ($\varphi < \pi/2$) forbidden by the principle of maximum plastic work

means that it must be convex, as must be the yield function. This is the *convexity rule*.

It is easy to prove that the normality and convexity rules together imply the maximum work principle.

(c) *Consequences of the above*.

The two relations

$$\begin{cases} \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \underline{\underline{\mathrm{h}}}(\underline{\underline{\sigma}}, \underline{Y})(\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : \mathrm{d}\underline{\underline{\sigma}} \\ \qquad\qquad\qquad\qquad \text{if } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } (\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : \mathrm{d}\underline{\underline{\sigma}} \geq 0 \\ \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = (\partial \mathrm{f}/\partial \underline{\underline{\sigma}})\mathrm{d}\lambda, \ \ \mathrm{d}\lambda > 0 \end{cases} \tag{3.222}$$

together imply, provided that Hill's principle is satisfied, that the flow tensor $\underline{\underline{\mathrm{h}}}$ is proportional to $\partial f/\partial \underline{\underline{\sigma}}$, and the proportionality factor $\mathrm{g}(\underline{\underline{\sigma}}, \underline{Y})$ is positive. Therefore for a standard material

$$\begin{cases} \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \mathrm{g}(\underline{\underline{\sigma}}, \underline{Y}) \dfrac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} \left(\dfrac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} : \mathrm{d}\underline{\underline{\sigma}} \right) \text{ if } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } \dfrac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} : \mathrm{d}\underline{\underline{\sigma}} \geq 0 \\ \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = 0 \text{ if } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) < 0 \text{ or } \mathrm{f}(\underline{\underline{\sigma}}, \underline{Y}) = 0 \text{ and } \dfrac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} : \mathrm{d}\underline{\underline{\sigma}} \leq 0 \end{cases} \tag{3.223}$$

For perfect plasticity (standard material with no hardening) we have simply

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = (\partial \mathrm{f}/\partial \underline{\underline{\sigma}})\, \mathrm{d}\lambda, \text{ for arbitrary } \mathrm{d}\lambda > 0 \text{ and } \mathrm{f} = 0, \ \mathrm{df} = 0 \tag{3.224}$$

As would be expected, restricting the materials we consider to those that satisfy Hill's principle – a very wide class, however, which includes all those for which the Schmid law holds at the level of constituents crystals – has enabled us to reduce drastically the amount of information needed to give a complete characterisation of the plastic behaviour, the yield function then determining the flow almost completely.

It is tempting then to look for a corresponding result for work-hardening, giving the evolution of the parameters also determined by the yield function. This has been done within the framework of *generalised normality* for *generalised standard*

materials, in which the formalism of the thermodynamics of irreversible processes is used (Germain et al. 1983). Taking as work-hardening parameters the flux variables V_k and their duals A_k in the expression for a pseudo-dissipation potential, which is assumed to be convex with respect to the flux variables, we have simultaneously

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = (\partial \mathrm{f}/\partial \underline{\underline{\sigma}})\,\mathrm{d}\lambda \text{ and } -\mathrm{d}V_k = (\partial \mathrm{f}/\partial A_k)\mathrm{d}\lambda \tag{3.225}$$

We give some examples in the following paragraphs.

3.5.2.4 Particular Models

Having laid down a general framework for the description of elastoplastic behaviour we can now consider constructing some particular models, aiming to set up simple reference systems defined by the smallest possible number of parameters. It helps to assume standard (possibly generalised) materials, since the definition of these, whilst restrictive, enables the amount of experimental data needed to be greatly reduced. So as the better to illustrate the methodology, we shall most often use only the von Mises criterion since this is particularly simple; extension to other criteria does not in general present any difficulty except for the case of non-regular yield surfaces: this requires special treatment, which we shall not go into here.

The von Mises criterion (3.202 and 3.203) is stated in terms of J_2, the second symmetric invariant of the stress deviator; it is convenient to use instead a derived quantity which has the dimensions of stress and becomes the axial stress in the case of simple tension: this is the *von Mises equivalent stress*, defined by

$$\sigma_{\mathrm{eq}} = \bar{\sigma} = \sqrt{3\,\mathrm{J}_2} = \sqrt{(3/2)\underline{\underline{s}} : \underline{\underline{s}}} \tag{3.226}$$

The initial (that is, before any strain-hardening) von Mises yield function is then written simply

$$f_0(\underline{\underline{\sigma}}) = \bar{\sigma} - \sigma_0 \tag{3.227}$$

where σ_0 is the initial elastic limit in simple tension. Starting with this we consider in turn isotropic, linear kinematic, non-linear kinematic and isotropic-kinematic work-hardening, noting in particular the different responses to cyclic loads. We have

$$\begin{cases} \dfrac{\partial f_0}{\partial \underline{\underline{\sigma}}} = \dfrac{\partial \bar{\sigma}}{\partial \underline{\underline{\sigma}}} = \dfrac{\partial \bar{\sigma}}{\partial \underline{\underline{s}}} = \dfrac{3\underline{\underline{s}}}{2\bar{\sigma}} \\ \mathrm{d}_Y \mathrm{f}_0 = \dfrac{\partial \mathrm{f}_0}{\partial \underline{\underline{\sigma}}} : \mathrm{d}\underline{\underline{\sigma}} = \dfrac{3\underline{\underline{s}} : \mathrm{d}\underline{\underline{\sigma}}}{2\bar{\sigma}} = \mathrm{d}\bar{\sigma} \end{cases} \tag{3.228}$$

(a) *Isotropic work-hardening: Prandtl-Reuss model.*

Let R be the scalar work-hardening parameter; we find the simple yield function

$$\mathrm{f}(\underline{\underline{\sigma}}) = \mathrm{f}_0(\underline{\underline{\sigma}}) - R = \bar{\sigma} - R - \sigma_0 \tag{3.229}$$

For a standard material the plastic flow rule is

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \mathrm{d}\lambda \frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} = \frac{3\underline{\underline{s}}}{2\bar{\sigma}} \mathrm{d}\lambda \tag{3.230}$$

and $\mathrm{d}\lambda > 0$ for plastic flow, that is, for $\bar{\sigma} = R + \sigma_0$ and $\mathrm{d}\bar{\sigma} \geq 0$. This is the *Lévy*[29]*-Mises law*, which, as in the general case (Eq. 3.223), can be written explicitly as

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = g(\underline{\underline{\sigma}}, R) \frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} \left(\frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} : \mathrm{d}\underline{\underline{\sigma}} \right) = g(\underline{\underline{\sigma}}, R) \frac{3\underline{\underline{s}}}{2\bar{\sigma}} \mathrm{d}\bar{\sigma} \tag{3.231}$$

Consistently with the yield function, the positive-valued flow function $g(\underline{\underline{\sigma}}, R)$ depends only on J_2 and R, or, what is the same thing, on $\bar{\sigma}$ and R; and since in plastic flow R differs from $\bar{\sigma}$ only by the constant σ_0 this means that it depends only on $\bar{\sigma}$. We can therefore write

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \frac{3\mathrm{g}(\bar{\sigma})}{2\bar{\sigma}} \underline{\underline{s}} \mathrm{d}\bar{\sigma} \text{ for } \bar{\sigma} = R + \sigma_0 \text{ and } \mathrm{d}\bar{\sigma} > 0 \tag{3.232}$$

Also in plastic flow

$$\mathrm{df} = \mathrm{d}\bar{\sigma} - \mathrm{d}R = 0, \text{ so } \mathrm{d}R = \mathrm{d}\bar{\sigma} (> 0) \tag{3.233}$$

It is important to know which flux variable p is associated with the work-hardening variable R in the case of a generalised standard material. This is such that

$$\mathrm{d}p = -\frac{\partial f}{\partial R} \mathrm{d}\lambda = \mathrm{d}\lambda = \mathrm{g}(\bar{\sigma}) \mathrm{d}\bar{\sigma} \tag{3.234}$$

Moreover we have

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \frac{9}{4} \frac{\underline{\underline{s}} : \underline{\underline{s}}}{\bar{\sigma}^2} (\mathrm{d}\lambda)^2 \ (\mathrm{d}\lambda > 0) \tag{3.235}$$

[29] Maurice Lévy (1838–1910) was a French mathematician and engineer.

and so

$$\mathrm{d}\lambda = \mathrm{d}p = \sqrt{\frac{2}{3}\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} : \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}} \tag{3.236}$$

This is precisely the expression that generalises the concept of cumulative plastic strain to three dimensions, using an equivalence in the von Mises sense, as was done previously for the stresses. The factor 2/3 enables the incremental axial plastic strain to be found in the case of simple tension. We have

$$p = \int_0^t \sqrt{\frac{2}{3}\dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}}(\tau) : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}}(\tau)}\,\mathrm{d}\tau \tag{3.237}$$

If we add the elastic strain we have the *Prandtl*[30]*-Reuss*[31] *law* (for isotropic elasticity):

$$\begin{cases} \mathrm{d}\underline{\underline{\varepsilon}} = \frac{1+\nu}{E}\mathrm{d}\underline{\underline{\sigma}} - \frac{\nu}{E}\mathrm{d}(\mathrm{Tr}\underline{\underline{\sigma}})\underline{\underline{\delta}} + \frac{3g(\bar{\sigma})}{2\bar{\sigma}}\underline{\underline{s}}\,\mathrm{d}\bar{\sigma} \text{ if } \bar{\sigma} = R + \sigma_0, \mathrm{d}\bar{\sigma} \geq 0 \\ \mathrm{d}\underline{\underline{\varepsilon}} = \frac{1+\nu}{E}\mathrm{d}\underline{\underline{\sigma}} - \frac{\nu}{E}\mathrm{d}(\mathrm{Tr}\underline{\underline{\sigma}})\underline{\underline{\delta}} \text{ if } \bar{\sigma} < R + \sigma_0 \text{ or } \bar{\sigma} = R + \sigma_0, \mathrm{d}\bar{\sigma} < 0 \end{cases} \tag{3.238}$$

Given the values of the elastic constants the law is defined completely if $\mathrm{g}(\bar{\sigma})$ is known; and this of course can only be found experimentally. This is easily done if a tensile test curve $\varepsilon^{\mathrm{p}} = \varphi(\sigma)$ is available; we then have

$$\mathrm{d}\varepsilon^{\mathrm{p}} = \mathrm{d}p = \mathrm{g}(\bar{\sigma})\mathrm{d}\bar{\sigma} = \mathrm{g}(\sigma)\mathrm{d}\sigma = \varphi'(\sigma)\mathrm{d}\sigma \tag{3.239}$$

and so $\mathrm{g}(\sigma) = \varphi'(\sigma)$ or $\mathrm{g}(\bar{\sigma}) = \varphi'(\bar{\sigma})$. Hence in plastic flow

$$\mathrm{d}\underline{\underline{\varepsilon}} = \frac{1+\nu}{E}\mathrm{d}\underline{\underline{\sigma}} - \frac{\nu}{E}\mathrm{d}(\mathrm{Tr}\underline{\underline{\sigma}})\underline{\underline{\delta}} + \frac{3\varphi'(\bar{\sigma})}{2\bar{\sigma}}\underline{\underline{s}}\,\mathrm{d}\bar{\sigma} \text{ if } \bar{\sigma} = R + \sigma_0, \mathrm{d}\bar{\sigma} \geq 0 \tag{3.240}$$

Note: In general, the isotropic model is insufficient in dealing with cyclic loading. If what is imposed is a stress, then as soon as during the first cycle there is *shakedown* (elastic adaptation) without cyclic plastic deformation. If the strain is prescribed in a one-dimensional tensile test then for large enough amplitudes stabilisation can occur only progressively; the stable cycle is either elastic or without work-hardening in the plastic flow regimes if hardening is saturated in monotonic loading.

[30] Ludwig Prandtl (1875–1953) was a German physicist and engineer.

[31] Endre Reuss (1900–1968) was a Hungarian mathematician.

(b) *Linear kinematic work-hardening: Prager's model* (Prager and Hodge 1951)
We now introduce a tensorial work-hardening variable $\underline{\underline{X}}$, of order 2, writing the yield function as

$$\mathrm{f}(\underline{\underline{\sigma}}) = \mathrm{f}_0(\underline{\underline{\sigma}} - \underline{\underline{X}}) = \overline{(\sigma - X)} - \sigma_0 \tag{3.241}$$

where

$$\overline{(\sigma - X)} = \sqrt{\frac{3}{2}(\underline{\underline{s}} - \underline{\underline{x}}) : (\underline{\underline{s}} - \underline{\underline{x}})} \tag{3.242}$$

and $\underline{\underline{x}}$ is the deviatoric part of $\underline{\underline{X}}$. We then have

$$\frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} = \frac{\partial \mathrm{f}_0}{\partial(\underline{\underline{\sigma}} - \underline{\underline{X}})} = \frac{3}{2}\frac{\underline{\underline{s}} - \underline{\underline{x}}}{\overline{(\sigma - X)}} \tag{3.243}$$

For reasons given in Sect. 3.2.2.2(b) we have to assume that the kinematic hardening is linear, which is equivalent to assuming a fixed proportionality between the increments $\mathrm{d}\underline{\underline{X}}$ and $\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}$, say

$$\mathrm{d}\underline{\underline{X}} = k\,\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = k(\partial \mathrm{f}/\partial \underline{\underline{\sigma}})\mathrm{d}\lambda \tag{3.244}$$

From the generic form of the work-hardening law in plastic flow (Eq. 3.210) we have

$$\mathrm{d}X_{ij} = K_{ij}\frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} : \mathrm{d}\underline{\underline{\sigma}} \tag{3.245}$$

which is equivalent to taking $\underline{\underline{K}}$ proportional to $\partial \mathrm{f}/\partial \underline{\underline{\sigma}}$. Since we have $\partial \mathrm{f}/\partial \underline{\underline{X}} = -\partial \mathrm{f}/\partial \underline{\underline{\sigma}}$, the consistency condition can be written

$$1 - (\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : \underline{\underline{K}} = 0 \tag{3.246}$$

whence

$$\begin{cases} \underline{\underline{K}} = \dfrac{\partial \mathrm{f}/\partial \underline{\underline{\sigma}}}{(\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : (\partial \mathrm{f}/\partial \underline{\underline{\sigma}})} \\ \mathrm{d}\underline{\underline{X}} = \dfrac{\partial \mathrm{f}/\partial \underline{\underline{\sigma}}}{(\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : (\partial \mathrm{f}/\partial \underline{\underline{\sigma}})}(\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : \mathrm{d}\underline{\underline{\sigma}} = k(\partial \mathrm{f}/\partial \underline{\underline{\sigma}})\mathrm{d}\lambda \end{cases} \tag{3.247}$$

Thus with (3.242) and (3.243), which entails $(\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : (\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) = 3/2$, we get

$$k\mathrm{d}\lambda = \frac{\underline{\underline{s}} - \underline{\underline{x}}}{\overline{\sigma - X}} : \mathrm{d}\underline{\underline{\sigma}} \tag{3.248}$$

and

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \frac{3}{2k\sigma_0^2}(\underline{\underline{s}} - \underline{\underline{x}})\left[(\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}}\right] \quad \text{if } \overline{\sigma - X} = \sigma_0 \text{ and } (\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}} \geq 0 \tag{3.249}$$

For identification from a tensile test we have

$$(\underline{\underline{s}}) = \frac{\sigma}{3}[M]; (\underline{\underline{x}}) = (\underline{\underline{X}}) = \frac{X}{3}[M]; (\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}) = \frac{\mathrm{d}\varepsilon^{\mathrm{p}}}{2}[M]; [M] = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{3.250}$$

and

$$\overline{(\sigma - X)} = |\sigma - X| \quad \text{with } |\sigma - X| = \sigma_0 \text{ during flow} \tag{3.251}$$

which gives from (3.249)

$$\mathrm{d}\varepsilon^{\mathrm{p}} = \frac{3}{2k\sigma_0^2}\left[\frac{2(\sigma - X)}{3}\right]^2 \mathrm{d}\sigma = \frac{2}{3}\frac{\mathrm{d}\sigma}{k} \tag{3.252}$$

We thus get a constant work-hardening modulus $C = 3k/2$; so finally Prager's[32] law can be written as follows:

$$\begin{cases} \begin{cases} \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \dfrac{9}{4C\sigma_0^2}(\underline{\underline{s}} - \underline{\underline{x}})\left[(\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}}\right] \\ \qquad\qquad \text{if } \overline{(\sigma - X)} = \sigma_0 \text{ and } (\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}} \geq 0 \\ \mathrm{d}\underline{\underline{X}} = \dfrac{3}{2\sigma_0^2}(\underline{\underline{s}} - \underline{\underline{x}})\left[(\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}}\right] \end{cases} \\ \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \mathrm{d}\underline{\underline{X}} = 0 \text{ if } \overline{(\sigma - X)} < \sigma_0 \text{ or } \overline{(\sigma - X)} = \sigma_0 \text{ and } (\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}} \leq 0 \end{cases} \tag{3.253}$$

This model, constructed so as to take account of the Bauschinger effect, cannot give any representation of the ratchet effect or the relaxation of the mean stress during cyclic loading: there is always some shakedown and the stable cycle is reached either immediately (*e.g.* when the cycle is tension/compression) or progressively (*e.g.* when it is tension-shear).

[32] William Prager (1903–1980) was a German-born US applied mathematician.

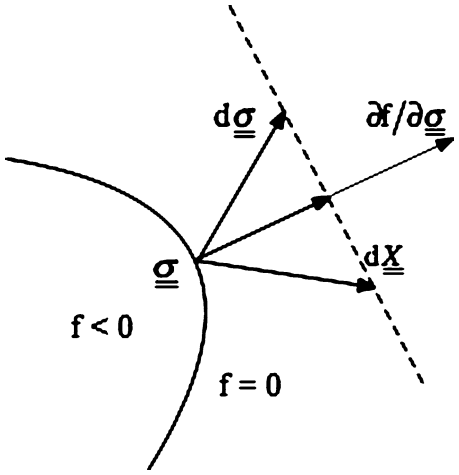


Fig. 3.133 Representation of the relation $\mathrm{df} = 0$ for $\mathrm{f}\left(\underline{\underline{\sigma}} - \underline{\underline{X}}\right)$

Note: For a generalised standard material the flux variable $\underline{\underline{\alpha}}$ associated with the work-hardening variable $\underline{\underline{X}}$ is such that

$$\begin{cases} \begin{cases} \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \dfrac{9}{4C\sigma_0^2}(\underline{\underline{s}} - \underline{\underline{x}})\left[(\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}}\right] \\ \qquad\qquad\qquad\qquad \text{if } \overline{(\sigma - X)} = \sigma_0 \text{ and } (\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}} \geq 0 \\ \mathrm{d}\underline{\underline{X}} = \dfrac{3}{2\sigma_0^2}(\underline{\underline{s}} - \underline{\underline{x}})\left[(\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}}\right] \end{cases} \\ \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \mathrm{d}\underline{\underline{X}} = 0 \text{ if } \overline{(\sigma - X)} < \sigma_0 \text{ or } \overline{(\sigma - X)} = \sigma_0 \text{ and } (\underline{\underline{s}} - \underline{\underline{x}}) : \mathrm{d}\underline{\underline{\sigma}} \leq 0 \end{cases} \tag{3.254}$$

Now $-\partial \mathrm{f}/\partial \underline{\underline{X}} = \partial \mathrm{f}/\partial \underline{\underline{\sigma}}$, and since the material is standard $\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = (\partial \mathrm{f}/\partial \underline{\underline{\sigma}})\mathrm{d}\lambda$. The variable $\underline{\underline{\alpha}}$ can thus be identified with the plastic strain $\underline{\underline{\varepsilon}}^{\mathrm{p}}$ and there is a linear relation between $\underline{\underline{X}}$ and $\underline{\underline{\alpha}}$ (the pseudo-dissipation potential is quadratic in $\underline{\underline{\alpha}}$).

For all kinematic hardening criteria which have a yield function of the form $f(\underline{\underline{\sigma}} - \underline{\underline{X}})$ we have, in plastic flow,

$$\mathrm{df} = (\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : \mathrm{d}\underline{\underline{\sigma}} + (\partial \mathrm{f}/\partial \underline{\underline{X}}) : \mathrm{d}\underline{\underline{X}} = (\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : (\mathrm{d}\underline{\underline{\sigma}} - \mathrm{d}\underline{\underline{X}}) = 0 \tag{3.255}$$

whence

$$(\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : \mathrm{d}\underline{\underline{\sigma}} = (\partial \mathrm{f}/\partial \underline{\underline{\sigma}}) : \mathrm{d}\underline{\underline{X}} \tag{3.256}$$

which is represented geometrically in Fig. 3.133. Prager's model makes $\mathrm{d}\underline{\underline{X}}$ collinear with $\partial \mathrm{f}/\partial \underline{\underline{\sigma}}$; to overcome the restrictions of this linear kinematic model,

others have been proposed – *e.g.* by Ziegler and by Mroz, see Lemaitre and Chaboche (1985) – which make different assumptions concerning the direction of $\mathrm{d}\underline{\underline{X}}$ in load space. We now consider an example.

(c) *Nonlinear kinematic hardening*.

One way to incorporate the ratchet effect into a kinematic model is to modify the relation between $\mathrm{d}\underline{\underline{X}}$ and $\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}$ so as to represent nonlinear work-hardening, in particular by introducing a supplementary dependence of $\underline{\underline{X}}$ on the cumulative plastic deformation p which is growing in the course of cycling, *e.g.*

$$\mathrm{d}\underline{\underline{X}} = (2/3)\, C\ \mathrm{d}\underline{\underline{\varepsilon}}^{p} - \gamma \underline{\underline{X}}\, \mathrm{d}p, \gamma > 0 \tag{3.257}$$

The back stress term $(-\gamma \underline{\underline{X}}\, \mathrm{d}p)$ expresses a transient memory of the loading path and makes it possible to cause the hardening modulus to decrease in the course of the flow. If the same dissipative potential is used as in the Prager model with $\underline{\underline{X}} = (2C/3)\ \underline{\underline{\alpha}}$, the material will no longer be generalised standard since then the relation $\mathrm{d}\underline{\underline{\alpha}} = \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}}$ would always hold. The generalised framework can be preserved by dissociating the yield function f from the flow and introducing a flow potential F distinct from the yield function: this is a case of "non-associated plasticity". Putting

$$\mathrm{F} = \overline{(\underline{\underline{\sigma}} - \underline{\underline{X}})} - \sigma_0 + \frac{3\gamma}{4C} \underline{\underline{X}} : \underline{\underline{X}} \tag{3.258}$$

we get

$$\begin{cases} \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \dfrac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} \mathrm{d}\lambda = \dfrac{\partial \mathrm{F}}{\partial \underline{\underline{\sigma}}} \mathrm{d}\lambda \\ \mathrm{d}\underline{\underline{\alpha}} = -\dfrac{\partial \mathrm{F}}{\partial \underline{\underline{X}}} \mathrm{d}\lambda = -\dfrac{\partial \mathrm{f}}{\partial \underline{\underline{X}}} \mathrm{d}\lambda - \dfrac{3\gamma}{2C} \underline{\underline{X}} \mathrm{d}\lambda \\ \quad = \dfrac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} \mathrm{d}\lambda - \dfrac{3\gamma}{2C} \underline{\underline{X}} \mathrm{d}\lambda = \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} - \dfrac{3\gamma}{2C} \underline{\underline{X}} \mathrm{d}p \end{cases} \tag{3.259}$$

since $\mathrm{d}p = \mathrm{d}\lambda$; finally, as expected,

$$\mathrm{d}\underline{\underline{X}} = \frac{2C}{3} \mathrm{d}\underline{\underline{\alpha}} = \frac{2C}{3} \mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} - \gamma \underline{\underline{X}} \mathrm{d}p \tag{3.260}$$

We can derive the following properties from this model:

1. The yield surface is moved in stress space during the flow, but remains always within a limit surface whose equation is

$$\bar{\sigma} = \sigma_0 + \frac{C}{\gamma} \tag{3.261}$$

2. Using (1) we can describe the ratchet effect associated with an unsymmetrical applied tension-compression load when stress is imposed and the relaxation of the mean stress when strain is controlled, when the mean strain is not zero.
3. There is neither cyclic hardening nor cyclic softening.

(d) *Combined isotropic and kinematic hardening.*

We take the yield function to be of the form

$$f(\underline{\underline{\sigma}}, R, \underline{\underline{X}}) = \overline{(\underline{\underline{\sigma}} - \underline{\underline{X}})} - R - \sigma_0 \tag{3.262}$$

involving kinematic (governed by the variation in $\underline{\underline{X}}$) and isotropic (governed by that of R, usually a function of the cumulative plastic strain p) work-hardening simultaneously. Without going into details, we can say that the form of dependence of $\underline{\underline{X}}$ on $\underline{\underline{\varepsilon}}^{\mathrm{p}}$ and p, used for nonlinear kinematic work-hardening, can be kept and generalised by replacing the constants C and γ by functions $C(p)$ and $\gamma(p)$. This brings in, in addition to the previously discussed properties of nonlinear kinematic work-hardening, the possibility of representing cyclic hardening and softening when there is shakedown.

A final refinement can be derived from observations of the dependence of the stable cycle on the previous history of the loading: for example, the fact that different cycles are reached according as this is one of increasing or decreasing levels of cyclic loading. An extra variable can be introduced to carry the memory of previous loadings: for example, of maximum plastic strains reached. By such means a fairly complete model of work-hardening and flow can be constructed, which can moreover be generalised by application of the same approach to other criteria than the von Mises one, in particular to non-isotropic criteria such as that of Hill.

In conclusion we must emphasise that all the models we have been discussing at least make use of the principle of maximum plastic work, *i.e.*, are restricted to standard materials. The physical basis for this is sound enough in the case of crystalline materials, but the situation is not at all the same for other materials, particularly soils, for which different models have to be developed.

3.5.3 Introduction to the Plastic Design of Structures

3.5.3.1 Outline

Calculation of the elastoplastic response of structures presents many difficulties resulting from the incremental and multi-branched nature of the constitutive equations, even for the simple schemes for work-hardening and flow that have been developed in the previous paragraphs. The *incremental* nature of the laws requires a step-by-step treatment, following the loading path; the variational methods needed for the development of efficient numerical procedures are based on *rate* formulations for which the relevant extremal theorems derive from the normality and convexity

properties of standard (simple or generalised) materials. Their *multibranched* nature requires the zones in elastic and in plastic regime, and the evolution of the boundaries of these zones, to be determined at each step.

We do not aim at developing here fully general methods which can be found in specialist books (*e.g.* Prager and Hodge 1951; Koiter 1960; Save et al. 1997; Hill 1998); to illustrate some of the problems that arise and to give some simple computational tools that are useful for studying the mechanics of heterogeneous materials, we limit our treatment to two types of situation in which important simplifications can be achieved:

– proportional loadings of structures
– limit yield analysis of rigid-perfectly plastic solids

Finally, we give introductory elements for dealing with nonlinear homogenisation and the plasticity of heterogeneous materials.

3.5.3.2 Proportional Loadings of Structures

We are concerned here with structures subjected to external monotonic forces, proportional to a single parameter. In some conditions we can assume that at all points of the structure the stress tensor also, either strictly or to a good enough approximation, is proportional to this parameter, the proportionality factor being a monotonic scalar function of (kinematic!) time t. With these restrictions the incremental equations can be integrated to give finite relations between the variables $\underline{\underline{\sigma}}$ and $\underline{\underline{\varepsilon}}^{\mathrm{p}}$, as in the case of elastic behaviour.

For simplicity, consider a Prandtl-Reuss model for local behaviour. The incremental equation for the plastic strain is

$$\mathrm{d}\underline{\underline{\varepsilon}}^{\mathrm{p}} = \frac{3\varphi'(\bar{\sigma})}{2\bar{\sigma}}\underline{\underline{s}}\,\mathrm{d}\bar{\sigma} \text{ if } \bar{\sigma} = R + \sigma_0, \mathrm{d}\bar{\sigma} \geq 0 \tag{3.263}$$

where $\underline{\underline{\varepsilon}}^{\mathrm{p}} = \varphi(\sigma)$ is the equation of the tensile monotonic stress-strain curve, for $\sigma \geq \sigma_0$. The assumption of radial loading entails the constancy of the tensor $\underline{\underline{s}}/\underline{\underline{\sigma}}$; if we write

$$\underline{\underline{\sigma}}(t) = \alpha(t)\underline{\underline{\Sigma}} \tag{3.264}$$

where $\underline{\underline{\Sigma}}$ is constant, then (3.264) still holds for the deviators, namely

$$\underline{\underline{s}}(t) = \alpha(t)\underline{\underline{S}} \tag{3.265}$$

and $\underline{\underline{s}} : \underline{\underline{s}} = \alpha^2(t)\underline{\underline{S}} : \underline{\underline{S}}$; so

$$\bar{\sigma} = |\alpha(t)| \sqrt{\frac{3}{2}\underline{\underline{S}} : \underline{\underline{S}}} \tag{3.266}$$

$$\frac{\underline{\underline{s}}}{\bar{\sigma}} = \frac{\alpha(t)\underline{\underline{S}}}{|\alpha(t)|\sqrt{\frac{3}{2}\underline{\underline{S}} : \underline{\underline{S}}}} = \frac{\pm\underline{\underline{S}}}{\sqrt{\frac{3}{2}\underline{\underline{S}} : \underline{\underline{S}}}} \text{ (constant tensor)} \tag{3.267}$$

Integration of (3.263) now gives the Hencky[33]-Mises law ("J_2 deformation theory"):

$$\underline{\underline{\varepsilon}}^{\mathrm{p}} = \int_{\sigma_0}^{\sigma_0+R} \frac{3\varphi'(\bar{\sigma})}{2\bar{\sigma}} \underline{\underline{s}}\, \mathrm{d}\bar{\sigma} = \frac{3\underline{\underline{s}}}{2\bar{\sigma}} \int_{\sigma_0}^{\sigma_0+R} \varphi'(\bar{\sigma})\mathrm{d}\bar{\sigma} = \left[\frac{3\underline{\underline{s}}}{2\bar{\sigma}}\varphi(\bar{\sigma})\right]_{\sigma_0}^{\sigma_0+R} = \frac{3\underline{\underline{s}}}{2\bar{\sigma}}\varphi(\bar{\sigma}) \tag{3.268}$$

Thus, as expected, we have derived a relation between the values of $\underline{\underline{\varepsilon}}^{\mathrm{p}}$ and $\underline{\underline{s}}$ at any instant, which involves the strain-dependent *secant plastic modulus* $\frac{2}{3}\bar{\sigma}/\varphi(\bar{\sigma})$. If we define an equivalent von Mises plastic strain $\bar{\varepsilon}^{\mathrm{p}}$ by

$$\bar{\varepsilon}^{\mathrm{p}} = \sqrt{(2/3)\underline{\underline{\varepsilon}}^{\mathrm{p}} : \underline{\underline{\varepsilon}}^{\mathrm{p}}} \tag{3.269}$$

we can even obtain a "universal " equation, valid for all monotonic radial loadings, that is, for "simple" or "proportional" loadings:

$$(\bar{\varepsilon}^{\mathrm{p}})^2 = \frac{2}{3}\underline{\underline{\varepsilon}}^{\mathrm{p}} : \underline{\underline{\varepsilon}}^{\mathrm{p}} = \frac{2}{3}\frac{9}{4}\frac{\underline{\underline{s}} : \underline{\underline{s}}}{\bar{\sigma}^2}\varphi^2(\bar{\sigma}) \tag{3.270}$$

so

$$\bar{\varepsilon}^{\mathrm{p}} = \varphi(\bar{\sigma}) \tag{3.271}$$

which is the same expression for the tensile curve as we started with, now in terms of equivalent quantities. This enables us to give a wider justification to the use of a "work-hardening exponent" which, starting from a tensile curve expression

$$\sigma = \sigma_0 + K(\varepsilon^{\mathrm{p}})^n \tag{3.272}$$

where n is the work-hardening exponent, keeps the same meaning for equivalent quantities in monotonic radial loading:

$$\bar{\sigma} = \sigma_0 + K(\bar{\varepsilon}^{\mathrm{p}})^n \tag{3.273}$$

[33] Heinrich Hencky (1885–1951) was a German engineer.

Note: The definition of equivalence for any criterion must take account of the particular criterion; thus for Tresca's, the best choice is

$$\bar{\sigma} = \sigma_1 - \sigma_3 \quad \bar{\varepsilon}^{\mathrm{p}} = \varepsilon_1^{\mathrm{p}} - \varepsilon_3^{\mathrm{p}} \quad \text{if } \sigma_1 > \sigma_2 > \sigma_3 \tag{3.274}$$

and for Hill's ($\underline{\underline{s}} : \mathbf{B} : \underline{\underline{s}} = \sigma_0^2$), it is

$$\bar{\sigma} = (\underline{\underline{s}} : \mathbf{B} : \underline{\underline{s}})^{1/2} \tag{3.275}$$

This situation of local proportional loading does not apply at all points of the structure unless the initial state was stress-free and the elastic strains can be neglected: these are necessary conditions for the structure problem to be solvable directly.

3.5.3.3 Introduction to Limit Analysis

Considering only standard, rigid-perfectly plastic media, we now attempt to set bounds to the conditions of incipient plastic flow. The limit analysis is based on two extremal theorems, with the problem stated in terms of velocities $\underline{v}^{\mathrm{g}}$ prescribed on a part S_{v} of the external surface S of the medium V, body forces $\underline{F}$ in V and surface forces $\underline{T}^{\mathrm{g}}$ prescribed on S_{T}, the complement of S_{v} on this surface.

Static approach. $\underline{\underline{\sigma}}^*$ is said to be *licit* if it is statically admissible for the problem under consideration ($\mathrm{div}\underline{\underline{\sigma}}^* + \underline{F} = 0$, $\underline{\underline{\sigma}}^* = \underline{\underline{\sigma}}^{*\mathrm{T}}$, $\underline{\underline{\sigma}}^*.\underline{n} = \underline{T}^{\mathrm{g}}$ on S_{T}) and plastically admissible ($\mathrm{f}(\underline{\underline{\sigma}}^*) \leq 0$ in V). Then, according to the principle of maximal work, with $\underline{\underline{\sigma}}$, $\underline{\underline{\varepsilon}}$ and $\underline{v}$ referring to a solution to the problem

$$(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : \dot{\underline{\underline{\varepsilon}}} \geq 0 \quad \Rightarrow \quad \int_V (\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : \dot{\underline{\underline{\varepsilon}}} \, \mathrm{d}V \geq 0 \tag{3.276}$$

with

$$\int_V (\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*) : \dot{\underline{\underline{\varepsilon}}} \, \mathrm{d}V = \int_S \underline{v}.(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*).\underline{n} \, \mathrm{d}S = \int_{S_{\mathrm{v}}} \underline{v}^{\mathrm{g}}.(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^*).\underline{n} \, \mathrm{d}S \tag{3.277}$$

whence

$$\mathrm{H}(\underline{\underline{\sigma}}^*) = \int_{S_{\mathrm{v}}} \underline{v}^{\mathrm{g}}.\underline{\underline{\sigma}}^*.\underline{n} \, \mathrm{d}S \quad \Rightarrow \quad \mathrm{H}(\underline{\underline{\sigma}}^*) \leq \mathrm{H}(\underline{\underline{\sigma}}) \tag{3.278}$$

This, called the *lower bound theorem*, is the first of our extremal theorems; it states that among all licit stress fields a solution to the problem maximises the functional $\mathrm{H}(\underline{\underline{\sigma}}^*)$.

Kinematic approach. $\underline{v}'$ is *licit* for the problem if it is kinematically admissible (*i.e.,* is continuous, and $\underline{v}' = \underline{v}^{\mathrm{g}}$ on S_{v}) and plastically admissible (we can associate with it $\underline{\underline{\sigma}}'$, $\mathrm{f}(\underline{\underline{\sigma}}') = 0$, with $\underline{\underline{\dot{\varepsilon}}}' = \dot{\lambda}(\partial \mathrm{f}/\partial \underline{\underline{\sigma}}), \dot{\lambda} > 0$). Then, even if $\underline{\underline{\sigma}}'$ is not unique, $\underline{\underline{\sigma}}' : \underline{\underline{\dot{\varepsilon}}}'$ is well defined; in fact, for $\underline{\underline{\sigma}}'_1 \neq \underline{\underline{\sigma}}'_2$, we have $(\underline{\underline{\sigma}}'_1 - \underline{\underline{\sigma}}'_2) : \underline{\underline{\dot{\varepsilon}}}' \geq 0$ and $(\underline{\underline{\sigma}}'_2 - \underline{\underline{\sigma}}'_1) : \underline{\underline{\dot{\varepsilon}}}' \geq 0$. So $(\underline{\underline{\sigma}}'_1 - \underline{\underline{\sigma}}'_2) : \underline{\underline{\dot{\varepsilon}}}' = 0$ and $\underline{\underline{\sigma}}'_1 : \underline{\underline{\dot{\varepsilon}}}' = \underline{\underline{\sigma}}'_2 : \underline{\underline{\dot{\varepsilon}}}'$.

Further,

$$\underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\sigma}}' : \underline{\underline{\dot{\varepsilon}}}' = \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}}' + (\underline{\underline{\sigma}} - \underline{\underline{\sigma}}') : \underline{\underline{\dot{\varepsilon}}}' \text{ with } (\underline{\underline{\sigma}} - \underline{\underline{\sigma}}') : \underline{\underline{\dot{\varepsilon}}}' \leq 0 \tag{3.279}$$

whence

$$\underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\sigma}}' : \underline{\underline{\dot{\varepsilon}}}' \leq \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}}' \text{ and } \int_V (\underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\sigma}}' : \underline{\underline{\dot{\varepsilon}}}')\mathrm{d}V \leq \int_V \underline{\underline{\sigma}} : (\underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\dot{\varepsilon}}}')\mathrm{d}V \tag{3.280}$$

with

$$\begin{aligned} \int_V \underline{\underline{\sigma}} : (\underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\dot{\varepsilon}}}')\mathrm{d}V &= \int_V \underline{F}.(\underline{v} - \underline{v}')\mathrm{d}V + \int_S (\underline{v} - \underline{v}').\underline{\underline{\sigma}}.\underline{n}\mathrm{d}S \\ &= \int_V \underline{F}.(\underline{v} - \underline{v}')\mathrm{d}V + \int_{S_{\mathrm{T}}} (\underline{v} - \underline{v}').\underline{T}^{\mathrm{g}}\mathrm{d}S \end{aligned} \tag{3.281}$$

whence, from (3.280) and (3.281)

$$\int_V (\underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \underline{\underline{\sigma}}' : \underline{\underline{\dot{\varepsilon}}}')\mathrm{d}V \leq \int_V \underline{F}.(\underline{v} - \underline{v}')\mathrm{d}V + \int_{S_{\mathrm{T}}} (\underline{v} - \underline{v}').\underline{T}^{\mathrm{g}}\mathrm{d}S \tag{3.282}$$

and finally

$$\mathrm{L}(\underline{v}') = \int_V \underline{\underline{\sigma}}' : \underline{\underline{\dot{\varepsilon}}}'\mathrm{d}V - \int_V \underline{F}.\underline{v}'\mathrm{d}V - \int_{S_{\mathrm{T}}} \underline{v}'.\underline{T}^{\mathrm{g}}\mathrm{d}S \quad \Rightarrow \quad \mathrm{L}(\underline{v}) \leq \mathrm{L}(\underline{v}') \tag{3.283}$$

This is our second extremal theorem, called the *upper bound theorem*; it states that among all licit velocity fields a solution to the problem minimises the functional $\mathrm{L}(\underline{v}')$.

As the solution is both statically and kinematically admissible, it turns out that $\mathrm{H}(\underline{\underline{\sigma}}) = \mathrm{L}(\underline{v})$; so the two theorems put the solution between upper and lower bounds:

$$\mathrm{H}(\underline{\underline{\sigma}}^*) \leq \mathrm{H}(\underline{\underline{\sigma}}) = \mathrm{L}(\underline{v}) \leq \mathrm{L}(\underline{v}') \tag{3.284}$$

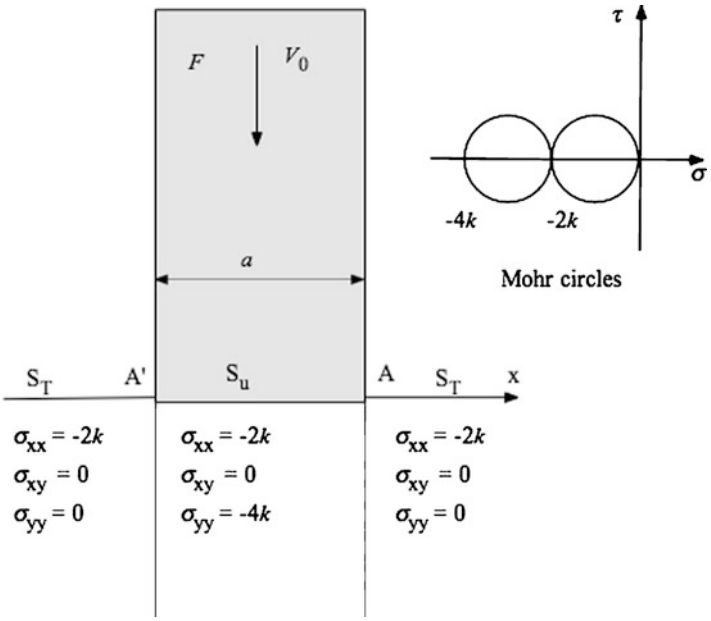


Fig. 3.134 Piecewise uniform stress field with three blocks and associated Mohr circles

Note: These theorems form the basis for the derivation of limit loads applied to structures described globally by load parameters Q_i and their associated kinematical parameters q_i (see hereafter the example of punching). They can also be used to derive bounds for the effective behaviour of plastic heterogeneous media (see Sect. 3.5.3.6).

3.5.3.4 An Example (Static Approach): Flat Punch on a Semi-Infinite Body (Plane Strain, Isotropic Criterion)

First we choose a stress field of three blocks (Fig. 3.134), piecewise uniform, everywhere at the flow limit (Mohr circle of radius k, the flow limit in simple shear). We have, per unit of thickness

$$\begin{cases} \int_{S_v} \underline{v}^g . \underline{\underline{\sigma}}^* . \underline{n} \, \mathrm{d}x = -\int_{AA'} \sigma_{yy} V_0 \, \mathrm{d}x = 4kaV_0 \\ \int_{S_v} \underline{v}^g . \underline{\underline{\sigma}} . \underline{n} \, \mathrm{d}x = FV_0 \end{cases} \tag{3.285}$$

and then $F \geq 4ka$.

With a piecewise uniform stress field of five blocks (Fig. 3.135) we find

$$\sigma^*{}_{yy}\left(AA'\right) = -2k(1+\sqrt{2}), \text{ whence } \mathrm{F} \geq 2(1+\sqrt{2})ka \tag{3.286}$$

but it is necessary, as is possible, to complete the field in the lower zones.

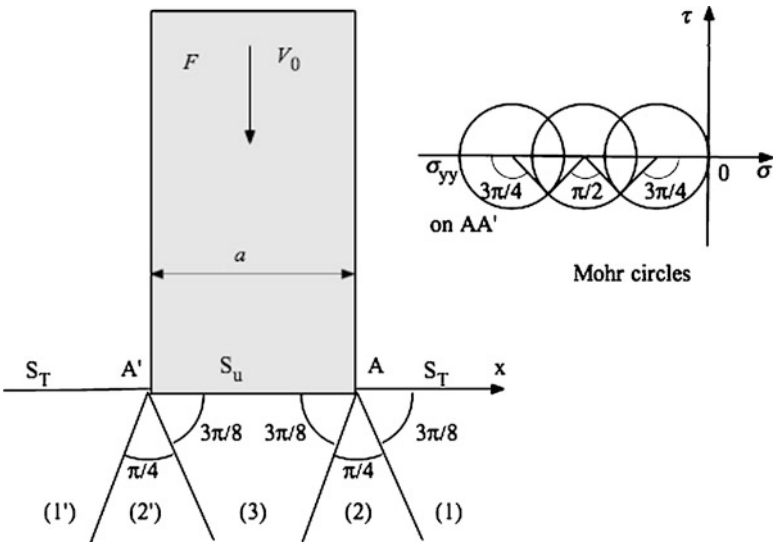


Fig. 3.135 Piecewise uniform stress field with five blocks and associated Mohr circles

Generalising the result to $(2n+1)$ blocks (Fig. 3.136), provided that we have completed the field throughout the whole of the half-space, which becomes more and more difficult, we have

$$\sigma_{yy}\left(AA'\right) = -2k[1 + n\sin(\pi/2n)] \text{ and } F \geq 2ka[1 + n\sin(\pi/2n)] \quad (3.287)$$

As $n \to \infty$ the inequality for F tends to $F \geq 2(1+\pi/2)ka$, with "Prandtl fans" in A and A'. We have not proved that the field can be extended throughout the body, but this can be established by a different method.

In order to derive an upper bound, a kinematic treatment is needed. A convenient and simple method for that has been developed for easy application to various problems of metal forming: the method of rigid blocks. It is presented as an exercise in Volume III and applied to the same problem of the flat punch. The more general and powerful method of slip lines is briefly developed now.

3.5.3.5 Outline of the Method of Slip Lines (Characteristics)

(a) This method leads to the geometrical construction of a network of curves from which the velocities and stresses in the deformed (or possibly deformed) zones can be calculated; the calculation will use the extremal theorems we proved in the previous section, most often the upper bound. The method is based on the fact that, for free incipient plastic flow *in plane strain conditions* of a standard rigid-perfectly plastic medium of von Mises or Tresca type (for which the criteria are equivalent under these conditions), the stress fields in the deformed zones can be determined independently of the velocity fields.

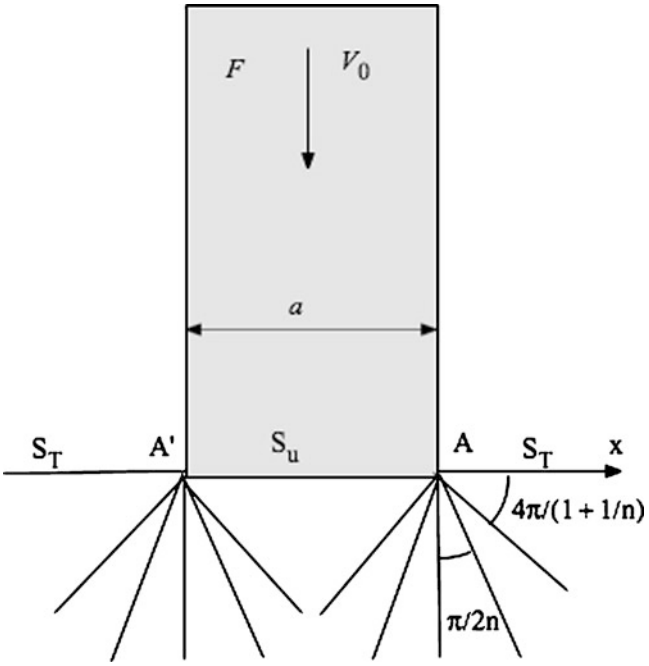


Fig. 3.136 Field with 2n + 1 blocks

It is easy to show that for these criteria – as indeed for all isotropic criteria that are independent of hydrostatic pressure – the principal stress normal to the (*x*,*y*) plane is intermediate. Let σ_1 be the maximum principal stress ($\sigma_1 > \sigma_2$); during flow

$$\begin{cases} \sigma_1 - \sigma_2 = 2k = \sqrt{(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2} \\ \dot{\varepsilon}_1 = -\dot{\varepsilon}_2 = \dot{\lambda} \text{ with } \dot{\lambda} > 0 \text{ arbitrary} \\ \sigma_{13} = \sigma_{23} = 0 \text{ (isotropy)} \end{cases} \tag{3.288}$$

so that the equilibrium equations in the deformed zones can be written

$$\sigma_{11,1} + \sigma_{12,2} = 0, \quad \sigma_{12,1} + \sigma_{22,2} = 0, \quad (\sigma_{33,3} = 0) \tag{3.289}$$

This gives three equations for the three unknowns $\sigma_{11}, \sigma_{12}, \sigma_{22}$; it is helpful to change the variables as follows (Fig. 3.137):

$$\begin{aligned} -p &= (\sigma_1 + \sigma_2)/2 = (\sigma_{11} + \sigma_{22})/2, \quad \theta = (\underline{Ox}, \underline{\sigma_1}) \\ R &= \sqrt{(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2} \end{aligned} \tag{3.290}$$

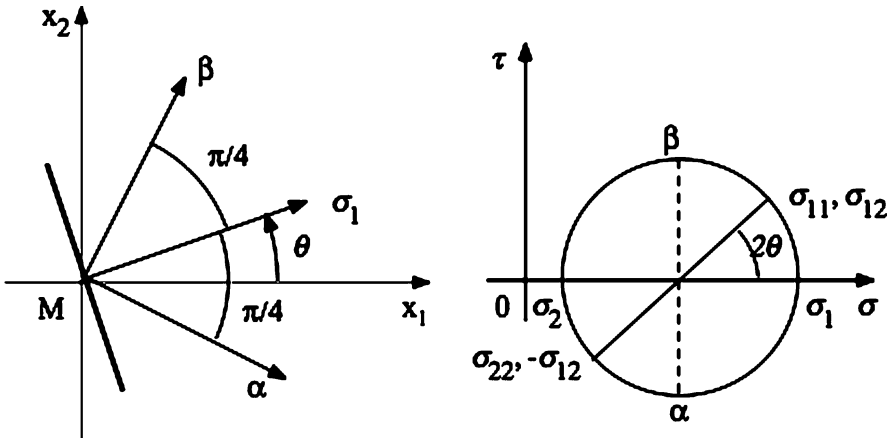


Fig. 3.137 Slip lines (notations)

$-p$ is the abscissa of the centre of the Mohr circle; from this circle we have

$$\sigma_{11} = -p + R\cos 2\theta, \quad \sigma_{12} = R\sin 2\theta, \quad \sigma_{22} = -p - R\cos 2\theta \tag{3.291}$$

and the equations for the stresses in the (assumed) deformed zones are

$$\begin{cases} R = k \\ -p_{,1} - 2k\theta_{,1}\sin 2\theta + 2k\theta_{,2}\cos 2\theta = 0 \\ -p_{,2} + 2k\theta_{,1}\cos 2\theta + 2k\theta_{,2}\sin 2\theta = 0 \end{cases} \tag{3.292}$$

Now consider the cross-sections M_α, M_β inclined respectively at $-\pi/4$ and $+\pi/4$ to σ_1 (considered as a vector): these support the maximum shear stress k at M. With these we can define, for the entire deformed zone, an orthogonal curvilinear lattice which we can then use as a co-ordinate system: in this the equilibrium equations become simply

$$-p_{,\alpha} - 2k\theta_{,\alpha} = 0, \quad -p_{,\beta} + 2k\theta_{,\beta} = 0 \tag{3.293}$$

as can be checked by putting $\theta = \pi/4$ in the original equations, so that the lines $M\alpha$, $M\beta$ coincide with the axes Mx_1, Mx_2. Integration is immediate, giving the *Hencky relations*:

$$p + 2k\theta = \text{const. on an } \alpha \text{ - line}, \quad p - 2k\theta = \text{const. on a } \beta \text{ - line} \tag{3.294}$$

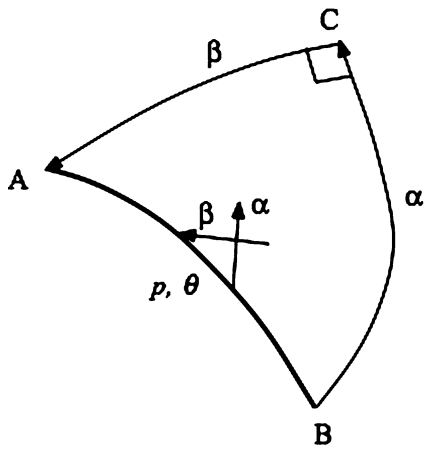


Fig. 3.138 Determination of the α and β lines in a deformed zone from known p and θ values on AB

Note: The α and β lines are the characteristic lines of a quasi-linear system of hyperbolic partial differential equations – it is for this reason that they are real. Using these, the values of p and θ, and thence of σ_{11}, σ_{12}, σ_{22}, at all points in a deformed zone can be calculated from a knowledge of their values on the boundary.

(b) *Examples*

(i) In Fig. 3.138 p and θ are known on the arc AB.

- the directions of the α, β lines can be drawn at every point on AB
- this having been done for two neighbouring points on AB, the values of p and θ at an internal point can be deduced, using the Hencky equations (finite differences)
- the process is repeated step by step in all the region ABC bounded by AB, the α line through B and the β line though A.

(ii) Two slip lines through a point are known and also the value of p at that point (Fig. 3.139)

- the value of θ and thus of p is deduced on PA and PB
- starting from two points close to P on Pα and Pβ, p and θ (and thus the stresses) are determined at the intersection of the two slip lines emanating from these points
- the procedure is repeated step by step throughout the region PABC, bounded by two α lines and two β lines.

(iii) Homogeneous field: p and θ are constant in a given region: the lattice of α and β lines thus consists of two families of orthogonal straight lines.

(iv) Semi-homogeneous field: one of the families of slip lines consists of straight lines, on each of which therefore p is constant. If these are the α (or β) lines then the quantity $p - 2k\theta$ (or $p + 2k\theta$) is constant throughout

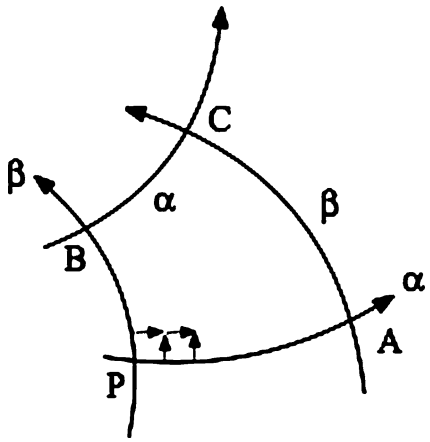


Fig. 3.139 Extrapolation of the slip lines field: p, α, β are known at P

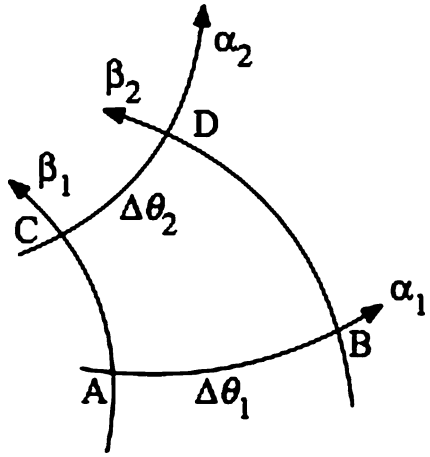


Fig. 3.140 Hencky's theorem $\Delta\theta_1 = \Delta\theta_2$

the region. This is the case, for example, for the Prandtl fan, with one family of concurrent straight lines and another of concentric circles.

(v) Hencky's theorem: for the situation of Fig. 3.140:

$$\Delta p_1 = -2k\Delta\theta_1 \text{ on } \alpha_1, \Delta p_2 = -2k\Delta\theta_2 \text{ on } \alpha_2$$

$$\Delta p_1 = p_B - p_A, \Delta p_2 = p_D - p_C$$

$$\Delta p_1 - \Delta p_2 = p_B - p_D + p_C - p_A = 2k(\theta_B - \theta_D + \theta_C - \theta_A) = 2k(\Delta\theta_1 - \Delta\theta_2)$$
$$= 2k(\Delta\theta_2 - \Delta\theta_1) \Rightarrow \Delta\theta_1 = \Delta\theta_2 \tag{3.295}$$

Thus if the part of one of the α lines lying between two β lines is straight, the same is true for all the others (semi-homogeneous field). Since two stress fields

can be continuously linked only along a characteristic, it follows in particular that a homogeneous field can only be linked to a semi-homogeneous field.

(c) Determination of velocities. The assumption of isotropy implies that $\underline{\underline{\dot{\varepsilon}}}$ and $\underline{\underline{\sigma}}$ have collinear eigen-vectors; therefore in the curvilinear co-ordinate system (α,β)

$$\dot{\varepsilon}_{\alpha\alpha} = \dot{\varepsilon}_{\beta\beta} = 0 \quad \dot{\varepsilon}_{\alpha\beta} = \dot{\lambda} > 0 \ (i.e., \text{ no extension along the "slip" lines}) \tag{3.296}$$

In terms of the velocities v_α, v_β this gives the Geringer[34] equations

$$\begin{cases} \mathrm{d}v_\alpha - v_\beta \, \mathrm{d}\theta = 0 & \text{on line } \alpha \\ \mathrm{d}v_\beta + v_\alpha \, \mathrm{d}\theta = 0 & \text{on line } \beta \end{cases} \tag{3.297}$$

$$\dot{\lambda} = (1/2)(v_{\alpha,\beta} - v_\beta \theta_{,\beta} + v_\alpha \theta_{,\alpha}) \geq 0 \quad (\text{plastic power } \geq 0)$$

We can use the *hodograph method*: for a plane problem we associate with the physical plane a velocity plane (the hodograph plane), in which the point $m(u,v)$ is the image of a physical point $M(x,y)$ at which the velocity components are (u,v) and if O is the (fixed) origin of the hodograph plane $\underline{Om} = \underline{v}(M)$. So, the (α, β) lattice has an orthogonal image (a, b) such that α in M and a in m are orthogonal, and similarly for β and b.

Since the discontinuities in velocity can be proved to be necessarily tangential to the discontinuity lines, the latter are necessarily the characteristics α and β, and hence the name *slip lines*. Further, a velocity discontinuity is constant along the line and must have the same orientation as its corresponding shear stress.

(d) For the flat punch referring to Fig. 3.141, a solution is

- in ACD: homogeneous field with $\theta = \pi/2, p = k$
- in ABC: semi-homogeneous field, Prandtl fan with opening π/2
- in A′BA: homogeneous field with $\theta = 0, p = p_{\mathrm{AA}'}$.

Applying the Hencky equation to an α line originating on AA′ we have

$$p_{\mathrm{AA}'} + 0 = k + \pi k = k(1 + \pi) \tag{3.298}$$

But $\sigma_2(\mathrm{AA}') = -p - k = -k(2 + \pi)$, and therefore

$$F \leq ka(2 + \pi) \tag{3.299}$$

[34]Hilda Geiringer (1893–1973), also known as Hilda von Mises, was an Austrian mathematics professor.

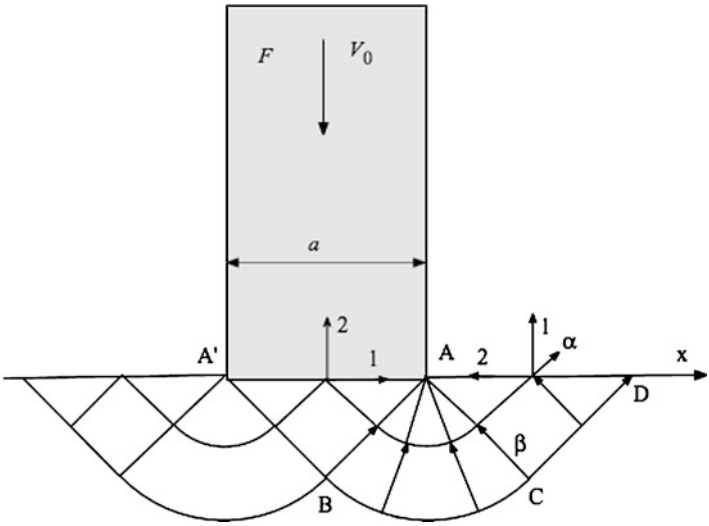


Fig. 3.141 Flat punch on a semi-infinite body: slip lines

This is an upper bound, as would be justified by the velocity analysis. The static analysis of Sect. 3.5.3.4(a) gave the same value as a lower bound, and therefore this is the *exact value*.

Note: This elementary solution can be used to assess the *hardness H* of a material, which is measured by the normal force per unit area insuring free incipient plastic flow. We have $H = F/a = k(2 + \pi) = \sigma_0(2 + \pi)/\sqrt{3}$ (by von Mises), so $H \cong 3\sigma_0$. Thus the yield stress in simple tension can be estimated by an indentation test (see Sect. 1.3.5.3 and in Volume II: Mechanics of Contacts).

3.5.3.6 Taylor's Model

The extremal theorems we have established for a standard rigid-perfectly material can be used to clarify the question raised in Sects. 3.5.1.3 and 3.5.2.1 concerning the plasticity criteria for a polycrystal whose grains (g) obey the Schmid law. We can treat such a polycrystal as a heterogeneous material and use the theorems to set upper and lower bounds for the flow limit, as we have already done for elastic behaviour of such materials. Since, as we have shown, polycrystals are a standard material, we can use the above analysis for a rigid-perfectly plastic behaviour.

A *static model*, equivalent to the Reuss model, can be constructed on the basis of a uniform stress field $\underline{\underline{\sigma}}^* =< \underline{\underline{\sigma}}^* >= \underline{\underline{\Sigma}}$. For simplicity, we shall consider only the case of simple tension and a Schmid law with the same critical resolved shear τ_c for all systems. Choosing $\underline{\underline{\sigma}}^*$ uniaxial such that $\underline{m}^g.\underline{\underline{\sigma}}^*.\underline{n}^g \leq \tau_c \ \forall g$, and therefore a uniaxial stress equal to $2\tau_c$, we get as lower bound $\bar{\Sigma}_0 \geq 2\tau_c$.

The equivalent of the Voigt model, when elasticity can be neglected, is the Taylor model (see Sect. 3.4.2.3(c)), in which the field $\underline{\underline{\dot{\varepsilon}}}'\,(=\underline{\underline{\dot{\varepsilon}}}'^{\mathrm{p}})$ is assumed to be uniform. If the problem is formulated in terms of a strain rate $\underline{\underline{\dot{E}}}(=\underline{\underline{\dot{E}}}^{\mathrm{p}})$ prescribed on the boundary we have $\underline{\underline{\dot{E}}} = \underline{\underline{\dot{\varepsilon}}}'$, which, together with the restriction of incompressibility, gives five conditions to be imposed on each grain. Thus a grain requires five independent slip systems to be activated before it can be deformed in this manner and in each grain (g) we have

$$\underline{\underline{\sigma}}' : \underline{\underline{\varepsilon}}' = \underline{\underline{\sigma}}' : \underline{\underline{E}} = \tau_{\mathrm{c}} \sum_{i=1}^{5} \dot{\gamma}_i^g \tag{3.300}$$

The upper bound theorem then gives

$$\frac{1}{V} \int_V \underline{\underline{\sigma}}' : \underline{\underline{\dot{\varepsilon}}}' \, \mathrm{d}V \geq \Sigma_0 \dot{E} \tag{3.301}$$

where $\dot{E}$ is the macroscopic axial plastic strain rate; thus

$$\tau_{\mathrm{c}} \int \left(\sum\nolimits_i \dot{\gamma}_i^g \right) \mathrm{f}(\underline{g}) \, \mathrm{d}\underline{g} \geq \Sigma_0 \dot{E} \tag{3.302}$$

where $\mathrm{f}(\underline{g})$ is the distribution function for the crystal orientation $\underline{g}$ (defined *e.g.* by 3 Euler angles); thus

$$\Sigma_0 \leq < M(\underline{g}) > \tau_{\mathrm{c}} \quad \text{with } M(\underline{g}) = \frac{\sum_i \dot{\gamma}_i^g}{\dot{E}} \tag{3.303}$$

where $M(\underline{g})$ is the *Taylor factor* for the orientation $(\underline{g})$.

For the FCC case there will be a large number of combinations of five active slip systems out of the 12 possibilities – though less than the total of $C_{12}^5 = 792$ possibilities, not all of which are independent. For the least upper bound for Σ_0 we have to choose, for each orientation (g), the combination of five systems {111} <110> that gives the smallest possible value for $\sum_i \dot{\gamma}_i^g$, that is, the one that gives the lowest plastic power; in a FCC polycrystal without texture this gives

$$\Sigma_0 \leq 3.067 \tau_{\mathrm{c}} \tag{3.304}$$

and finally

$$2 \leq \Sigma_0 / \tau_{\mathrm{c}} \leq 3.067 \tag{3.305}$$

If the polycrystal is isotropic we know in advance the form of $\underline{\underline{\dot{E}}}$ that corresponds to flow in simple tension: it is

$$\left(\underline{\underline{\dot{E}}}\right) = \dot{E} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix} \tag{3.306}$$

from which the Taylor factors can be determined directly. This is not so if the polycrystal has a crystallographic texture and we have to express it in terms of several parameters, for example

$$\left(\underline{\underline{\dot{E}}}\right) = \dot{E} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -q & 0 \\ 0 & 0 & -1+q \end{pmatrix} \tag{3.307}$$

We must then develop a double minimisation procedure to find the optimum value of q, or of the *Lankford coefficient* $R = q/(1-q)$, and for this value the combination of slip systems that gives the lowest plastic power.

Note: the extremal nature of the Taylor model as yielding an upper bound for the flow stress has been proved by Hill with help of the above derived extremal theorems (3.284) long after it was proposed as an estimate by Taylor on a physical basis (1938). On the contrary, no such extremal nature could be proved for the Sachs[35] model (Sachs 1928) which was derived from the approximate assumption of an equal maximum shear stress, the CRSS, reached simultaneously by every grain orientation; nevertheless the computed Sachs estimate of the tensile yield stress for isotropic FCC polycrystals (namely $2.24\tau_c$) lies between Reuss and Taylor bounds.

3.5.4 *Introduction to the Plasticity of Heterogeneous Materials*

The Taylor model can be considered as the first significant contribution to nonlinear continuum micromechanics. Nevertheless, almost one century later and despite important advances in this field, especially during the last decade, this topic can still be considered as an open one and is the matter of intensive research. We only aim here at presenting the main trends of the modern plasticity of heterogeneous materials and at referring to specialist literature for more details.

The solution by Eshelby of the inclusion problem (Sect. 2.7) served as a key for trespassing the known limits of the Taylor model, according to the idea that the

[35] George (Georg Oskar) Sachs (1896–1960) was a Russian-born German and American metallurgist.

plastic strain could be considered as a "stress-free strain", in the sense of Eshelby, so that elastoplasticity could be treated as elasticity with eigenstrains. Unfortunately, this idea is not fully relevant when the considered material is plastically flowing. As long as the overall plastic flow can be considered as frozen, it leads to interesting new results, which have already been partly reported in Sect. 2.5.2 (heterogeneous elasticity with initial strains). In what follows, these results are first adapted to plasticity and somewhat extended in Sect. 3.5.4.1. The limits of validity of the assimilation of plastic strain to an eigenstrain and some ways of improvement are then discussed in Sect. 3.5.4.2. A brief outline of more recent variational approaches opening the way to fully renewed treatments is finally given in Sect. 3.5.4.3.

3.5.4.1 Elastoplasticity Treated as Elasticity with Eigenstrains

(a) *Mean values.*
The results derived in Sect. 2.5.2 for linear elasticity with initial strains cannot be applied without care to the plastic flow of heterogeneous plastic materials, because the plastic flow and then the plastic strain are stress-dependent whereas, by definition, a "stress-free strain" is not: outside the elastic domain, the mechanical behaviour is obviously nonlinear elastoplastic. Nevertheless, a number of previously obtained properties, which do not explicitly refer to an overall plastic flow, can be saved: the residual stress $\underline{\underline{\sigma}}^{\mathrm{r}}$ can be defined in the same way by referring to an elastic unloading and the relation between the local and the overall plastic strain rates can be derived through the same arguments which have been used for the derivation of (2.118), which now reads

$$\dot{\underline{\underline{E}}}^{\mathrm{p}} = < \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} : \mathbf{B} > = < \mathbf{B}^{\mathrm{T}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > \tag{3.308}$$

where $\mathbf{B}$ is still associated to a purely elastic process (Mandel 1972).

(b) *Plastic dissipation.*
Similarly to the derivation of (2.120) for the elastic stored energy, which is still valid for elastoplasticity, we can also compare the local plastic dissipation $d = \underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}}$ and the macroscopic one D. We have

$$\begin{aligned} D &= < d > = < \underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > = < \left(\underline{\underline{\Sigma}} : \mathbf{B}^{\mathrm{T}} + \underline{\underline{\sigma}}^{\mathrm{r}} \right) : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > \\ &= \underline{\underline{\Sigma}} : < \mathbf{B}^{\mathrm{T}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > + < \underline{\underline{\sigma}}^{\mathrm{r}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > = \underline{\underline{\Sigma}} : \dot{\underline{\underline{E}}}^{\mathrm{p}} + < \underline{\underline{\sigma}}^{\mathrm{r}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > \end{aligned} \tag{3.309}$$

Here again, the compatibility of the residual strain rate field $\dot{\underline{\underline{\varepsilon}}}^{\mathrm{r}} = \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} + \mathbf{s} : \dot{\underline{\underline{\sigma}}}^{\mathrm{r}}$ and the equilibrium of the residual stress field $\underline{\underline{\sigma}}^{\mathrm{r}}$ lead, through Hill's lemma, to

$$\begin{aligned} < \underline{\underline{\sigma}}^{\mathrm{r}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > &= < \underline{\underline{\sigma}}^{\mathrm{r}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{r}} > - < \underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \dot{\underline{\underline{\sigma}}}^{\mathrm{r}} > \\ &= < \underline{\underline{\sigma}}^{\mathrm{r}} > : < \dot{\underline{\underline{\varepsilon}}}^{\mathrm{r}} > - < \underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \dot{\underline{\underline{\sigma}}}^{\mathrm{r}} > \\ &= - < \underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \dot{\underline{\underline{\sigma}}}^{\mathrm{r}} > \end{aligned} \tag{3.310}$$

and then to

$$D = < \underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > = \underline{\underline{\Sigma}} : \dot{\underline{\underline{E}}}^{\mathrm{p}} - < \underline{\underline{\sigma}}^{\mathrm{r}} : s : \dot{\underline{\underline{\sigma}}}^{\mathrm{r}} > \tag{3.311}$$

This result expresses the fact that, whereas the local plastic power $\underline{\underline{\sigma}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}}$ is fully dissipated, the macroscopic one $\underline{\underline{\Sigma}} : \dot{\underline{\underline{E}}}^{\mathrm{p}}$ contains one part, namely $< \underline{\underline{\sigma}}^{\mathrm{r}} : \mathbf{s} : \dot{\underline{\underline{\sigma}}}^{\mathrm{r}} >$, which remains stored and not dissipated: it corresponds to the variation of the elastic energy of the residual stresses.

Note that, using the same arguments, we can also obtain the relation

$$\dot{\underline{\underline{\Sigma}}} : \dot{\underline{\underline{E}}}^{\mathrm{p}} = < \dot{\underline{\underline{\sigma}}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} > + < \dot{\underline{\underline{\sigma}}}^{\mathrm{r}} : \mathbf{s} : \dot{\underline{\underline{\sigma}}}^{\mathrm{r}} > \tag{3.312}$$

If the constituents of an RVE of the considered heterogeneous material have a perfectly plastic (*i.e.*, no hardening) standard behaviour with a regular yield surface, the term $< \dot{\underline{\underline{\sigma}}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} >$ in (3.312) vanishes since $\dot{\underline{\underline{\sigma}}} : \dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} = 0$ everywhere (either $\dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} = 0$ or $\dot{\underline{\underline{\sigma}}}$ is normal to $\dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}}$, due to the normality rule). So, $\dot{\underline{\underline{\Sigma}}} : \dot{\underline{\underline{E}}}^{\mathrm{p}} \geq 0$, since $\mathbf{s}$ is positive definite, which means, because the overall behaviour is standard too, as proved below, that there is a macroscopic hardening originating in the interactions between the constituents. An experimental proof of this conclusion is given by the fact that the initial part of the tensile curve of a polycrystal in the plastic regime is generally concave even though each plastically active grain at this stage experiences a quasi non-hardening flow such as the stage I of FCC single crystals (see Fig. 3.63). The same phenomenon is responsible for the major part of the kinematic hardening of polycrystals.

(c) *Macroscopic elastic domain and the principle of maximum plastic work.*
Let us consider an RVE of a heterogeneous material with standard constituents. In addition, for the sake of simplicity, these constituents are elastic-perfectly plastic, so that the local yield surface $C(\underline{x})$ at any point $\underline{x}$ is fixed. For a given macroscopic stress tensor $\underline{\underline{\Sigma}}$ associated with the local fields $\underline{\underline{\sigma}}(\underline{x})$ and $\underline{\underline{\sigma}}^{\mathrm{r}}(\underline{x})$, the macroscopic elastic domain is defined as the set of macroscopic stress tensors $\underline{\underline{\Sigma}}^*$ which can be reached from $\underline{\underline{\Sigma}}$ along purely elastic paths. On such paths, where the local stresses are denoted $\underline{\underline{\sigma}}^*(\underline{x})$, the residual stresses $\underline{\underline{\sigma}}^{\mathrm{r}}(\underline{x})$ remain constant. Since the process is elastic, we have

$$\begin{cases} \underline{\underline{\sigma}}^*(\underline{x}) = \mathbf{B}(\underline{x}) : \underline{\underline{\Sigma}}^* + \underline{\underline{\sigma}}^{\mathrm{r}}(\underline{x}) \\ \underline{\underline{\sigma}}(\underline{x}) = \mathbf{B}(\underline{x}) : \underline{\underline{\Sigma}} + \underline{\underline{\sigma}}^{\mathrm{r}}(\underline{x}) \\ \underline{\underline{\sigma}}(\underline{x}) - \underline{\underline{\sigma}}^*(\underline{x}) = \mathbf{B}(\underline{x}) : (\underline{\underline{\Sigma}} - \underline{\underline{\Sigma}}^*) \end{cases} \tag{3.313}$$

The macroscopic elastic domain $C^{\text{eff}}(\{\underline{\underline{\sigma}}^{\text{r}}\})$ can be defined as the intersection of all the domains $C(\underline{x})$ after translation by the residual stresses $\underline{\underline{\sigma}}^{\text{r}}(\underline{x})$ and transformation by $\mathbf{B}^{-1}(\underline{x})$.

During a macroscopic plastic flow $\dot{\underline{\underline{E}}}^{\text{p}}$, associated with the local flow $\dot{\underline{\underline{\varepsilon}}}^{\text{p}}(\underline{x})$, we have locally, from the principle of maximum plastic work at the local level

$$\left(\underline{\underline{\sigma}}(\underline{x}) - \underline{\underline{\sigma}}^{*}(\underline{x})\right) : \dot{\underline{\underline{\varepsilon}}}^{\text{p}}(\underline{x}) \geq 0 \qquad \text{f}(\underline{\underline{\sigma}}^{*}) \leq 0 \tag{3.314}$$

where $\text{f}(\underline{\underline{\sigma}})$ is the local yield function; this reads also, according to (3.313)

$$\left(\underline{\underline{\Sigma}} - \underline{\underline{\Sigma}}^{*}\right) : \mathbf{B}^{\text{T}} : \dot{\underline{\underline{\varepsilon}}}^{\text{p}}(\underline{x}) \geq 0 \qquad \text{f}(\underline{\underline{\sigma}}^{*}) \leq 0 \tag{3.315}$$

Averaging over the RVE, and using (3.308) and the overall yield function $\text{F}(\underline{\underline{\Sigma}})$, we get

$$\left(\underline{\underline{\Sigma}} - \underline{\underline{\Sigma}}^{*}\right) : \dot{\underline{\underline{E}}}^{\text{p}} \geq 0 \qquad \text{F}(\underline{\underline{\Sigma}}^{*}) \leq 0 \tag{3.316}$$

This proves that the principle of maximum plastic work is then also valid at the macroscopic level. In particular, this result shows that a polycrystal whose grains obey the Schmid law, which implies this principle at the local level, is a standard material at the macroscopic scale (unless specific non standard deformation modes occur at the grain boundaries).

3.5.4.2 Estimates for the Effective Elastoplastic Behaviour of Polycrystals

The Taylor model has been used intensively for the analysis of various processes of metal forming (rolling, deep drawing . . .), especially for the prediction of the formation of crystallographic textures and of the associated plastic anisotropy at finite strain. Despite a number of satisfying results, it proved, as expected from its extremal nature ("upper bound"), to be generally too stiff. The extension proposed by Lin (1957) to consider an elastoplastic behaviour instead of a rigid-plastic one and to assume a uniform *total* strain did not lead to change this conclusion; actually, for isotropic elasticity and isochoric plasticity, we have

$$\begin{aligned} \underline{\underline{\varepsilon}}^{\text{e}} + \underline{\underline{\varepsilon}}^{\text{p}} = \underline{\underline{E}}^{\text{e}} + \underline{\underline{E}}^{\text{p}} \quad \text{Tr}(\underline{\underline{\varepsilon}}^{\text{e}}) = \text{Tr}(\underline{\underline{E}}^{\text{e}}) \\ \Rightarrow \quad \underline{\underline{\sigma}} = \underline{\underline{\Sigma}} + 2\mu(\underline{\underline{E}}^{\text{p}} - \underline{\underline{\varepsilon}}^{\text{p}}) \end{aligned} \tag{3.317}$$

With $\left|\underline{\underline{\sigma}} - \underline{\underline{\Sigma}}\right| / \mu$ of the order of 10^{-3} to 10^{-2}, this proves that, from a practical point of view, Lin's and Taylor's models almost coincide: actually, they were found to yield almost identical predictions.

As indicated above, new approaches followed the resolution by Eshelby (1957) of the inclusion problem: Budiansky, Hashin and Sanders (1960) proposed to model the very commencement of the plastic regime by considering the first plastically active grains as isolated (spherical) inclusions I in an elastic matrix, according to Eshelby's "dilute approximation" approach (remind Sect. 2.8.2 which reported on an application of this approach to heterogeneous elasticity). For isotropic elasticity and isochoric plasticity, this treatment leads to the relations

$$\begin{cases} \underline{\underline{\varepsilon}}_{\mathrm{I}} = \underline{\underline{E}} + \mathbf{S}_{\mathrm{I}}^{\mathrm{Esh}} : \underline{\underline{\varepsilon}}_{\mathrm{I}}^{\mathrm{p}} = \beta \underline{\underline{\varepsilon}}_{\mathrm{I}}^{\mathrm{p}} \\ \underline{\underline{\sigma}}_{\mathrm{I}} = \underline{\underline{\Sigma}} - 2\mu(1-\beta)\underline{\underline{\varepsilon}}_{\mathrm{I}}^{\mathrm{p}} \end{cases} \tag{3.318}$$

One year later, Kröner (1961) crossed the path towards the fully developed plastic flow by extending this treatment to the plastic regime: the matrix was supposed to undergo the uniform plastic strain $\underline{\underline{E}}^{\mathrm{p}}$ and the self-consistent scheme was used instead of the dilute approximation since the volume fraction of the plastically active grains is no more negligible. Under the same assumptions of isotropic elasticity and isochoric plasticity and using spherical inclusions as representative of every phase (r), defined by the same crystal orientation, this corresponds to a particular case of the problem solved by Eq. 2.112 of Chap. 2, namely

$$\underline{\underline{\varepsilon}}^{\mathrm{HI}} = (\mathbf{I} + \mathbf{P} : \delta\mathbf{c}^{\mathrm{HI}})^{-1} : \left[\underline{\underline{E}} + \mathbf{P} : (\mathbf{c}^{\mathrm{HI}} : \underline{\underline{\varepsilon}}^{\mathrm{F}} - \mathbf{C} : \underline{\underline{E}}^{\mathrm{F}})\right] \tag{2.112bis}$$

This problem referred to an ellipsoidal elastic inhomogeneity with a uniform eigenstrain $\underline{\underline{\varepsilon}}^{\mathrm{F}}$ embedded in an infinite elastic matrix with the uniform eigenstrain $\underline{\underline{E}}^{\mathrm{F}}$, uniformly deformed by $\underline{\underline{E}}$ at infinity. Changing HI into (r), $\underline{\underline{\varepsilon}}^{\mathrm{F}}$ into $\underline{\underline{\varepsilon}}^{\mathrm{p}}$, $\underline{\underline{E}}^{\mathrm{F}}$ into $\underline{\underline{E}}^{\mathrm{p}}$ and $\mathbf{P}$:$\mathbf{C}$ into $\mathbf{S}^{\mathrm{Esh}}$ and putting $\mathbf{c}^{\mathrm{HI}} = \mathbf{C}$, we find

$$\underline{\underline{\varepsilon}}_{r} = \underline{\underline{E}} + \mathbf{S}_{r}^{\mathrm{Esh}} : (\underline{\underline{\varepsilon}}_{r}^{\mathrm{p}} - \underline{\underline{E}}^{\mathrm{p}}) \tag{3.319}$$

For spherical inclusions, isotropic elasticity and isochoric plasticity, this equation reads simply

$$\underline{\underline{\varepsilon}}_{r} = \underline{\underline{E}} + \beta\,(\underline{\underline{\varepsilon}}_{r}^{\mathrm{p}} - \underline{\underline{E}}^{\mathrm{p}}) \tag{3.320}$$

or equivalently

$$\underline{\underline{\sigma}}_{r} = \underline{\underline{\Sigma}} + 2\mu(1-\beta) : (\underline{\underline{E}}^{\mathrm{p}} - \underline{\underline{\varepsilon}}_{r}^{\mathrm{p}}) \tag{3.321}$$

which is the so-called "Kröner interaction law".

It is consistent with a full assimilation of the plastic strain to an eigenstrain: whereas this treatment was a reasonable assumption for the incipient plastic regime, when most of the grains (and then the "matrix") are still elastic, it becomes debatable at larger strain when elastoplastic interactions between grains cannot reduce any longer to the elastic ones which are concerned with Eshelby's inclusion problem. As a matter of fact, predictions derived from Kröner's model proved to be, except near the initial yield stress, very close to those derived from an assumption of uniform (either plastic or total) strain: this could be guessed from a comparison between (3.317) and (3.321): with $\beta \simeq 1/2$, $\left|\underline{\underline{\sigma}} - \underline{\underline{\Sigma}}\right| / \mu$ is only divided by two when going from Lin's to Kröner's model, which does not change its order of magnitude and the associated practical conclusion of the prediction of quasi-uniform plastic strain and too stiff estimates.

Four years later, Hill (1965) proposed a completely different use of Eshelby's solution: within the same self-consistent approach as Kröner, which is adequate to the granular morphology of polycrystals, he proposed first to linearise the elastoplastic behaviour at the local and the overall scales by use of the tangent (multibranch) instantaneous moduli, denoted $\mathbf{l}_r$ and $\mathbf{L}^{\mathrm{SC}}$, respectively:

$$\left\{ \begin{array}{l} \dot{\underline{\underline{\sigma}}}_r = \mathbf{l}_r : \dot{\underline{\underline{\varepsilon}}}_r \\ \dot{\underline{\underline{\Sigma}}} = \mathbf{L}^{\mathrm{SC}} : \dot{\underline{\underline{E}}} \end{array} \right. \tag{3.322}$$

and then to estimate the strain-rate concentration tensors of this linearised version from the solution of Eshelby's *inhomogeneity* (instead of *inclusion*) problem. This can be done as explained in the case of linear elasticity in Sect. 2.8.2.2 whose resulting Eq. 2.194, namely

$$\mathbf{C}^{\mathrm{SC}} = < \mathbf{c} : (\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : \delta\mathbf{c}^{\mathrm{SC}})^{-1} > : < (\mathrm{I} + \mathbf{P}^{\mathrm{SC}} : \delta\mathbf{c}^{\mathrm{SC}})^{-1} >^{-1} \tag{2.194bis}$$

now reads

$$\mathbf{L}^{\mathrm{SC}} = < \mathbf{l} : \left[\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : (\mathbf{l} - \mathbf{L}^{\mathrm{SC}})\right]^{-1} > : < \left[\mathbf{I} + \mathbf{P}^{\mathrm{SC}} : (\mathbf{l} - \mathbf{L}^{\mathrm{SC}})\right]^{-1} >^{-1} \tag{3.323}$$

For similar and aligned ellipsoids, the concentration equation reduces to

$$\dot{\underline{\underline{\sigma}}}_r = \dot{\underline{\underline{\Sigma}}} - \mathbf{L}^* : (\dot{\underline{\underline{\varepsilon}}}_r - \dot{\underline{\underline{E}}}) \tag{3.324}$$

where Hill's constraint tensor $\mathbf{L}^*$ is linked with $\mathbf{L}^{\mathrm{SC}}$ and $\mathbf{S}^{\mathrm{EshSC}}$ in the same way as $\mathbf{C}^*$ was linked with $\mathbf{C}$ and $\mathbf{S}^{\mathrm{Esh}}$ in (2.113) for linear elasticity. This result clearly illustrates the elastoplastic nature of the intergranular interactions according to Hill's model, whereas they were purely elastic according to Kröner's one.

Consequently, Hill's predictions for the effective behaviour are expected to be softer than those resulting from Taylor's, Lin's or Kröner's approaches. As a matter of fact, a number of applications of Hill's model to metal forming, initiated by Hutchinson (1970) and extended at finite strain (Iwakuma and Nemat-Nasser 1984; Lipinski et al. 1990), have shown this model to constitute a significant improvement compared with the former ones. Another way to understand the reason for that is given by the simplified isotropic version of Hill's model proposed by Berveiller and Zaoui (1979) where the *deformation* theory of plasticity (see Sect. 3.5.3.2) is used instead of the *flow* theory. Under the assumption of local and global proportional loading, isotropic elasticity and phase distribution and isochoric plastic strain, this so-called "secant approach" (since the constitutive equations now relate finite, and not infinitesimal, strains and stresses) leads to the following concentration equation:

$$\underline{\underline{\sigma}}_r = \underline{\underline{\Sigma}} + 2\mu(1-\beta)\mathrm{f}(\bar{E}^{\mathrm{p}}) : (\underline{\underline{E}}^{\mathrm{p}} - \underline{\underline{\varepsilon}}^{\mathrm{p}}_r) \tag{3.325}$$

In this equation, the function f depending on the equivalent macroscopic plastic strain $\bar{E}^{\mathrm{p}}$, defined by (3.269), is found to be equal to one as long as $\bar{E}^{\mathrm{p}} = 0$ and then to be rapidly decreasing by one or two orders of magnitude in the plastic regime. This means that Kröner's interaction law (3.321) is recovered only in the elastic domain and that, in the plastic regime, the plastic strain deviation can increase significantly with respect to Kröner's model prediction, yielding, as wished, a far softer estimate of the effective response.

That is the reason why Hill's approach rapidly gained acceptance so as to have been considered for a long time the standard for deriving nonlinear estimates. This is no longer true because of the development of new variational treatments for nonlinear elasticity. Nevertheless, by introducing the method of linearisation of the constitutive equations, Hill's approach basically opened the way to these new treatments themselves. We give in conclusion a brief outline of these treatments (Zaoui 2002).

3.5.4.3 New Methods for Bounding and Estimating the Effective Behaviour in Nonlinear Elasticity; Consequences for Elastoplasticity

A connexion between nonlinear elasticity and elastoplasticity ("deformation theory") can be established when proportional loading paths are considered (see Sect. 3.5.3.2). In addition, Hill's incremental linearisation method can be applied to nonlinear elasticity and combined with any linear homogenisation scheme; the same statement holds for the secant linearisation method quoted above. That is why some advances achieved in the field of homogenisation for nonlinear elasticity can be used for elastoplasticity.

This happened when new variational methods for nonlinear elasticity were proposed, following the pioneer work of Talbot and Willis (1985), by several authors (Ponte Castañeda1991; Suquet 1993; Willis 1994), and provided rigorous

Hashin-Shtrikman -type upper bounds for the effective elastic potential. These new bounds could then be compared with several estimates for the overall responses predicted by several of the previous models, including Hill's one, when applied to nonlinear elasticity, as they can be: this showed in particular that Hill's model may violate these bounds, at least for some values of the material parameters. This conclusion suffices to conclude that Hill's incremental linearisation method is still too stiff and has to be replaced by softer ones. That is the reason why a number of alternative linearisation methods were proposed (the "tangent", "affine", "second-order"... formulations) and their predictions compared with the bounds, with varying success.

A detailed description of these approaches and of their results would lie out of the scope of this book. Interested readers can usefully refer to the concerned research literature, *e.g.* the general review by Ponte Castañeda and Suquet (1998). Nevertheless, one important conclusion can be stressed. As suggested above, the homogenisation theory for nonlinear elastic heterogeneous materials has strongly benefited in the last decade from the development of these new variational approaches. These advances can be directly transferred to the field of viscoplasticity without elasticity (see Chap. 4) because the constitutive equations relate the stress and the strain-rate tensors in a way which is strictly similar to the relation between the stress and the strain tensors for nonlinear elasticity; in both cases, the local and global behaviour can be defined by *one* potential only. This is no more true for elastoplasticity or elasto-visco-plasticity for which *two* potentials are needed for the description of the constitutive *hereditary* behaviour. In these cases, the variational approaches for nonlinear elasticity do not work any longer; so, bounds for the overall response cannot be derived.

Though reference to nonlinear elasticity can be somewhat useful for suggesting new routes for deriving improved estimates for the effective behaviour, there is no guarantee that they would not lead to a bound violation if new bounds happen to be discovered in the future. From a physical point of view, the main difficulty lies in the so-called "long-range memory effect". Information on this effect and on the way to deal with it in simple cases is given in Chap. 5 on Viscoelasticity.

References

Ashby MF (1970) Deformation of plastically non-homogeneous materials. Philos Mag 21:399–424

Bacon DJ, Osetsky YN, Rodney D (2009) Dislocation-obstacle interactions at the atomic level. In: Hirth JP and Kubin LP (eds) Dislocations in solids. Elsevier North-Holland Amsterdam: 1–90

Bain EC (1924) The nature of martensite. Trans Am Inst Miner Metall Eng 70:25

Berveiller M, Zaoui A (1979) An extension of the self-consistent scheme to plastically flowing polycrystals. J Mech Phys Solids 26:325–344

Bilby BA (1960) Continuous distribution of dislocations. In: Sneddon IN, Hill R (eds) Progress in solid mechanics I. North-Holland, Amsterdam, pp 329–398

Bogers AJ, Burgers WJ (1964) Partial dislocation in the 110 planes in the B.B.C. lattice and the transition of the F.C.C. into the B.C.C. lattice. Acta Metall 12:255–261

Bouaziz O (2011) Strain-hardening of twinning induced plasticity steels doi:10.1016/j.scriptamat.2011.11.029
Bouaziz O, Guelton N (2001) Modelling of TWIP effect on work-hardening. Mater Sci Eng A 319–321:246–249
Bouaziz O, Allain S, Scott CP, Cugy P, Barbier D (2011) High manganese austenitic twinning induced plasticity steels: a review of the microstructure properties relationships. Curr Opin Solid State Mater Sci 15:141–168
Brown LM (1964) Selfstress of dislocations and the shape of extended nodes. Philos Mag 10(105):441
Brown LM, Tholen AR (1964) The shape of 3-fold extended nodes. Discuss Faraday Soc 38:35
Budiansky B, Hashin Z, Sanders JL (1960) The stress field of a slipped crystal and the early plastic behavior of polycrystalline materials. In: Plasticity Proceeding of 2nd Symposium Naval Structural Mechanism, Pergamon Press, Oxford, p 239
Caillard D (2010a) Kinetics of dislocations in pure Fe. Part I. In situ straining experiments at room temperature. Acta Mater 58:3504–3515
Caillard D (2010b) Kinetics of dislocations in pure Fe. Part II. In situ straining experiments at low temperature. Acta Mater 58:3493–3503
Carlton CE, Ferreira PJ (2007) What is behind the inverse Hall-Petch effect in nanocrystalline materials? Acta Mater 55:3749–3756
Chen M, Ma E, Hemker KJ, Sheng H, Wang Y, Cheng X (2003) Deformation twinning in nanocrystalline aluminium. Science 300:1275
Christian JW, Mahajan H (1995) Deformation twinning. Prog Mater Sci 39:1–157
Coujou A (1987) Observation in situ de la propagation d'une micromacle (mécanismes d'intersection). Philos Mag Lett 56:13–21
Coujou A, Beneteau A, Clement N (1992) Observation in situ dans un MET du cisaillement d'un joint de macle cohérent Σ3 par une micro-macle de déformation. Acta Metall Mater 40:337–344
Deo C, Tomé C, Lebensohn R, Maloy S (2008) Modeling and simulation of irradiation hardening in structural ferritic steels for advanced nuclear reactors. J Nucl Mater 377:136–140
Domain C, Monnet G (2005) Simulation of screw dislocation motion in iron by molecular dynamics simulations. Phys Rev Lett 95:215506–215510
Efstathiou C, Sehitoglu H (2010) Strain hardening and heterogeneous deformation during twinning in Hadfield steel. Acta Metall Mater 58:1479–1488
El-Sherik AM, Erb U (1995) Synthesis of bulk nanocrystalline nickel by pulsed electrodeposition. J Mater Sci 30:5743–5749
Erb U (1995) Electrodeposited nanocrystals: synthesis, properties and industrial applications. Nanostruct Mater 6:533–538
Eshelby JD (1957) The determination of the elastic field of an ellipsoidal inclusion and related problems. Proc R Soc Lond A 241:376–396
Fisher FD, Reisner G, Werner E, Tanaka K, Cailletaud G, Antretter T (2000) A new view on transformation induced plasticity (TRIP). Int J Plasticity 16:723–748
Foreman AJE, Makin MJ (1966) Dislocation movement through random arrays of obstacles. Philos Mag 14:911–924
Franciosi P, Berveiller M, Zaoui A (1980) Latent hardening in copper and aluminium single crystals. Acta Metall 28:273–283
Frank FC, Read WT (1950) Multiplication processes for slow moving dislocations. Phys Rev 79:772
Frederiksen SL, Jacobsen KW (2003) Density functional theory of screw dislocation core structures in bcc metals. Philos Mag 83:365–375
Friedel J (1964) Dislocations. Pergamon Press, Oxford, p 264
Garat V, Cloué J-M, Poquillon D, Andrieu E (2008) Influence of Portevin-Le Chatelier effect on rupture mode of alloy 718 specimens. J Nucl Mater 375:95–101
Germain P, Nguyen QS, Suquet P (1983) Continuum thermodynamics. J Appl Mech 50:1010–1020
Gil Sevillano J (1993) The cold-worked state. Materials Science Forum 113–115:19–28

Greenwood GW, Johnston RH (1965) The deformation of metals under small stresses during phase transformations. Proc R Soc A 283:403–422
Hall EO (1951) The deformation and ageing of mild steel. II characteristics of the Lüders deformation. Proc R Soc Lond B 64:742–747
Head AK (1953) Edge dislocations in inhomogeneous media. Proc Phys Soc Lond B 66:793
Hill R (1950) The mathematical theory of plasticity. Clarendon Press, Oxford
Hill R (1965) Continuum micro-mechanics of elastoplastic polycrystals. J Mech Phys Solids 13:89–101
Hill R (1998) The mathematical theory of plasticity. Oxford University Press, Oxford
Hirsch PB, Howie A, Nicholson RB, Pashley DW, Whelan MJ (1965) Electron microscopy of thin crystals. Butterworth, London
Hutchinson JW (1970) Elastic-plastic behavior of polycrystalline metals and composites. Proc. R. Soc. London A319:101–127
Iwakuma T, Nemat-Nasser S (1984) Finite elastic-plastic deformation of polycrystalline metals. Proc. R. Soc. London A394:87–119
Jaoul B (2008) Étude de la plasticité et application aux métaux. Mines Paris, Paris
Jouffrey B and Dorignac D (1992) Observation of monolayer Guinier-Preston zones in Al-at%1.7Cu J Phys 2:1067–1074
Karlik M, Jouffrey B (1996) Guinier-Preston platelets interaction with dislocations. Study at the atomic level. J Phys III 6:825–829
Kocks UF, Argon AS, Ashby M, Kocks UF, Argon AS, Ashby MF (1975) Thermodynamics and kinetics of slip. Pergamon, Oxford
Koehler JS (1941) On the dislocation theory of plastic deformation. Phys Rev 60:397–410
Koiter WT (1960) General theorems for elastic-plastic solids. In: Sneddon IN, Hill R (eds) Progress in solid mechanics I. North-Holland, Amsterdam, pp 163–221
Kondo K (1952) On the geometrical and physical foundations of the theory of yielding. In: Proceedings of the 2nd Japan National Congress for Applied Mechanics, Tokyo, pp 41–47
Kröner E (1958) Kontinuumstheorie des Versetzungen und Eigenspannungen. Springer, Berlin
Kröner E (1961) Zur plastischen Verformung des Vielkristalls. Acta Metall 9:155–161
Leblond JB, Devaux J, Devaux JC (1989) Mathematical modelling of transformation plasticity of steel; I case of ideal plasticity. Int J Plasticity 5:551
Lecroisey F (1971) Transformations martensitiques induites par déformation plastique dans le système Fe-Ni-Cr-C PhD thesis Université de Nancy 15 March
Lecroisey F, Thomas B (1970) Variation of intrinsic stacking fault energy with temperature in Fe-18-Cr-12 –Ni Alloys. Phys Statu Solidi A 2:K217
Lemaitre J, Chaboche JL (1985) Mécanique des matériaux solides. Dunod, Paris
Li J, Wang C-Z, Chang JP, Cai W, Butalov VV, Ho KM, Yip S (2004) Core energy and Peierls stress of screw dislocations in bcc molybdenum. A periodic-cell tight-binding study. Phys Rev B 70:104113
Le Chatelier F and Portevin A (1923) Sur un phénomène observé lors de l'essai de traction d'alliages en cours de transformation C R Acad Sci Paris 176:506–510
Liao XZ, Zhou F, Lavernia EJ, He DW, Zhu YT (2003) Deformation twinning in nanocrystalline Al. Appl Phys Lett 95:5062
Lin TH (1957) Analysis of elastic and plastic strains of a face-centered cubic crystal. J Mech Phys Solids 5:143–149
Lipinski P, Krier J, Berveiller M (1990) Elastoplasticité des métaux en grandes déformations: comportement global et évolution de la structure interne. Rev Phys Appl 25:361–388
Madec R, Devincre B, Kubin LP, Hoc T, Rodney D (2003) The role of collinear interaction in dislocation-induced hardening. Science 301:1879–1882
Magee CL (1966) Transformation kinetics, microplasticity and aging of martensite in Fe-31Ni. PhD thesis, Carnegie Inst Tech Pittsburgh
Mandel J (1972) Plasticité classique et viscoplasticité. CISM Lecture Notes No. 97 Springer, Wien
Marchini JM, Wyon G (1962) Observation de spirales sur des surfaces d'aluminium polies électrolytiquement. Acta Metall 10:915–924

Mecking H, Kocks UF (1981) Kinetics of flow and strain-hardening. Acta Metall 29:1865–1875
Mendelev MI, Han S, Srolovitz DJ, Ackland GJ, Sun DY, Asta M (2003) Development of new interatomic potentials appropriate for crystalline and liquid iron. Philos Mag 83:3977–3994
Meyers MA, Mishra A, Benson DJ (2006) Mechanical properties of nanocrystalline materials. Prog Mater Sci 51:427–556
Mughrabi H (1988) Dislocation clustering and long-range internal stresses in monotonically and cyclically deformed metal crystals. Rev Phys App 23:367–379
Müllner P, Solenthaler C, Uggowitzer PJ, Speidel MO (1994) Brittle fracture in austenitic steel. Acta Metall Mater 42:2211–2217
Nabarro FRN (1947) Dislocations in a simple cubic lattice. Proc Phys Soc 59:256–272
Nabarro FRN (1967) Theory of dislocations. Clarendon, Oxford
Orowan E (1948) International symposium on stresses in metals and alloys session III discussion. Institute of Metals London, p 451
Peach M, Koehler JS (1950) The forces produced on dislocations and the stress fields produced by them. Phys Rev 80:436–439
Petch NJ (1953) The cleavage strength of polycrystals. J Iron steel Inst 174:25–28
Peierls RE (1940) The size of a dislocation. Proc Phys Soc London 52:34–37
Ponte Castañeda P (1991) The effective mechanical properties of nonlinear isotropic composites. J Mech Phys Solids 39:45–71
Ponte Castañeda P, Suquet P (1998) Nonlinear composites. Adv Appl Mech 34:171–302
Pope DP, Ezz SS (1984) Mechanical properties of Ni_3Al and nickel-base alloys with high volume fraction of γ'. Int Metall Rev 29:136–167
Prager W, Hodge PG (1951) Theory of perfectly plastic solids. Wiley, New York
Preston GD (1938) The diffraction of X-rays by an age-hardening alloy of aluminium and copper; the structure of an intermediate stage. Philos Mag 25:855–871
Read WT (1953) Dislocations in crystals. McGraw-Hill, New York, pp 46–54
Rehbinder PA (1947) New physico-chemical phenomena in the deformation and mechanical treatment of solids. Nature 159:866
Rémy L (1978) Kinetics of FCC deformation twinning and its relationship to stress-strain behaviour. Acta Metall 26:443–451
Rémy L (1981) The interaction between slip and twinning systems and the influence of twinning on the mechanical behaviour of FCC metals and alloys. Met Trans 12A:387–408
Reppich B (1993) in Plastic Deformation and Fracture of Materials ed Mughrabi H Volume 6 of Materials Science and Technology (VCH Publ Inc Weinheim Germany)
Robertson A, Erb U, Palumbo G (1999) Practical applications of electrodeposited nanocrystalline materials. Nanostruct Mater 12:1035–11040
Rollet AD, Kocks UF (1993) A review of the stages of work-hardening. Solid State Phenomena 35–36:1–18
Rosenfield AR, Hahn GT, Bement AL Jr, Jaffee RI (1968) Dislocation dynamics. McGraw Hill, New York
Ryen O, Laukli HI, Holdmedal B, Nes E (2006) Large strain work-hardening of aluminium alloys and the effect of Mg in solid solution. Met Mat Trans A 37:2007–2013
Saada G (2005) Hall-Petch revisited. Mater Sci Eng A 400–401:146–149
Saada G, Dirras G (2009) Mechanical properties of nanograined metallic polycrystals. In: Hirth JP and Kubin LP (eds) Dislocations in solids. Elsevier North-Holland Amsterdam: 198–248
Sachs G (1928) Zur Ableitung einer Fliessbedingung. Z Ver Deut Ing 72:734–736
Sanders PG, Youngdahl CJ, Weertman JR (1997) The strength of nanocrystalline metals with and without flaws. Mater Sci Eng A 234–236:77–82
Save M, Massonnet C, de Saxce G (1997) Plastic analysis of plates, shells and disks. North Holland/Elsevier, Amsterdam
Schmid E and Boas W (1935) Kristallplastizität Springer Berlin
Snoek J (1941) Effect of small quantities of carbon and nitrogen on the elastic and plastic properties of irons. Physica 8:711–733
Stroh AN (1962) Steady state problems in anisotropic elasticity. J Math Phys 41:77

Suquet P (1993) Overall potentials and extremal surfaces of power law or ideally plastic materials. J Mech Phys Solids 41:981–1002
Suzuki H (1952) Sci Rep Res Inst Tohoko Univ A4:455–463
Talbot DRS, Willis JR (1985) Variational principles for inhomogeneous nonlinear media. IMA J Appl Math 35:39–54
Taylor GI (1938) Plastic strain in metals. J Inst Met 62:307–324
Teodosiu C, Raphanel JL, Tabourot L (1993) Finite element simulation of the large elastoplastic deformation of multi-crystals. In: Teodosiu C, Raphanel JL, Sidoroff F (eds) Large plastic deformations. Balkema, Rotterdam, pp 153–168
Tham YW, Fu MW, Hng HH, Yong MS, Lim KB (2007) Bulk nanostructured processing of aluminum alloy. J Mater Proc Technol 192–193:575–581
Thompson N (1953) Dislocation nodes in face-centred cubic lattices. Proc Phys Soc Lond B 66:481–492
Valiev RZ (ed) (1996) Ultrafine grained materials prepared by severe plastic deformation. Annales de Chimie Science des Matériaux 21:369
Venables JA (1964) The nucleation and propagation of deformation twins. J Phys Chem Solids 25:693–700
Ventelon L, Willaime F (2007) Core structure of Peierls potential of screw dislocations in α-iron from first principles: cluster versus dipole approaches. J Comput Aided Mater Des 14(supp 1):85–94
Verdier M, Fivel M, Groma I (1998) Mesoscopic scale simulation of dislocations dynamics in fcc metals: Principles and applications. Model Simul Mater Sci Eng 6:755–7770
Videau JC, Cailletaud G, Pineau A (1996) Experimental study of the transformation induced plasticity in a Cr-Ni-Mo-Al-Ti steel. J Physique IV Colloque C1, Supplm J de Physique III 6:C1-465
Willems B, Nistor LC, Ghica C, Van Tendeloo G (2005) Strain mapping around dislocations in diamond and cBN. Phys State Solid (a) 11:2224–2228
Willis JR (1994) Upper and lower bounds for nonlinear composite behaviour. Mater Sci Eng A 75:7–14
Zaoui A (2002) Continuum Micromechanics: Survey. J Eng Mech 128:808–816

Chapter 4
Elastoviscoplasticity

Abstract This chapter is devoted to the elastoviscoplastic (or creep) behaviour of materials, in particular metals and metallic alloys. Ceramic materials and polymers are only briefly mentioned. Typical experiments used to investigate the viscoplastic behaviour of these materials are firstly described showing the diversity in the response, depending on test conditions. These experiments are mainly based on uniaxial test conditions, although multiaxial loading is also considered. Then the physical mechanisms responsible for viscoplasticity are analysed. A distinction between low temperature ($T < 0.3\ T_{\mathrm{m}}$) and high temperature ($T > 0.3\ T_{\mathrm{m}}$) viscoplasticity is made. The formalism used to describe thermal activation for low temperature viscoplasticity is described in some detail, in particular for precipitation hardening. Several high temperature creep mechanisms are described: dislocation creep and diffusion creep. Deformation mechanism maps including all these mechanisms are introduced and described for several materials including metals and ceramics. The last part is devoted to the introduction of the macroscopic models for viscoplasticity. The chapter ends with a brief presentation of the metallurgical factors, which can be used to improve the creep strength of structural metallic alloys.

4.1 Introduction

The type of behaviour to be discussed in this chapter has in common with elastoplasticity the fact that a permanent deformation remains after the load has been removed. On the other hand, viscosity now prevents the occurrence of instantaneous plastic deformations: time controls the inelastic deformation and we have what is called *rate-dependent plasticity*. Apart from this, viscoplastic behaviour is very similar to ordinary plasticity, already treated in Chap. 3 of this volume; but it is distinguished from this by the separation of the deformation into an elastic part and an inelastic viscoplastic part, say $\varepsilon = \varepsilon^{\mathrm{e}} + \varepsilon^{\mathrm{v}}$, with, in general, the absence of reversibility of deformation, even delayed, for the inelastic part.

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1_4, © Springer Science+Business Media B.V. 2012

Viscoplastic behaviour occurs especially in metals and alloys when tested at high temperatures, not less than one-third of the melting point T_m. It is then accompanied by diffusion phenomena with different characteristic distances: those related to dislocations climb (intragranular creep) are seen in the *creep* which results from the competition between work-hardening and recovery, or at greater distances those related to grain boundaries in *diffusional creep*. *Low-temperature viscoplasticity* also is possible, related to thermal activation of the plastic deformation. We define these various terms at the appropriate points later in the chapter. As well as to metals, viscoplasticity theory applies to polymers when these are loaded beyond the point at which their behaviour can continue to be viscoelastic, and also, in studies of the evolution of mechanical behaviour over very long periods, to soils, rocks and ice.

As in the study of plasticity, we make the simplifying assumptions of linear elasticity, small deformations and absence of damage. The present chapter gives:

- A description of the tests that enable the basic features of viscoplastic behaviour and the associated phenomena to be seen.
- An account of the physical models of viscoplasticity (thermal activation, creep). Particular attention is paid to the field of validity of each equation derived, so that the limits of the models can be predicted – for example, when results are to be extrapolated to longer times or higher temperatures.
- An account of phenomenological mechanical models of viscoplasticity, bringing out the physical significance of each term introduced; here we discuss the theories of plasticity and viscoplasticity in combination, and also, for reasons of computational convenience, the use of viscoplasticity rather than time-independent plasticity.
- A discussion of methods for reinforcing against creep, showing how, once the deformation mechanisms have been understood, means for preventing against excessive deformation and rupture can be found.

Only viscoplasticity is considered here, creep damage being left to Chap. 8 in Volume II.

4.2 Typical Experimental Results

The basic features of viscoplasticity appear under steady loading, with either the force or the deformation held constant after the load is applied – corresponding to creep and stress relaxation experiments, respectively. The importance of such tests is their simplicity, but in general they do not show the complete behaviour since either the stress or the plastic deformation remains practically constant each time. We must therefore consider more elaborate tests, with changes made incrementally and showing changes in the strain rate, or with the load increasing in steps. One of the most important requirements for a good characterisation of the mechanisms is the separation of the effects of deformation from those of time on the state of the material.

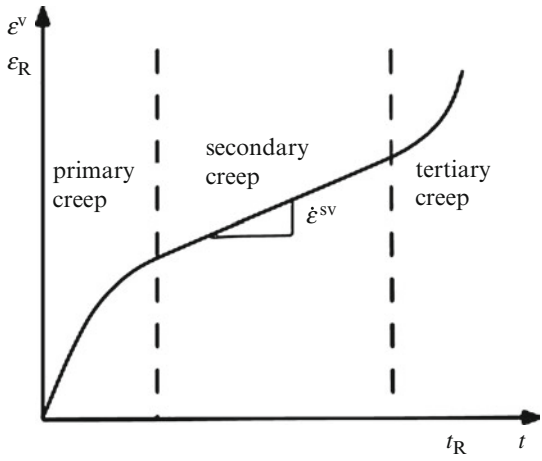


Fig. 4.1 The three stages in a creep test

4.2.1 One-Dimensional Response

4.2.1.1 Creep

Creep tests are carried out on cylindrical or flat test pieces under constant load; the time for a test can be anything from a few hours to several years, according to the application. The results are given as a time-deformation curve (Fig. 4.1). To the extent that the load is usually applied as a dead weight the stress is not constant throughout the test, but this is usually ignored and the test characterised by the initial value of the axial stress, $\sigma_0 = F/S_0$, where F is the applied force and S_0 the initial cross-section of the test piece. There are three main stages in the response to the load, more or less clearly distinguished according to the material and the temperature:

Primary stage, during which the strain rate falls; this corresponds to an increase in the resistance of the material
Secondary stage, in which the strain rate remains constant and equal to $\dot{\varepsilon}^{sv}$.
Tertiary stage, when the rate increases; significant mechanical damage appears in this stage, related to cavitation for example, or to a softening of the material, induced by localisation of the strain on the scale of the microstructure. In many cases, *tertiary creep* is associated with the formation of a neck in the tensile creep specimen similar to that formed in a tensile specimen (see Sect. 1.3.3.3). As the load is maintained constant this produces an acceleration of overall creep strain rate. Under these conditions of *structural tertiary creep*, the time to failure, t_R and the applied stress, σ are simply related by $\dot{\varepsilon}^{sv} t_R = 1/M$ where M is the Norton exponent (see Eq. 4.1c) (Hoff 1953).

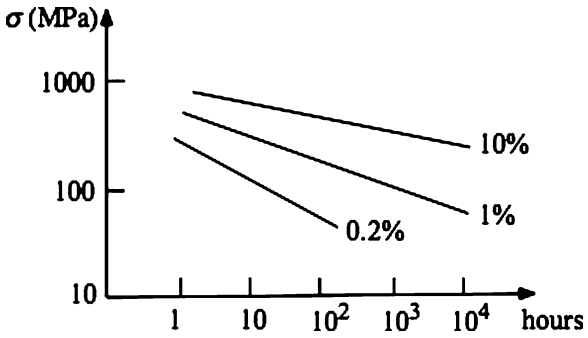


Fig. 4.2 Curves of constant deformation in creep

At low temperatures it is usually the primary and secondary creep that dominates; as the temperature rises the secondary stage becomes established sooner, and tertiary creep becomes more important.

The results of a set of tests can be presented by plotting all the curves together on the same sheet; but it can be useful to plot lines of constant deformation in the time-stress plane, showing, for each initial stress, the time required to reach a certain deformation: Fig. 4.2, for example, shows these lines for creep strains of 0.2%, 1% and 10%. This has the advantage of making it easy to compare different materials, or to assess the effect of temperature: thus material A will have a "50°C creep difference" from material B if their curves are separated by this amount.

At temperatures below 0.3 T_m (T_m being the melting point) only primary creep occurs, with the reduction in strain rate with time given by either a power or a logarithmic law, *e.g.*:

$$\dot{\varepsilon} = At^{1/3} \text{ (Andrade's law)} \tag{4.1a}$$

$$\dot{\varepsilon} = A \log (1 + t/t_0) \text{ (logarithmic creep)} \tag{4.1b}$$

where A and t_0 depend on the material.

These expressions should not be confused with the constitutive equations, since they cannot give a correct description of the way the deformation changes in response to a real variation of loading, such as partial or total unloading. We shall discuss later how work-hardening can be expressed, in terms of viscoplastic deformation for example.

Another way to present the results, restricted to secondary creep, is to show on a log-log plot of strain rate *vs.* stress the minimum rate measured in each test. In any small interval of stress the points will lie on a straight line, enabling the tests to be interpreted according to *Norton's law* (Fig. 4.3):

$$\dot{\varepsilon}^{SV} = (\sigma/K)^M \tag{4.1c}$$

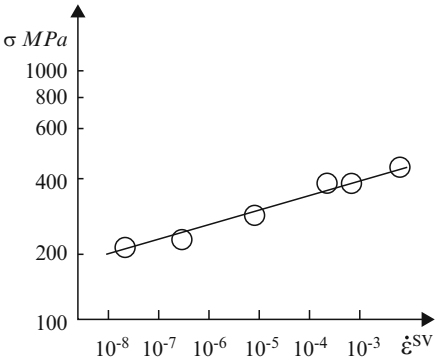


Fig. 4.3 Interpretation of creep test according to Norton's law (data points are hypothetic)

Table 4.1 Values of Norton's exponent M, creep activation energy Q and self-diffusion energy for pure metals (Mukherjee et al. 1969)

Metal	M	Q, kcal/mol	Self-diffusion energy, kcal/mol
Al	4.4	34	34
Cu	4.8	48.4	47.1
Au	5.5	48 ± 5	41.7
Ni	4.6	66.5	66.8
Pb	4.2	24.2 ± 2.5	24.2
Ta	4.2	114 ± 4	110
Cd	4.3	19 ± 2	19.1
Zn	6.1	21.6	24.3

Whilst there is only one relation that is valid for describing the steady state, and which, in particular, takes no account of work-hardening, (4.1c) is often used in calculations as though it were an established constitutive equation. The coefficient M, the numerical value of which decreases with increasing temperature, is "Norton's exponent"; Table 4.1 gives the values for a number of pure metals. Physical models, as we shall show in our later discussion (Sect. 4.3 below), predict values of 1 for diffusional creep and 4−5, according to the mechanism involved, for dislocation creep. However, these models are valid only over a limited range, and for low temperatures an internal stress has to be introduced; applying the rule as it is leads to exponents that can reach 40−50, particularly in the case of complex alloys that are not simple solid solutions.

It can be useful to keep to the same type of law, modified to include both stress and temperature; such a law, involving an activation energy and the temperature (in K) is

$$\dot{\varepsilon}^{\mathrm{SV}} = \left(\frac{\sigma}{K}\right)^{M} \exp\left(-\frac{Q}{\mathrm{k}T}\right) \tag{4.2}$$

As Table 4.1 shows, the activation energy is the same as the energy of *self-diffusion* (see this volume, Annex 2) in the case of pure metals. For more complex

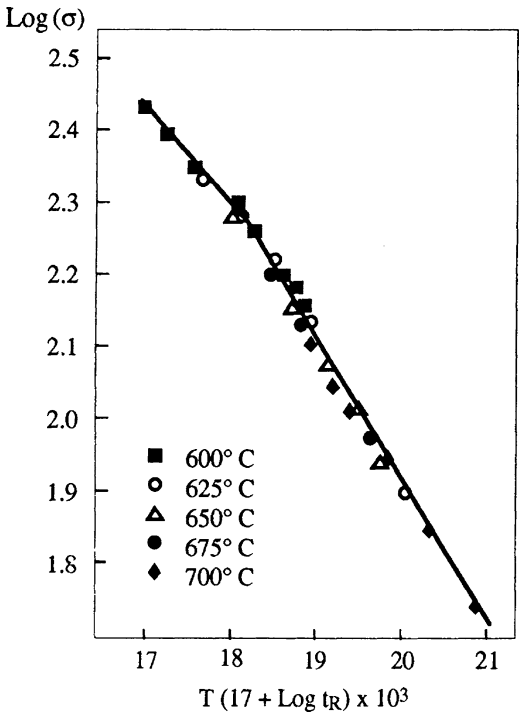


Fig. 4.4 Larson-Miller plot for austenitic stainless steel Z03CND 17–12, containing 2.5% molybdenum

materials the term cannot be given a precise physical meaning: Q can depend, for example, on the stress applied. Nevertheless, it continues to be used, in particular to establish time-temperature equivalences; these have a bearing on lifetime calculations and can be helpful in using the results of short-time tests at high temperatures to estimate times to failure at lower temperatures for which, when the relevant times are so long. If we accept the *Monkman-Grant law* (see Volume II) that states that the lifetime t_R is a power function of the rate of steady-state creep, that is $\dot{\varepsilon}^{SV} {t_R}^{\alpha} = \text{const.}$, we find $t_R = C \exp(Q/kT)$. A parametric representation is then possible provided that C and Q do not both depend on the stress: this is the case, for example, for diffusional creep, where only C varies with stress and the parameter $P' = \log t_R - \text{const}/T$ can be used to represent the creep data. If only Q varies with stress the *Larson-Miller parameter* $P = T(\log t_R + \text{const})$ can be used, as shown in Fig. 4.4 (see Larson and Miller 1952 and McClintock and Argon 1966).

4.2.1.2 Tensile Curves

In contrast to the case in which the plasticity does not depend on time, no single stress–strain curve represents tensile loading at different rates. The "rate effect" is usually to increase the stress for a given deformation, as shown in Fig. 4.5. There

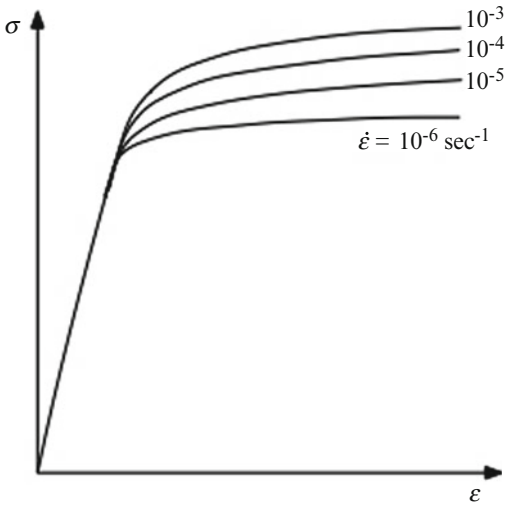


Fig. 4.5 Effect of the strain rate in a tensile test

is a tendency to saturation at very high and at very low rates, suggesting limiting cases of instantaneous plasticity at very high rates and at zero rate. If we define a *critical stress* $\sigma_c(\varepsilon^v, \alpha)$, the stress corresponding to the zero-rate case, where α is the conventional representation of the work-hardening variables, and a *viscous stress* σ_v, which depends on the viscoplastic strain rate and is zero at zero rate, we can write the stress-strain relation for different strain rates as:

$$\sigma = \sigma_c\left(\varepsilon^v, \alpha\right) + \sigma_v\left(\dot{\varepsilon}^v, \alpha\right) \tag{4.3}$$

The existence of this *viscous stress* can be evidenced by changing temporarily the loading rate in the course of a test. As Fig. 4.6 shows, there is a tendency to return to the undisturbed curve after resuming the original rate.

These observations do not always apply. Some materials show an "inverse strain rate effect" in certain regions of temperature and loading rate, in which the resistance to deformation increases with decreasing strain rate. These effects are usually associated with instabilities in the modes of deformation. These are shown at the macroscopic level by the *Portevin-Le Chatelier effect*, characterised by drops in the work-hardening curve (Fig. 4.7) when a powerful enough servomechanism is able to hold the strain rate constant (see Sect. 3.4.3.1e in this volume).

If it is the loading rate that is imposed, the curve has broad plateaux so long as nothing keeps the strain rate constant when instabilities occur, as is the case in Fig. 4.8.

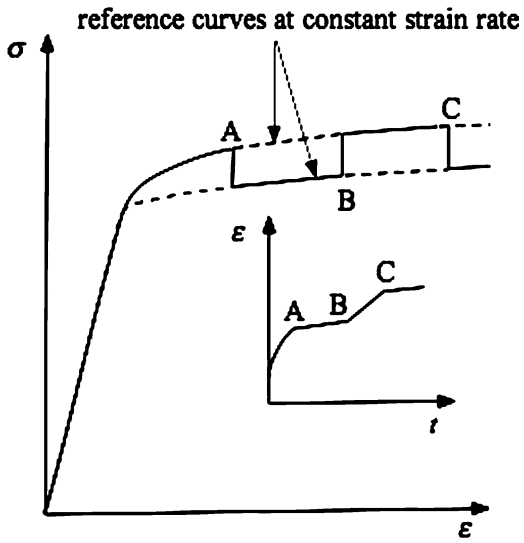


Fig. 4.6 Effect of changing the strain rate in a tensile test

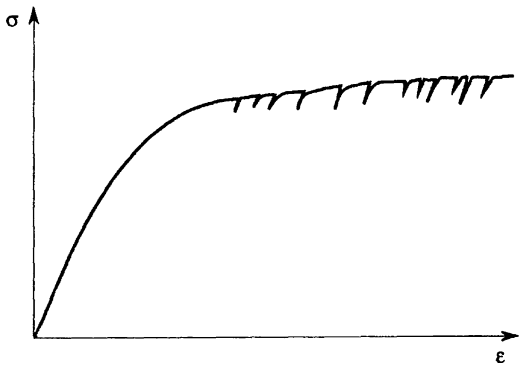


Fig. 4.7 Portevin-Le Chatelier effect in a tensile test under strain control

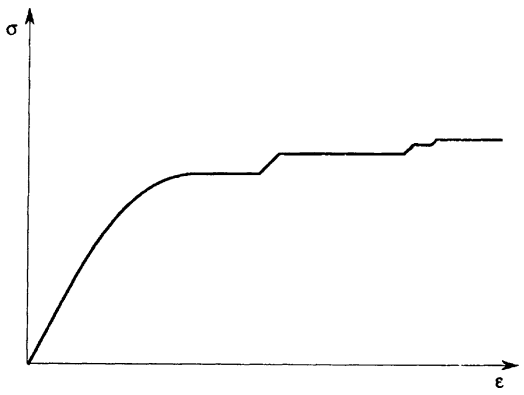


Fig. 4.8 Portevin-Le Chatelier effect in a tensile test under stress control

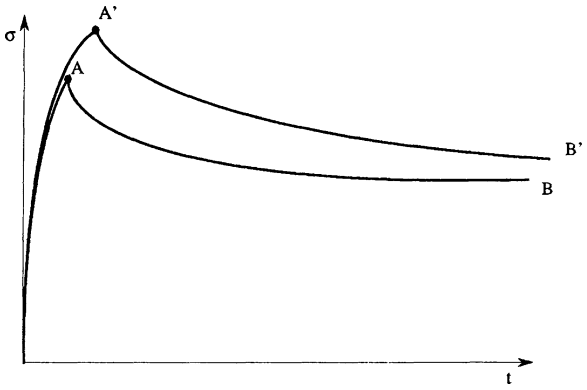


Fig. 4.9 Stress relaxation test. The total strain is maintained constant from A or A′ and then stress relaxation takes place

4.2.1.3 Stress Relaxation

In a stress relaxation test a load is applied, generally at a controlled strain rate, followed by holding the deformation constant. During the period of constant total deformation the stress falls while the viscoplastic deformation continues to increase (Fig. 4.9): this confirms that there is a fundamental difference from time-independent plasticity, since the points on the curve are *outside* the elastic domain, if indeed there is one. During the relaxation the rates of change of stress and viscoplastic deformation are related by:

$$\dot{\sigma} = -E\dot{\varepsilon}^{\mathrm{v}} \tag{4.4}$$

where E is the Young modulus in the tensile direction.

In the case where, in the absence of servocontrol of some local value, it is not the deformation but a displacement of the test piece that is held constant, the stiffness of the cross-head must be taken into account, and an equivalent, smaller modulus E^* used.

It is often more difficult here than in the case of creep test to carry out stress relaxation tests of very long duration, mainly because of the difficulty of controlling the temperature. Even so, stress relaxation tests have the advantage of enabling a wide range of viscoplastic strain rates to be covered, typically from 10^{-2} to 10^{-12} s^{-1}. The residual value reached when the fall in the stress has practically stopped (rates of less than 10^{-10} s^{-1}) corresponds to the upper limit σ_{c} of the elastic domain, introduced in Sect. 4.2.1.2. Classically, for a given material, this limit is smaller the higher the temperature. But its value remains often sufficiently high for being unreasonable to work without an elastic domain. The different forms obtained for the stress relaxation curves can be derived from identified creep laws. In most cases, simple laws that do not involve internal stresses are not able to describe both creep and stress relaxation simultaneously.

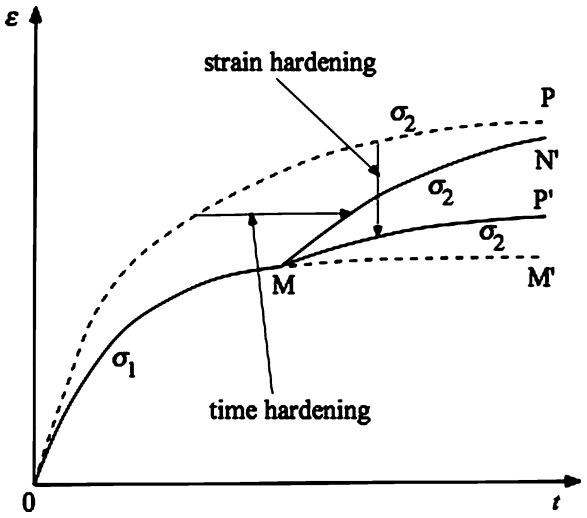


Fig. 4.10 Comparison of the time-hardening and strain-hardening assumptions. Curve OMM' is for stress σ_1; curve OP for stress σ_2; point M corresponds to the shift from stress σ_1 to stress σ_2. Strain-hardening translates OP parallel to the strain axis; time-hardening to the time axis

4.2.1.4 Stress Changes During Creep Tests

The application of sudden changes to the applied stress during primary creep enables the work-hardening in this stage to be investigated. The results of such tests can be studied from a plot on a viscoplastic strain-time diagram. The two classical assumptions are that the best indicator of the state of work-hardening is provided, (i) by the time and (ii) by the deformation. This leads to deriving curves for different values of stress by means of translations parallel either to the time axis (*time-hardening*) or to the strain axis (*strain-hardening*): see Fig. 4.10. The second assumption is the one most often used, the first not having any firm foundation (a problem that arises is what to do with the origin if the test is interrupted?). It is expressed in terms of Norton's law, modified to allow the constant K to vary with the viscoplastic stress:

$$\dot{\varepsilon}^{\mathrm{v}} = \left[\frac{\sigma}{K\left(\varepsilon^{\mathrm{v}}\right)}\right]^{M} \tag{4.5}$$

Usually $K(\varepsilon^{\mathrm{v}})$ is taken as a power function.

In the *dip test*, used to locate the elastic domain, the loading path is more elaborate. It starts with the material in the secondary creep phase, under the stress σ_0, and it is the variables corresponding to this load state that are to be investigated. As Fig. 4.11 shows, a sequence of unloadings of greater or lesser amplitude is made, returning after each to the original stress σ_0 so as not to depart too far from the steady state being studied.

The following effects are normally seen:

- After a small unloading (A in the diagram) viscoplastic flow occurs, possibly after a latency interval: this is *creep hesitation*.

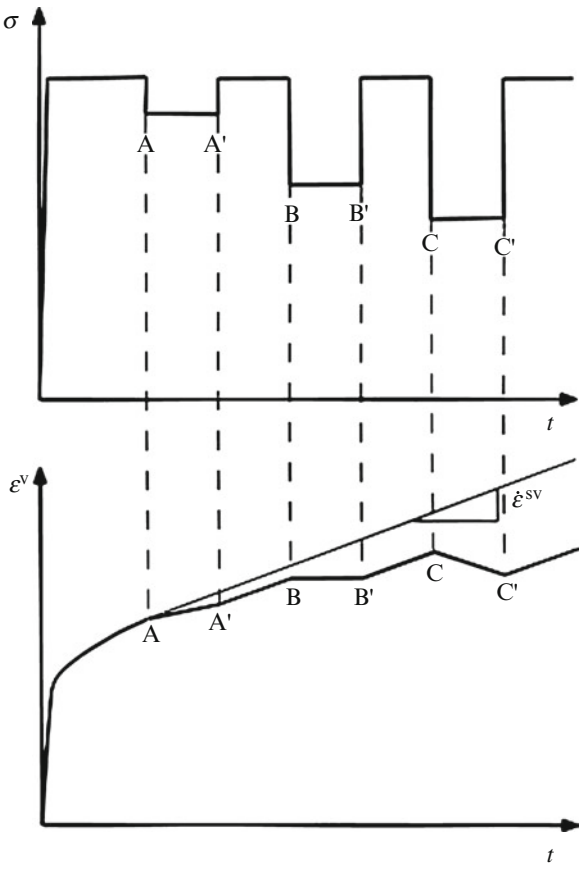


Fig. 4.11 The various stages in a dip test

- After a moderate unloading (B) there is no further flow, and the stress state is inside the elastic domain.
- After a large unloading compressive (C) viscoplastic flow may occur, the applied stress remaining positive (*i.e.* tensile). This shows that the elastic region has been crossed and the effective stress that determines the strain rate has changed sign.

Tests of this type give an indication that only laws that involve a threshold term can give a correct representation of the behaviour of a material. Norton's law can be modified with this aim in view by replacing σ by $\sigma - \sigma_s$:

$$\dot{\varepsilon}^{SV} = \left(\frac{\sigma - \sigma_s}{K}\right)^M \tag{4.6}$$

Thus kinematic work-hardening occurs in viscoplasticity as in time-independent plasticity. The formulations that lead to (4.5) and (4.6) respectively are in fact complementary rather than alternative, since both may be required to model general viscoplastic behaviour.

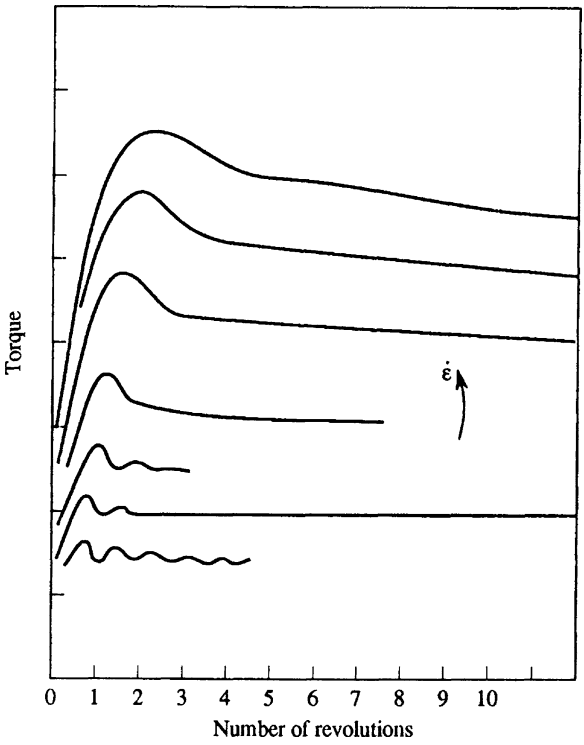


Fig. 4.12 Dynamic recovery during torsion tests

4.2.1.5 Recovery

At high temperatures viscoplasticity is accompanied by other temperature-dependent phenomena that evolve with time; recovery is one of these.

Static recovery is the name given to the phenomena associated with thermally activated microstructural rearrangements, dislocations in particular, which occur during annealing (see Sect. 3.4.2.2g). Recovery usually results in the partial or complete loss of the work-hardening built up in the course of the deformation; to account for this a fading memory term must be included in the work-hardening laws, as a function of time. The effect is seen at very low strain rates, even zero.

Dynamic recovery, on the other hand, is directly related to the deformation process and appears in the course of this when the rate is high, say from 1 to 100 s^{-1}. In particular, it occurs during torsion testing of solid cylindrical bars in which the strains can be 100% or more. These tests are largely used to simulate metal forming operations at elevated temperature. Results of such tests are given in Fig. 4.12 for different strain rates, increasing from bottom to top (Sellars and Teggart 1966). The peak of the curve corresponds, in the case of materials with low stacking-fault energy such as stainless steels, to a recrystallisation, and in those for which this is high, such as aluminium alloys, to a polygonisation in the grains

(see Sect. 3.4.2.2i); the oscillations after the peak are due to alternations between work-hardening and recrystallisation or polygonisation.

Detailed interpretation of such tests is difficult because new dislocation-free material is formed continuously, so that the state of the work-hardening becomes non-uniform through the test piece. Nevertheless, the test enables different materials to be classified relative to one another.

4.2.1.6 Ageing

There can be other phenomena at the microstructural level, for example increases in the volume fraction of precipitates already formed, or new precipitation, of carbides for example. In contrast to recovery, these can lead to hardening, which becomes evident for example in a tensile test in which the test piece is held for a time at high temperature and zero load (Fig. 4.13). Reapplying the load can display a temperature past history effect characterised by a rise above the initial stress–strain curve. There can also be softening effects, which demonstrates that in certain cases tertiary creep can be due to modifications of the deformation behaviour and not to damage.

4.2.1.7 Cyclic Tests

As well as materials that obey the laws of instantaneous plasticity, viscoplastic materials can be characterised with respect to their behaviour under cyclic loading (see Sects. 1.4 and 3.5. in this volume). The cyclic work-hardening curve can be defined as before, but now depends on the loading rate. The *Bauschinger effect* is seen again in the stress–strain loops, which can be more rounded than those for an elastoplastic material to the extent that positive viscoplastic flow continues as the tensile load is reduced (Fig. 4.14). Among tests currently in use, in particular for calibrating the damage laws (see Volume II) are those of cyclic relaxation (b) and cyclic creep (c).

4.2.2 Multiaxial Loading

To the extent that uniaxial tests raise doubts concerning the concept of an elastic domain it is reasonable to ask what might be retained from the formalism of time-independent plasticity in viscoplastic behaviour. Some answer can be derived from results of experiments with multiaxial loads. In plasticity experiments the delineation of the elastic domain, either initially or after the first loading, can be made by recording the start of plastic flow in certain directions in a loading plane, for example σ_{11}, σ_{12} or σ_{11}, σ_{22}. In the corresponding tests on a viscoplastic material

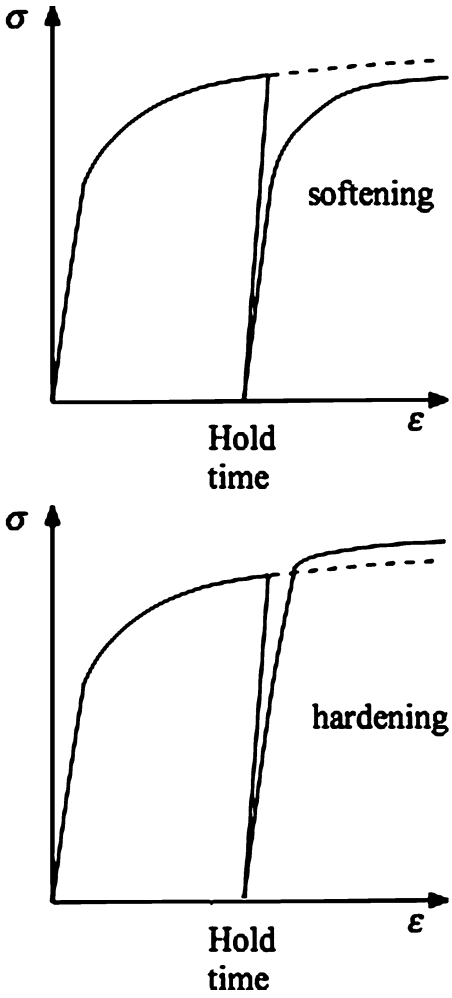


Fig. 4.13 Ageing effect in a tensile test

the strain rates are measured, and the test is in fact one of partial unloadings in multi-axial creep. Figure 4.15 is for a biaxial creep test in the σ_{11}, σ_{12} plane in which, after arriving at a stable secondary creep regime, the loads are reduced by varying amounts over short intervals (Fig. 4.15) and the magnitude and direction of the viscoplastic strain rate are measured. The results plotted in Fig. 4.15 show that the vector representing the direction of the flow is in general normal to the equipotentials in the stress plane, located with respect to a domain displaced from the origin. The magnitude of the flow vector is greater, the further the working point is from the centre of this region.

These experimental data fit very well into a scheme that uses at the same time the von Mises criterion, kinematic work-hardening and the normality rule.

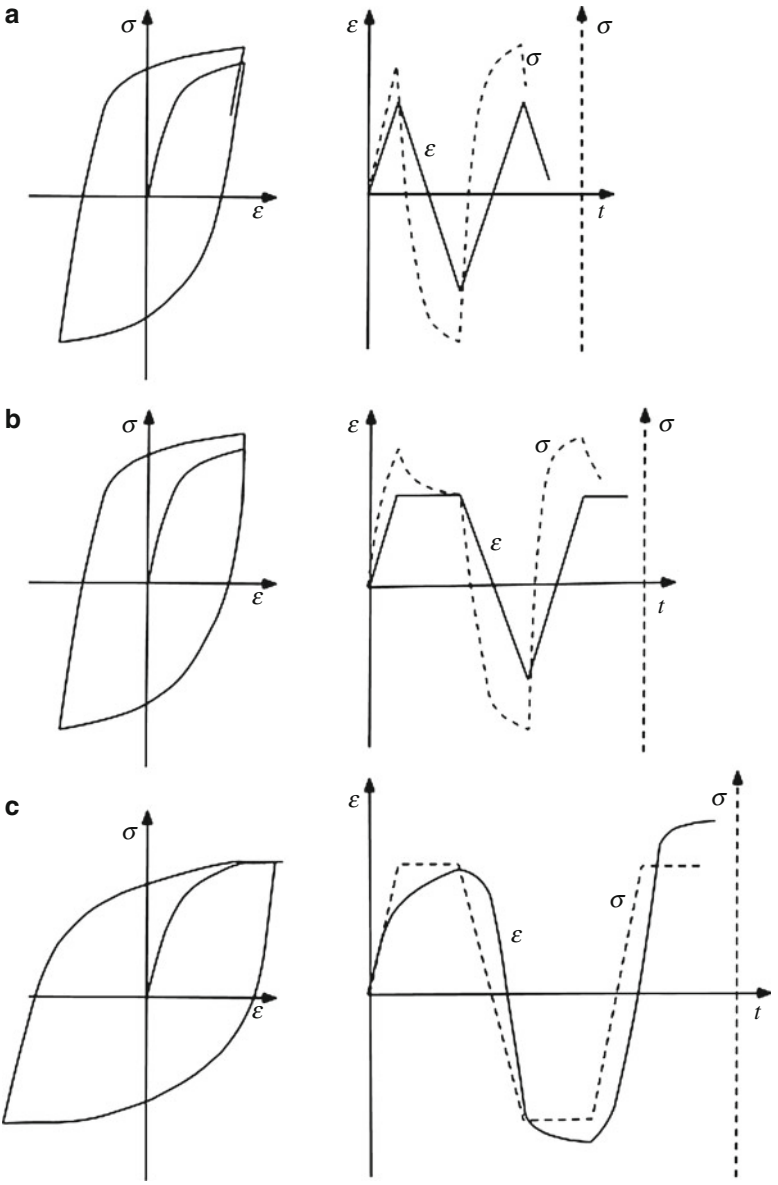


Fig. 4.14 Viscoplastic stress–strain cycles; (**a**) low-cycle fatigue test; (**b**) relaxation test with intervals at constant extension; (**c**) creep test with intervals at constant tension and constant compression

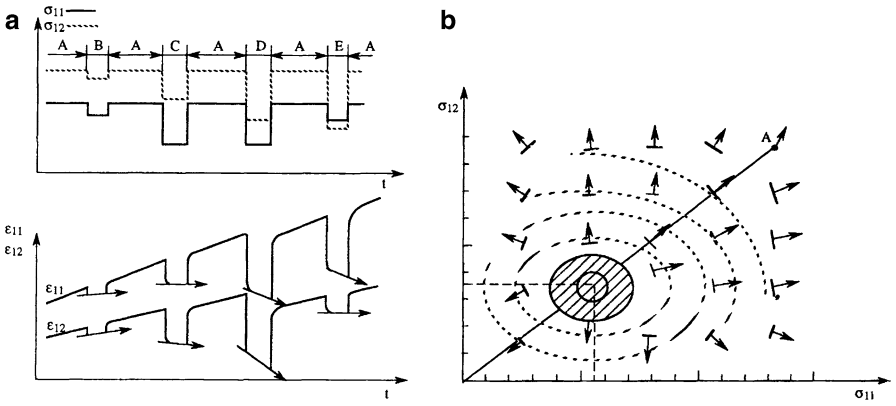


Fig. 4.15 Tension-torsion test with partial unloadings; (**a**) principle of the test; (**b**) equipotentials and flow directions in the σ_{11}, σ_{12} plane

4.2.3 Summary

The experiments we have been describing show that for a viscoplastic material the stress has to be treated as the sum of a critical stress for plastic flow and a "viscous" stress. The first, which can vanish in certain conditions, in particular at high temperature, depends primarily on the deformation (more generally, on the work-hardening) and possibly also on the time over which the time-dependent recovery mechanisms are active. The second depends primarily on the strain rate and possibly on the deformation itself (in formulations of strain-hardening type) (Sect. 4.2.1.4). Thus there are two ways of describing the hardening of a viscoplastic material: as modifying the elastic domain (*additive* hardening, since the changes raise the stress threshold); or as reducing the rate of flow outside the elastic domain (*multiplicative* hardening, since the work-hardening terms multiply the factor that normalises the rate of viscoplastic flow). In models, the work-hardening that affects the critical stress for plastic flow will follow a plastic-type law and can be isotropic or kinematic, linear or nonlinear. That which affects the *viscous* stress will, by its very nature, be isotropic. As in plasticity, kinematic work-hardening has to be used in order to model cyclic loading correctly.

The problem of activating viscoplastic flow is simpler than the corresponding problem in plasticity, since now there are only two regimes to be considered: elastic, when the working point is inside the elastic domain and viscoplastic otherwise. Further, for any given point outside the elastic domain the viscoplastic strain rate is completely independent of the rate of change of stress and depends only on the instantaneous values of the stress tensor components and of the work-hardening variables. There is no longer a consistency condition. Thus there is one equation lacking for the evaluation of the strain rate, which gives more freedom in defining this.

The choice between the descriptions we have given will be guided by a consideration of the physical mechanisms we describe next, in particular those concerning the threshold and the law of viscosity.

4.3 The Physical Mechanisms Responsible for Viscoplasticity

In this section we look at the physical mechanisms that produce viscoplastic behaviour in crystalline materials. At low temperatures, $T/T_m < 0.3$, only thermal activation of plasticity is responsible for the rate effects observed in a number of metallic alloys. The dislocations meet obstacles in the course of their movements, which they can overcome thanks not only to the forces acting on them but also to thermal activation; in this temperature region diffusion is too slow to operate. At higher temperatures, $T/T_m > 0.5$, it is diffusion that is responsible for viscoplasticity, whether over short or long distances (*dislocation or diffusional creep* respectively). We consider the two temperature regions in succession.

4.3.1 Low-Temperature Activation of Plastic and Viscoplastic Deformation

4.3.1.1 Introduction

When carrying out tensile tests on certain metals to measure the yield strength R_p at very low temperatures ($T/T_m < 0.2$) we find that it is strongly dependent on the temperature and on the strain rate; this is shown in Fig. 4.16. More generally, the flow stress, measured for example in a tensile test, is a function of strain rate, as indicated in Sect. 4.2.3 above (see Fig. 4.5). The best example of this is provided by metals and metallic alloys with BCC structures; in this case the cause of such a behaviour is the lattice friction forces or Peierls-Nabarro forces (see Sect. 3.4.2.1) under the influence of the thermal activation. We shall show later that other obstacles may be by-passed by the dislocations under the combined effect of this and the applied stress, but first we shall give the general formalism for the thermal activation of plastic deformation.

4.3.1.2 Thermal Activation of Plastic Deformation

(a) Basic mechanisms. Let ρ_M be the density of mobile dislocations with Burgers vector b and v their mean velocity; from Sect. 3.3.3.3 the corresponding strain rate is:

$$\dot{\gamma}^v = \rho_M b v \tag{4.7}$$

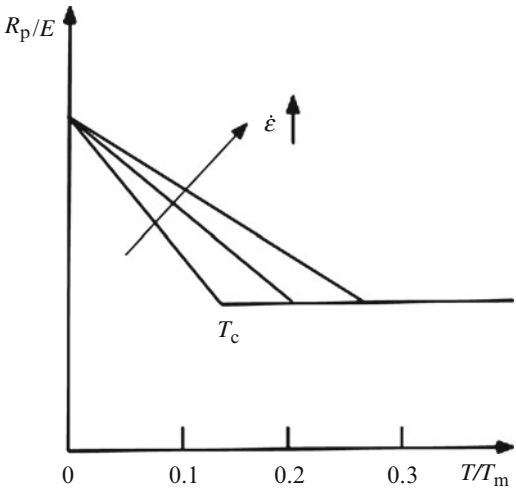


Fig. 4.16 Diagram showing that at low temperature ($T/T_m < 0.2$) the yield strength varies with both temperature and strain rate. Above a critical temperature T_c is the "athermal" region

This important relation holds whether the displacement of the dislocations is controlled by slip or by climb. An equivalent form which is better suited to the activation formalism is:

$$\dot{\gamma}^v = N_m \frac{\Delta A}{V} b\nu \tag{4.8}$$

where N_m is the number of activation sites, that is, those regions in the crystal where a segment of a dislocation is held by an obstacle, which the combination of the applied stress and the thermal agitation can help to overcome; $b\Delta A/V$ is the elementary strain produced when the dislocation crosses a potential barrier that is opposing its free movement (ΔA is the area swept out and V is the volume of the crystal); and ν is the frequency of activation, that is, the number of times per second that a site is activated and the dislocation succeeds in overcoming the obstacle.

Two cases arise, according to the distance λ travelled by the dislocation in crossing the potential barrier (Fig. 4.17). If this is a small multiple of the length b then ΔA is of the order of λl, where l is the length of the liberated segment, and it is possible for the segment to jump back to its original location. But if it is large, corresponding to the distance between precipitates, $10^2-10^3 b$, return is not possible.

Over the years 1955–1960 the results of the theory of chemical reaction were applied to the basic physical phenomena underlying the deformation of crystalline solids; this required a number of simplifying assumptions, which we now describe.

(b) General form of expressions relating dislocation rate to stress and temperature
Taking first the case in which the return jump is possible, the effective frequency will be the algebraic sum of the forward (ν^+) and backward (ν^-) frequencies:

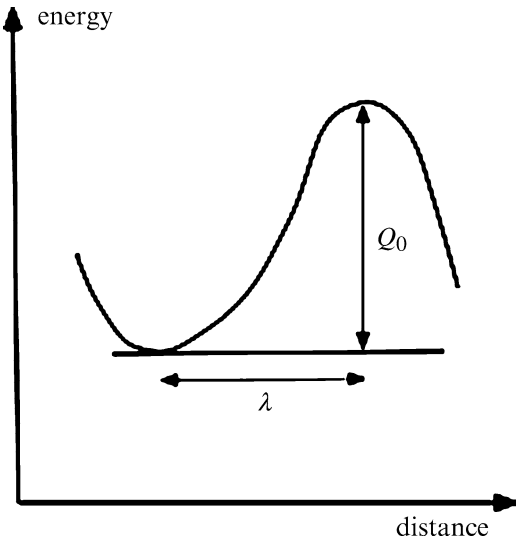


Fig. 4.17 Energy – distance diagram, activation energy

$$\bar{\nu} = \nu^+ - \nu^- = \nu_0 \exp\left(-\frac{Q_0 - \tau bA}{\mathrm{k}T}\right) - \nu_0 \exp\left(-\frac{Q_0 + \tau bA}{\mathrm{k}T}\right)$$
$$= 2\nu_0 \exp\left(-\frac{Q_0}{\mathrm{k}T}\right) \sinh \frac{\tau bA}{\mathrm{k}T} \tag{4.9}$$

Here Q_0 is, to a first approximation, the height of the potential barrier: we shall give the precise significance of this quantity later, and of A, the area swept by the dislocation. ν_0 is the jump frequency. If the probability of return is very low we have:

$$\bar{\nu} = \nu_0 \exp\left(-\frac{Q_0}{\mathrm{k}T}\right) \exp \frac{\tau bA}{\mathrm{k}T} \tag{4.10}$$

Substitution of these results in (4.8) enables us, under certain conditions, to give the general form for the relation $\dot{\varepsilon}^{\mathrm{v}} = f(T, \sigma)$.

(c) *Internal stress and effective stress*

As we indicated above when presenting some typical experimental results, it is often found necessary to introduce the concept of a threshold; metal physicists usually express this as an internal stress, which we wrote as σ_{c}, so that the "*effective*" (σ_{eff}) or "*viscous*" (σ_{v}) *stress* is written[1]:

$$\sigma_{\mathrm{v}} = \sigma_{\mathrm{eff}} = \sigma - \sigma_{\mathrm{c}} \tag{4.11}$$

[1]The counterpart of the stress σ at the grain level is the resolved shear stress denoted τ. Shifting from the micro level, when studying elementary mechanisms, to the macro level of experimental results is however somewhat problematic as will be explained later.

Table 4.2 Classes of various obstacles according to their range

Long range	Short range	Activation area (short range obstacles)
Frank network	Peierls troughs	$10\ b^2$
Attractive junctions	Repulsive junctions	$100\ b^2$
Precipitates	Solute atoms	$100\ b^2$

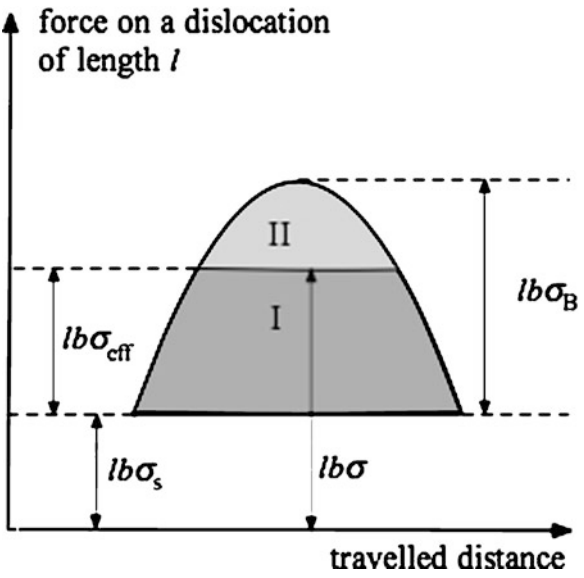


Fig. 4.18 Force – distance diagram, illustrating the concept of internal stress and viscous or effective stress. Shaded area I represents the energy provided by the effective stress, area II that provided by thermal activation

We shall return to this relation when we come to consider very general models for viscoplastic behaviour; for the moment we need only say that σ_c is an internal stress, at each point in the crystal the sum of the friction stresses and of all the long-range (that is, varying like $1/r$) stresses associated with crystalline defects in the solid – that is, mainly the effects of the other dislocations. Thus the internal stress depends only on the temperature, through the intermediary of the elastic moduli, and on the sub-structure and thus on the level of the deformation. It oscillates in space, with a wavelength of the order of the dimensions of the sub-structure (cells or subgrains) – much too long for the barriers formed by the maxima of the stress field to be surmountable by means of the thermal fluctuations. We denote by σ_s the (non-zero) mean value of this stress. Typical values of the activation area (see below) corresponding to three long range obstacles are given in Table 4.2.

The force-distance diagram, as in Fig. 4.18, can be helpful in representing the overcoming of an obstacle by thermal activation; from this:

- if $\sigma < \sigma_s$ the dislocation cannot move at all between the obstacles
- if $\sigma > \sigma_s$ the dislocation can experience a certain displacement and take up a position of stable equilibrium in contact with an obstacle, which will exert on it

a short-range returning force. The stress that contributes to the by-passing of the obstacle is: $\sigma_{\text{eff}} = \sigma - \sigma_s$; this provides a part of the energy required (area I in Fig. 4.18), the remainder (area II) coming from the thermal activation.
- if the temperature is zero a stress $\sigma = \sigma_s + \sigma_B$ is required for crossing the barrier.
- if the temperature is above a critical value T_c, so that thermal activation alone can provide all the energy represented by the area under the curve, we are in the *athermal plateau* region of Fig. 4.16, with $\sigma = \sigma_s$.
- there is an intermediate temperature T at which the combination of effective stress and thermal agitation is just sufficient for the obstacles to be surmounted at the frequency which allows the strain rate to be equal to the rate prescribed.

(d) *Thermal activation magnitudes*

Two things are necessary to enable us to take advantage of the relations derived from thermodynamics and to compare the thermal activation values measured in experiments with those predicted by the formal theory. First, the theory must be based firmly on the laws of thermodynamics; and second, the treatment must go from the elementary activation event of a dislocation overcoming an obstacle to the behaviour on the macroscopic scale. We shall take up these two points in order, but in this section we shall give most attention to the first.

To a first approximation the change in Gibbs free energy between the initial position of the dislocation (state 1), corresponding to a stable equilibrium under the action of the stress $\sigma_{\text{eff}} = \sigma - \sigma_s$, and its final position (state 2), corresponding to an unstable equilibrium when it has crossed the barrier locally, can be written:

$$\Delta G_A = \int_1^2 (\tau_B - \tau_{\text{eff}})\, bl\, dx \tag{4.12a}$$

Thus, at least in principle, we can compute ΔG_A if we know the function $\sigma_B = f(x)$; but we must know also the way the length $l(x)$ changes during the activation.

Since $\sigma_{\text{eff}} = \sigma - \sigma_s$ and σ is constant, (4.12a) can be written:

$$\Delta G_A = \int_1^2 [\tau_B(x,T) + \tau_S(x,T)]\, bl(x)dx - \tau b \int_1^2 l(x)dx \tag{4.12b}$$

in which the second term gives the work done by the applied stress. If we take σ and T as independent variables and σ_s as a constant we can write

$$d(\Delta G) = \left(\frac{\partial \Delta G}{\partial \sigma}\right)_T d\sigma + \left(\frac{\partial \Delta G}{\partial T}\right)_\sigma dT \tag{4.13}$$

from which we can define the activation entropy by:

$$\Delta S_A = -\left(\frac{\partial \Delta G_A}{\partial T}\right)_\sigma \tag{4.14}$$

Now $\partial \Delta G/\partial \sigma$ has the dimensions of volume and we should prefer to define an activation *area*, to which it is often possible to give a precise physical meaning; so we define

$$A^* = -\frac{1}{b}\left(\frac{\partial \Delta G_A}{\partial \tau}\right)_T \tag{4.15}$$

The activation enthalpy is:

$$\Delta H_{\mathrm{A}} = \Delta G_{\mathrm{A}} + T\Delta S_{\mathrm{A}} \tag{4.16}$$

which after substituting the value for ΔS_{A} gives:

$$\Delta H_{\mathrm{A}} = \partial\left(\frac{\Delta G_A}{T}\right)_\sigma \Big/ \partial\left(\frac{1}{T}\right) \tag{4.17}$$

We should note here that, if in this elementary process, the length l of the segment of the dislocation is independent of the distance x travelled, then the activation area $A^* = l.\Delta x$ has a very precise physical meaning: it is the area swept by the dislocation in moving from its initial position to that corresponding to the maximum of ΔG_{A}. In general, however, $\Delta x = \mathrm{f}(\tau)$; but in spite of this restriction it is convenient to write (4.12) in the simpler form:

$$\Delta G_A = \int_1^2 (\tau_{\mathrm{B}} + \tau_{\mathrm{S}}) bl\,\mathrm{d}x - \tau bl\Delta x \tag{4.18}$$

In fact, using relations given by the theory of reaction rates, we get, for the case in which there is no possibility of a return jump:

$$\bar{\nu} = \nu_0 \exp\left[-\frac{\int_1^2 (\tau_{\mathrm{B}} + \tau_{\mathrm{S}})\, bl\,\mathrm{d}x}{\mathrm{k}T}\right] \exp\frac{\tau bA^*}{\mathrm{k}T} \tag{4.19}$$

The integral in the numerator is the change in Gibbs free energy: $\Delta G = \Delta H_0 - T\Delta S$; so:

$$\bar{\nu} = \nu_0 \exp\frac{\Delta S}{\mathrm{k}} \exp\left[-\frac{\Delta H_0}{\mathrm{k}T}\right] \exp\frac{\tau bA^*}{\mathrm{k}T} \tag{4.20}$$

and using (4.8) we have for the macroscopic strain rate

$$\dot{\gamma}^{\mathrm{v}} = N_{\mathrm{m}} \frac{\Delta A}{V} b\nu_0 \exp\frac{\Delta S}{\mathrm{k}} \exp\left[-\frac{\Delta H_0}{\mathrm{k}T}\right] \exp\frac{\tau bA^*}{\mathrm{k}T} \tag{4.21}$$

If there is possibility of a back jump this becomes:

$$\dot{\gamma}^{\mathrm{v}} = 2N_m \frac{\Delta A}{V} b \nu_0 \exp \frac{\Delta S}{\mathrm{k}} \exp \left[-\frac{\Delta H_0}{\mathrm{k}T} \right] \sinh \frac{\tau b A^*}{\mathrm{k}T} \tag{4.22}$$

Three important remarks must be made at this stage.

1. The empirical relation (4.2), used to describe stationary creep, can often be written:

$$\dot{\varepsilon}^{\mathrm{v}} = \dot{\varepsilon}_0 \exp \left(-\frac{Q}{\mathrm{k}T} \right) \tag{4.23a}$$

which enables an apparent activation energy Q to be defined as:

$$Q = -\mathrm{k} \frac{\partial \log \left(\dot{\varepsilon}^{\mathrm{v}} / \dot{\varepsilon}_0 \right)}{\partial \left(1/T \right)} \tag{4.23b}$$

If now in (4.21) and (4.22) we put:

$$\dot{\varepsilon}_0 = N_m \frac{\Delta A}{V} b \nu_0 \exp \frac{\Delta S}{\mathrm{k}} \tag{4.24}$$

and assume that the terms in $\dot{\varepsilon}^{\mathrm{v}}$ are independent of T, we have $Q = \Delta H_0 - \sigma\, bA^*$. This shows that in the empirical relation (4.23a) the apparent activation energy is indeed a function of the applied stress.

2. Norton's law (4.1c) involves an exponent M; in rheological studies the parameter:

$$m = 1/M = \left(\frac{\partial \log \sigma}{\partial \log \dot{\varepsilon}^{\mathrm{v}}} \right)_T \tag{4.25}$$

is often used in investigations of rate sensitivity. If $\dot{\varepsilon}_0$ is independent of σ then m is directly related to the activation area and:

$$\begin{aligned} m &= b\sigma A^* / \mathrm{k}T \\ A^* &= \frac{\mathrm{k}T}{b} \left[\frac{\partial \log \left(\dot{\varepsilon}^{\mathrm{v}} / \dot{\varepsilon}_0 \right)}{\partial \sigma} \right]_T \end{aligned} \tag{4.26}$$

3. To give a physical meaning to the quantities m and Q just introduced we return to the simplifying assumptions that l is constant and $\dot{\varepsilon}^{\mathrm{v}}$ is independent of T and σ, the first of which may in some respects seem inconsistent with the fact that in general A^* is a function of τ. However, in the process of deriving the law of macroscopic deformation from the elementary mechanisms, the real problems have not been delt with; in particular, much more attention needs to be given to:

   – finding a more faithful representation of the scale change from the macroscopic stress state to the actual stresses acting on the dislocation
   – relating the deformation rate in a slip plane to the mean over the many events

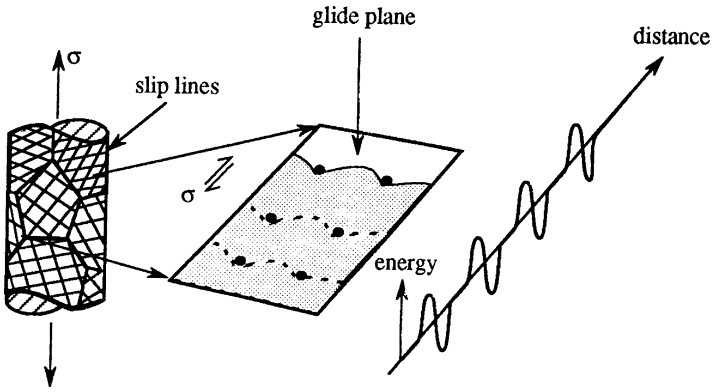


Fig. 4.19 Discrete obstacles in a slip plane, opposing the passage of a dislocation (Adapted from Frost and Ashby 1982)

– expressing the macroscopic deformation as the mean over many slip planes.
– for all these reasons we can only claim to measure mean values, given by:

$$\Delta\bar{G}_A = -\mathrm{k}\left[\frac{\partial \log\left(\dot{\varepsilon}^{\mathrm{v}}/\dot{\varepsilon}_0\right)}{\partial\left(1/T\right)}\right]_T$$
$$\bar{A} = \frac{\mathrm{k}Tm}{b\sigma} = \frac{\mathrm{k}T}{b}\left[\frac{\partial \log\left(\dot{\varepsilon}^{\mathrm{v}}/\dot{\varepsilon}_0\right)}{\partial\sigma}\right]_T \tag{4.27}$$

4.3.1.3 Nature and Resistance of Obstacles

The obstacles met by a dislocation in the course of gliding can be put into two main categories, shown diagrammatically in Figs. 4.19 and 4.20:

– discrete, either already existing, such as precipitates or solid-solution atoms,
– or induced by the deformation, such as trees in the forest of dislocations,
– diffuse, such as lattice forces, which are overcome by the dislocation.

The strain rate is imposed either by the kinetics of overcoming discrete obstacles (Fig. 4.19) or by that of development and propagation of the kinks (Fig. 4.20). Table 4.3 gives the orders of magnitude of the energies and stresses ($\tau_0 = \tau_{\mathrm{B}} + \tau_{\mathrm{s}}$) necessary to overcome the obstacle in the absence of thermal activation, that is, at absolute zero.

We now show how the formalism of thermal activation of plastic deformation can be used to treat three examples concerning obstacles of the two main classes.

(a) *Pre-existing obstacles: precipitates*
Hardening by precipitation provides an example of pre-existing obstacles (see Exercise 4.6). We recall that if these are perfectly coherent with the matrix then

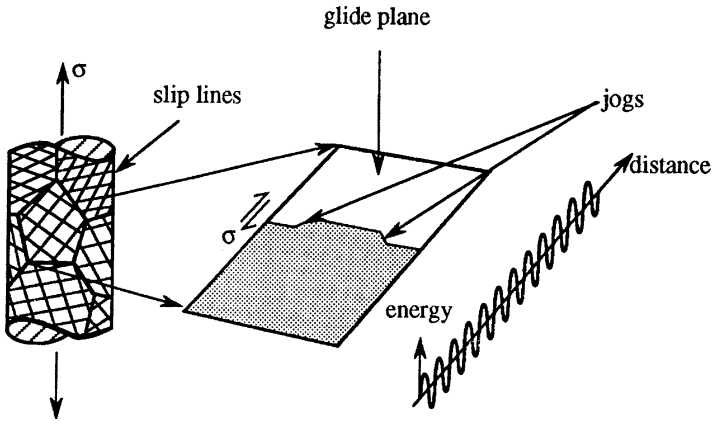


Fig. 4.20 Diffuse obstacles formed by kink nucleation and propagation of a dislocation through the Peierls troughs (Adapted from Frost and Ashby 1982)

Table 4.3 Energies and stresses needed to overcome various obstacles (Ashby and Jones 1980)

Obstacle Force	Energy Q_0	τ_0 (0 K)	Examples
Strong	$\approx \mu b^3$	$\approx \mu b/l$	Precipitates by-passed by dislocations
Medium	$\approx 0.2\mu b^3$	$\approx \mu b/l$	Dislocations of the forest Small sheared precipitates, Irradiation defects
Weak	$< 0.2\mu b^3$	$<< \mu b/l$	Lattice forces Solid solution

the dislocation can shear them if they are small or by-passing them if they are above some critical size. In the case of ordered precipitates, it is of the order of t/γ_A where t is the line tension of the dislocation and γ_A is the anti-phase energy corresponding to the shearing of the particles (see Sect. 3.4.3.2).

Whether the precipitates are sheared or by-passed, it is easy to show from the equilibrium equation that the dislocation has two equilibrium positions: one corresponding to the start (minimum on the energy curve), the other at the break-out (maximum of the energy curve) (Fig. 4.21, b). The way in which the dislocation crosses the energy jump ΔG_A, when the stress increases (and therefore when the temperature decreases) is indicated schematically in Fig. 4.21. The exact expressions giving the total energy and the form of the curves of the diagram are (Exercise 4.6):

– for particle shearing:

$$\Delta E = t\left(\Delta l_m + \Delta l_p\right) + \gamma_A \Delta S_p - \tau b\left(\Delta S_m + \Delta S_p\right) \tag{4.28}$$

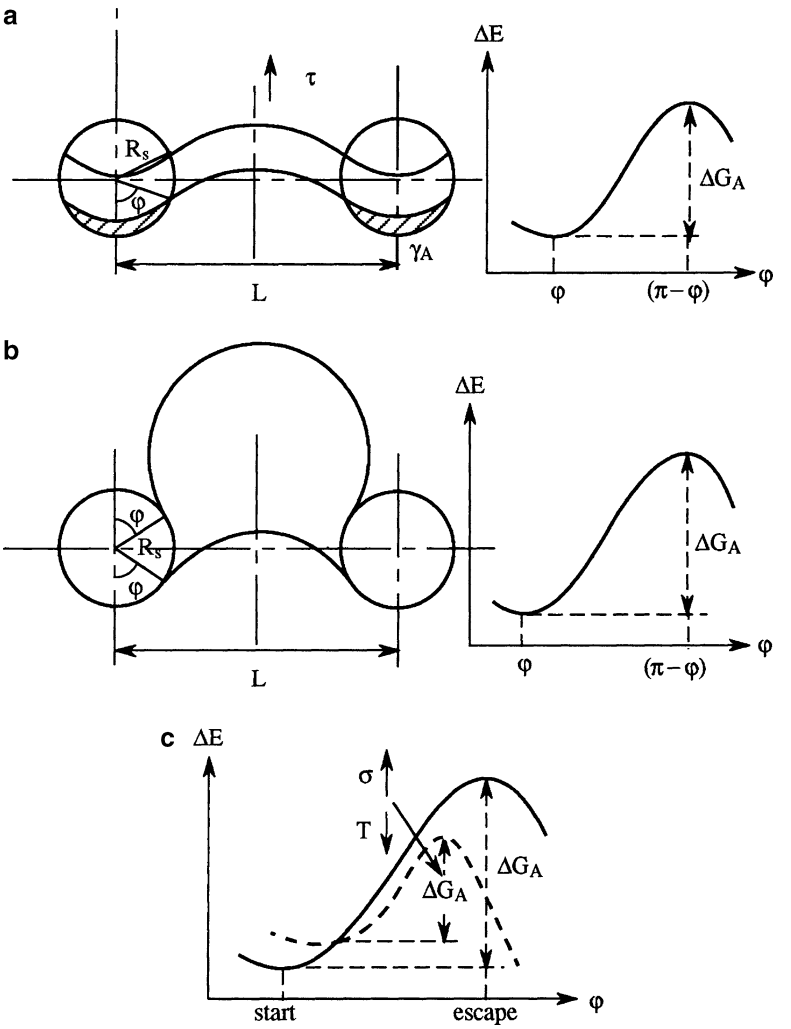


Fig. 4.21 Equilibrium positions for the crossing of two precipitates by a dislocation line by (**a**) shearing; (**b**) by-passing; (**c**) positions of the corresponding energy extrema and the effect of temperature (and thus of a given strain rate) (see also exercise 4.6)

– for by-passing:

$$\Delta E = t\Delta l_{\mathrm{m}} - \tau b \Delta S_{\mathrm{m}} \tag{4.29}$$

where Δl_{m}, Δl_{p} are the increases in length of the part of the dislocation in the matrix and the precipitate respectively, relative to some arbitrarily-chosen reference position, and ΔS_{m}, ΔS_{p} are the increases in the areas swept out by this dislocation.

Since for obstacles of this type the probability of a return jump is very low we can use (4.10) to determine the strain rate. Thus:

$$\dot{\gamma}^{\mathrm{v}} = N_{\mathrm{m}} \frac{\Delta A b}{V} \nu_0 \exp\left(-\frac{Q_0 - \tau b A^*}{\mathrm{k}T}\right) \tag{4.30}$$

where N_{m} is a constant and $\Delta A \approx L\lambda$. The values of Q_0 and A are easily determined for the two modes from the expressions already given (see Sect. 4.3.1.2). Thus for shearing under high stress, meaning in the region of absolute zero, it can be shown that Q_0 has the expected form, that is:

$$Q_0 = \pi R_{\mathrm{S}}^2 \gamma_{\mathrm{A}} \tag{4.31}$$

and

$$\begin{aligned} A^* &= 2L R_{\mathrm{S}} \left[1 - (\tau/\tau_0)^2\right]^{1/2} \\ \tau_0 &= 2\gamma_{\mathrm{A}}/R_{\mathrm{S}} \end{aligned} \tag{4.32}$$

The height of the activation barrier can be written:

$$\Delta G_{\mathrm{A}} = Q_0 [1 - (\tau/\tau_0)]^{3/2} \tag{4.33}$$

At the macroscopic level, the strain rate can be expressed in terms of temperature and stress, thus:

$$\dot{\varepsilon}^{\mathrm{v}} = \dot{\varepsilon}_0^{\mathrm{v}} \exp\left[-\frac{Q_0}{\mathrm{k}T}(1 - \sigma/\sigma_0)^{3/2}\right] \tag{4.34}$$

where $\dot{\varepsilon}_0^{\mathrm{v}}$ is constant.

This relation, with the others of the same type developed in this section, is commonly used to represent the variation of the yield strength R_{p} with temperature and strain rate.

(b) *Discrete obstacles, deformation-induced*: *trees in the forest of dislocations*.
The hardening caused by attractive and repulsive junctions between dislocations has been studied in Sect. 3.4.2. We showed that if there is no thermal activation a critical stress σ_0 has to be applied to move a dislocation through the forest. This critical stress is given by:

$$\tau_0 = \alpha \mu b \sqrt{\rho} \tag{4.35}$$

with α approximately between 1/3 and 1/4.

The energy required to overcome the junctions is of the order of $\kappa \mu b^3$ with: $1/5 < \kappa < 1$.

In the Orowan equation (4.7) the density of obstacles, in contrast to (a) above, is an increasing function of the deformation. With the mean velocity of the dislocations written as:

$$\bar{v} = \beta b v_0 \exp\left(-\frac{Q}{\mathrm{k}T}\right) \tag{4.36}$$

where β is a numerical factor and with the relation between the density of mobile dislocations and applied stress written as:

$$\rho_{\mathrm{M}} = \alpha'(\tau/\mu b)^2 \tag{4.37}$$

where α' is a constant of the order of 1, the strain rate can in general be written:

$$\dot{\gamma}^{\mathrm{v}} = \alpha'\beta b^2 v_0 \left(\frac{\tau}{\mu b}\right)^2 \exp\left(-\frac{Q}{\mathrm{k}T}\right) \tag{4.38}$$

The activation energy is a decreasing function of the applied stress, as in the case of pre-existing discrete obstacles; if we assume this to be linear, that is:

$$Q = Q_0(1 - \tau/\tau_0) \tag{4.39}$$

we obtain:

$$\dot{\gamma}^{\mathrm{v}} = \dot{\gamma}_0 \exp\left[-\frac{Q_0}{\mathrm{k}T}(1 - \tau/\tau_0)\right]$$
$$\dot{\gamma}_0 = \alpha'\beta v_0(\tau/\mu)^2 \tag{4.40}$$

At the macroscopic level similar expressions are found replacing the resolved shear stress τ by the stress σ, the shear modulus μ by Young modulus E and the shear strain rate $\dot{\gamma}$ by the strain rate $\dot{\varepsilon}$.

(c) *Diffuse obstacles: low-temperature viscoplasticity controlled by lattice resistance*.

We first recall that, at low temperatures ($T/T_{\mathrm{m}} < 0.2$), the velocity dislocations in most metallic materials, particularly with BCC structures, is controlled by obstacles of another type, those associated with the atomic structure (the Peierls-Nabarro forces). The energy of a dislocation varies with its position, as shown in Fig. 4.20; it advances by propagating simple or double kinks and in general the velocity with which it moves from one trough to another is controlled by the rate at which kinks are generated. As before, the activation energy can be written:

$$Q = Q_0[1 - (\sigma/\sigma_0)^p]^q \tag{4.41}$$

where $Q_0 \cong 0.2\mu b^3$. σ_0 is the elastic limit at absolute zero and p and q are of the order of 1; according to Frost and Ashby (1982) the choices $p = 3/4$, $q = 4/3$ give the best agreement.

The density of kinks is an increasing function of the strain. If we assume that it varies in the same way as that of trees in the forest, we find that the strain rate is:

$$\dot{\varepsilon}^{\mathrm{v}} = \dot{\varepsilon}_0 \exp\left[-\frac{Q_0}{\mathrm{k}T}(1-\sigma/\sigma_0)^{3/4}\right]^{1/4}$$
$$\dot{\varepsilon}_0 = \dot{\varepsilon}'_0(\sigma/\mu)^2 \tag{4.42}$$

The relations we have just derived do not apply in the case of the very high strain rates (above 10^4 s^{-1}) that are found in certain conditions of shock loading or in certain forming operations.

4.3.1.4 Drag Force on Fast Moving Dislocations

Dislocation Velocity

In the electron microscope, the dislocations are usually seen to move very quickly from obstacle to obstacle. We want to determine the relation between the dislocation velocity and the force applied on them. The first determination of dislocation velocity was performed by Gilman (Johnston and Gilman 1959) on LiF crystals. He took advantage of the etch pit that is created at the emergence point of a dislocation on the surface of the crystal. Once the dislocation had moved away from the position where such an etch pit was created, its sides continued to grow leaving a flat bottom contrasting with the pyramidal shape of fresh etch pits. It was thus possible to measure the distance travelled by a dislocation from a primitive etch pit to a new one under a stress pulse of given duration. As the displacement of dislocations can be a jerky one, this experimental technique provides a mean velocity. The result is that the dislocation velocity in LiF crystals is proportional to τ^{25} up to a velocity of the order of 10 ms^{-1}. Above that the stress dependency of the velocity is not so high, but never reaching the shear wave velocity in keeping with the relativistic behaviour described in Sect. 3.3.7.8 (Fig. 4.22). Determination of dislocations velocity was then achieved for Fe-3.35Si (Stein and Low 1960) yielding a stress exponent m of 41 at room temperature and of 38 at 198 K. The same behaviour has been found for a number of materials. Another empirical relation between dislocation velocity and stress is an exponential function. Table 4.4 gathers various results (see also Fig. 4.23). Recent *in-situ* TEM experiments were performed by Caillard (2010a and b) on pure iron. This author clearly showed the influence of temperature on the dislocation movement. He measured the velocity-stress dependence of screw dislocations at the scale of a single dislocation source, and compared those results with macroscopic measurements of activation areas.

Equation 4.8 shows that the dependence of the proof strength on strain rate should follow the same kind of law. This is the case (Fig. 4.24). For strain rates higher than about 10^4 s^{-1} a much faster increase of the stress needed to achieve a given strain rate is observed.

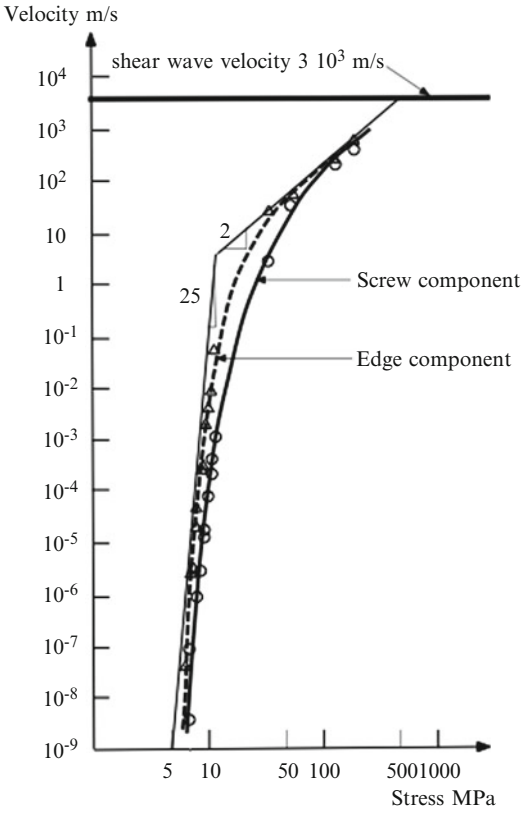


Fig. 4.22 Stress dependence of dislocation velocity in LiF (Adapted from Johnston and Gilman 1959)

Table 4.4 Various experimental results of dislocation velocity versus stress, $v = (\tau/\tau_0)^m$

Material	τ_0 (Pa)	m
Al		1
Cu	$2.7 \cdot 10^3$	1
Cu-1Al		17 at 77 K
		11 at 373 K
Pb		1
Zn	$2.94 \cdot 10^4$	1
Basal plane screw		3.4 at 300 K
edge		3.5 at 300 K; 1.2 at 77 K
Pyramidal screw		16 at 300 K; 7.1 at 77 K
edge		27 at 300 K; 11 at 77 K
Fe-3.35Si		30.1
		38 at 198 K
		41 at 373 K
Mo		6.4
LiF		25

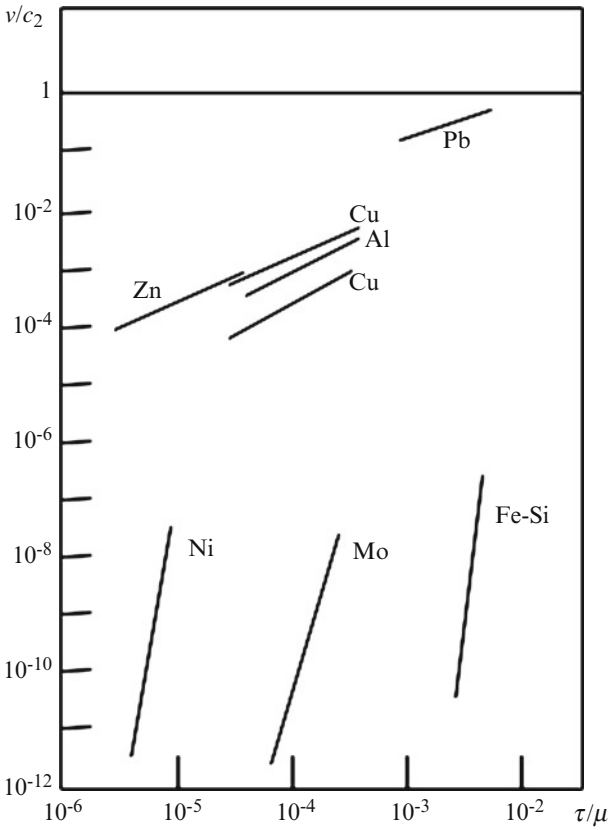


Fig. 4.23 Stress dependence of dislocation velocity for various materials (Adapted from Meyers and Chawla 1984)

Drag Mechanisms

If we evaluate the total energy of the dislocations in a strain-hardened material with respect to the work done by the applied stress, we find that the ratio is of the order of only 10%. Ninety percent of the work is dissipated as heat and it is common experience to feel the temperature increase of a metal after deformation. This means that there are dissipative phenomena occurring while moving the dislocations. They are linked with a drag force on the dislocations. The viscous behaviour would be expressed as:

$$v = \frac{\tau b}{B} \tag{4.43}$$

with B the damping coefficient in Nsm^{-1}.

The lower part of the velocity-stress curve with high m exponents falls out of this formalism. It can be interpreted as a thermally activated process. The upper part of

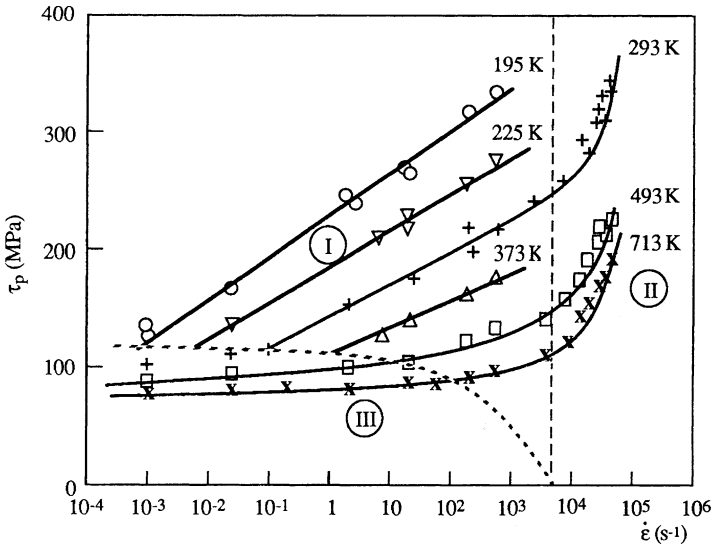


Fig. 4.24 Effect of temperature and strain rate on the shear yield strength for mild steel (Campbell and Ferguson 1970)

the velocity-stress curve, with a stress exponent approaching 1 is interpreted as the effect of a drag force on the dislocation.

A number of damping mechanisms were proposed (Nabarro 1967). There remain two of them, which yield high enough drag forces to cope with experimental observations (Mason 1968). They are electron and phonon viscosities. They result from the conversion of the moving strain field of a dislocation into heat owing to interactions with free electrons or with phonons, in the same way as sound waves are attenuated. As shown in Fig. 4.25 the electron viscosity is the controlling mechanism at low temperatures, whereas it is the phonon viscosity at higher temperatures, in particular at room temperature.

Substituting (4.43) in (4.8) we get:

$$\dot{\gamma}^{\mathrm{v}} = \frac{\rho_{\mathrm{M}} b^2}{B} \tau \tag{4.44}$$

This relation is transposed at the macroscopic level by substituting $\dot{\varepsilon}^{\mathrm{v}}$ to $\dot{\gamma}^{\mathrm{v}}$ and σ to τ.

At very high rates the ratio ρ_{M}/B is practically independent of stress and of temperature, so that this relation becomes an expression of Newtonian viscous flow (Fig. 4.26).

Summary As we have just shown, this type of low-temperature viscoplasticity is observed especially in materials with covalent bonds or with BCC crystal structures. Figure 4.24 represents the effects of strain rate on the yield strength of a mild steel, for various temperatures. Three regions can be distinguished, as indicated:

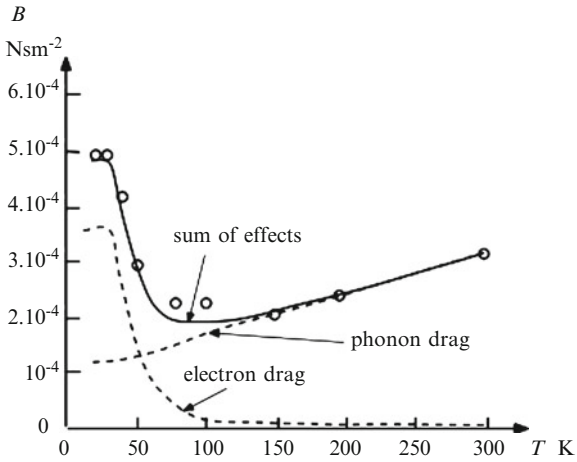


Fig. 4.25 Dislocation drag coefficient for aluminium (Adapted from Mason 1968)

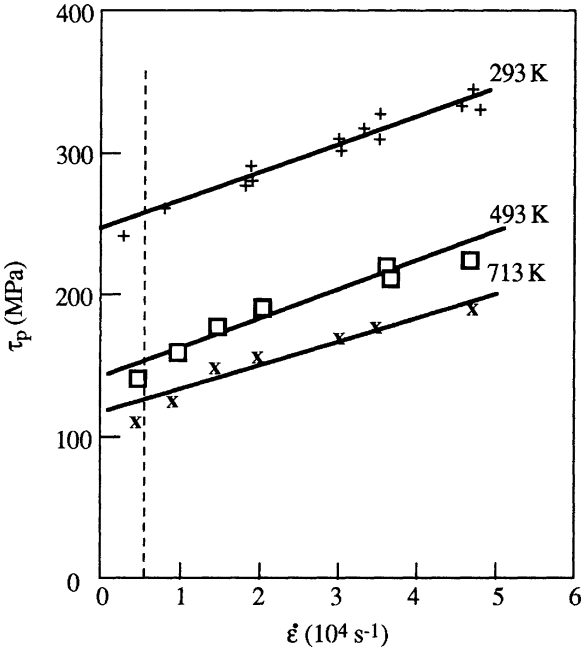


Fig. 4.26 Yield strength in shear τ_p as a function of strain rate, for a mild steel at very high strain rates ($>10^4$ s^{-1}) (Campbell and Ferguson 1970)

1. Low temperature (<293 K) and low-to-moderate strain rates ($<10^3$ s^{-1})
2. Very high rates ($>10^4$ s^{-1}), in which the yield strength is effectively a linear function of the rate; this is the region of phonon drag.
3. A region in which the yield strength varies very little with rate and temperature, corresponding to the athermal plateau of Fig. 4.16. The activation energy is now independent of temperature and applied stress and is sufficiently high for the thermal activation energy kT to be comparable to what is needed to overcome the obstacles.

4.3.2 *Physical Models of High-Temperature Viscoplasticity in Metals*

As we stated in the introduction to Sect. 4.3, at high temperatures ($T > 0.5\ T_m$) viscoplasticity involves diffusion phenomena: short-range *dislocation creep* controlled by the movements of the dislocations and longer-range *diffusional creep* over distances of the order of the grain size, when the deformation is caused by movement of the material due to trans- or intergranular diffusion. We shall study these two modes of deformation in order, together with the physical models, which have been suggested for their interpretation. These models are based mostly on simple creep experiments – which are why the terms dislocation creep and diffusional creep are used.

It is beyond the scope of this chapter to describe in detail all the many models that have been suggested for creep. More details can be found in text books devoted to creep (see *e.g.* Garofalo, 1965; Poirier, 1985). We therefore limit our treatment to introducing the concepts underlying two main classes of models: *recovery creep* and *creep controlled by dislocation climb*; further, we shall consider only steady-state creep.

Diffusional creep, which acts at the higher temperatures and lower stresses, enables viscous deformation to occur without the influence of dislocations; inelastic deformation occurs by transport of material along grain boundaries (*Coble creep* see Coble 1963) or, at still higher temperatures ($T > 0.8\ T_m$), within the grains (*Herring-Nabarro creep* see Herring (1950), Nabarro (1947, 1952)). The behaviour is thus perfectly viscous and the threshold constant σ_s in Eq. 4.6 is practically zero.

4.3.2.1 Dislocation Creep

(a) Recovery creep. Many models have been based on the original idea of Bailey and Orowan, according to which a stationary state is reached when the work-hardening that normally results from the accumulation of deformation and the increase in density of dislocations is compensated for by the recovery, a phenomenon regulated primarily by time and diffusion (see Sect. 3.4.2.2i).

When the diffusion is fairly rapid the obstacles to slip can vanish, which means that in Fig. 4.18 we can set $\sigma_B = 0$. The only barrier that remains is that due to the internal stress, and therefore $\sigma = \sigma_s$. The deformation is no longer activated thermally in the strict sense of that term, but the rate depends on that of the recovery, which is activated thermally through the intermediary of the diffusion.

If the creep strain rate is to remain constant so must the internal stress σ_s; thus:

$$\frac{d\sigma_S}{dt} = \left(\frac{\partial \sigma_S}{\partial \varepsilon}\right)_t \dot{\varepsilon}^v + \left(\frac{\partial \sigma_S}{\partial t}\right)_\varepsilon = 0 \tag{4.45}$$

If we define $r = -(\partial\sigma_S/\partial t)_\varepsilon$, $h = (\partial\sigma_S/\partial\varepsilon)_t$ for the recovery and work-hardening rates respectively we have:

$$\dot{\varepsilon}^v = r\,/h \tag{4.46}$$

The models proposed for this region differ from one another mainly in the way r and h are evaluated and in the assumptions made in obtaining their values.

An approach due largely to McLean (1962) shows that creep-recovery models lead naturally to the empirical Eq. 4.1c. His original model assumes that the internal stress arises solely from a three-dimensional network of dislocations of mean length λ; this could correspond to the cells observed in creep at high stress. The internal stress is then inversely proportional to λ, that is:

$$\sigma_S = \mu b\,/\lambda \tag{4.47}$$

Further, it is assumed that the deformation results from the gliding of dislocations in this network of cells, so that the Orowan relation can be written:

$$\varepsilon = \rho_M b\lambda \tag{4.48}$$

If we assume that the density of mobile dislocations (ρ_M) is proportional to the overall dislocation density we can write $\rho_M = \alpha/\lambda^2$ where α is the constant of proportionality; and taking this together with (4.47) we get:

$$h = (\partial\sigma_S\,/\partial\varepsilon)_t = \mu\,/\alpha = \text{constant} \tag{4.49}$$

In this relation it is assumed that α itself is constant. Various approaches have been made in order to determine the velocity of recovery, of which the following is the simplest. A network of cells will tend to increase its size so as to reduce the total length of dislocations and therefore the stored energy. Friedel (1964) assumes that this growth is controlled by the climbing of the arcs of dislocations, which exchange vacancies. The flow of vacancies is stimulated by the difference in free energy of those in equilibrium in the neighbourhood of the different arcs.

The recovery rate shown by the increase of cell size with time ($\mathrm{d}\lambda/\mathrm{d}t$) will be proportional to the diffusion coefficient D_v and the driving force $\sigma\Omega/\mathrm{k}T$, where $\Omega \approx b^3$ is the volume occupied by a vacancy, and this rate will increase with decreasing cell size. Thus we can write:

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = D_v \frac{1}{\lambda}\frac{\sigma_S b^3}{\mathrm{k}T} \tag{4.50}$$

so that:

$$r = -\frac{\partial\sigma_S}{\partial t} = -\frac{\partial\sigma_S}{\partial\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \frac{D_v}{\mathrm{k}T}\frac{{\sigma_S}^4 b}{\mu^2} \tag{4.51}$$

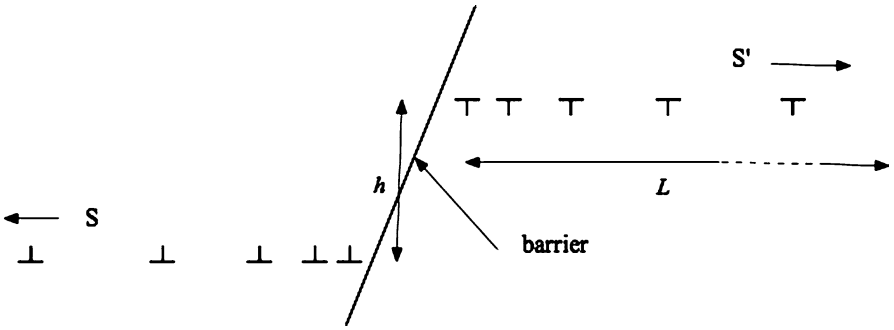


Fig. 4.27 Mechanism of recovery creep, by climb of dislocations located in pile-ups

From which, with (4.46), we find:

$$\dot{\varepsilon}^{\mathrm{v}} = \alpha \frac{D_0 b}{\mathrm{k}T} \frac{\sigma^4}{\mu^3} \exp\left(-\frac{Q_{\mathrm{D}}}{\mathrm{k}T}\right) \tag{4.52}$$

where Q_{D} is the self-diffusion energy.

Equation 4.52 can be written as, adopting the form proposed by Dorn (1957):

$$\frac{\dot{\varepsilon}^{\mathrm{v}} \mathrm{k}T}{D_{\mathrm{v}} \mu b} = \alpha \left(\frac{\sigma}{\mu}\right)^4 \tag{4.53}$$

Models of this type lead to laws of creep that are in reasonably good agreement with results found in experiments with pure metals, but as we have remarked previously, the exponent in the stress equation can differ from the value 4, and other models give values between 4 and 5. However, it is important to bear in mind that the agreement of the value of the exponent M deduced from a model with that found by experiment is never of itself a proof of the validity of the model.

(b) *Creep controlled by dislocation climb.* The mechanism of dislocation climb can enable recovery to take place. One model in particular, due initially to Weertman (1968), is based on the questionable assumption that this mechanism acts through the intermediary of a configuration of dislocations such as in Fig. 4.27, in which there are two pile-ups of edge dislocations of opposite signs, emitted by sources S and S', a distance h apart and resting against a barrier which can be either a grain boundary or another pile-up of dislocations. The stresses at the head of the pile-up, in contrast to the case of low-temperature plasticity, can be relieved by the climb and annihilation of the dislocations there. When a dislocation at the head has climbed the source can become active again and emit a new dislocation. Thus as in the general model of recovery creep there is competition between the production and the loss of dislocations.

An estimate of the corresponding strain rate can be derived from the equation for the dislocation climb:

$$\dot{\varepsilon}^{\mathrm{v}} = bANc\,/h \tag{4.54}$$

where A is the area swept by the dislocation in moving between the source and the head of a pile-up of length L, approximately L^2, N is the number of sources per unit volume and c is the mean velocity of climb.

We know that if lattice friction can be neglected the local stress at the head of the pile-up is (Sect. 3.3.8.2):

$$\sigma_1 \frac{\pi\,(1-\nu)\,L}{\mu b}\sigma^2 \simeq \frac{2L\sigma^2}{\mu b} \tag{4.55}$$

The component of this stress normal to the slip plane favours dislocation climb. Vacancies will be created or destroyed according as it is a tension or a compression; if C_0 is the equilibrium concentration of vacancies in the neighbourhood of the pile-up this is expressed by:

$$C_S = C_0 \exp\left(\pm\frac{2L\sigma^2 b^2}{\mu \mathrm{k}T}\right) \tag{4.56}$$

A flow of vacancies will be set up between the pile-ups of opposite signs, giving a velocity of climb:

$$c = \frac{1}{AN}\frac{2D_{\mathrm{v}}\sigma^2 b}{\mu \mathrm{k}T} \tag{4.57}$$

Proof of these relations follows the same lines as for diffusional creep, which we shall give later in Sect. 4.3.2.2.

The height h of the climb between the pile-ups is derived from the force needed to separate a pair of dislocations of opposite signs, a distance h apart: this is:

$$\sigma = \frac{\mu b}{4\pi\,(1-\nu)\,h} \tag{4.58}$$

from which we find:

$$\dot{\varepsilon}^{\mathrm{v}} = \frac{2D_{\mathrm{v}}\sigma^2 b^2}{h\mu \mathrm{k}T} = 8\pi\,(1-\nu)\,\frac{D_{\mathrm{v}}\sigma^3 b}{\mu^2 \mathrm{k}T} \tag{4.59}$$

If we assume that both the number of sources and the sizes of the pile-ups remain constant, this gives a creep law with a smaller exponent than is normally found in pure metals. Weertman (1968) suggests that such an assumption is not valid, and that L^2 is proportional to σ, giving an exponent 4.4, which is closer to what is observed; but there is no theoretical basis for this. Further, a serious objection to this model is that the dislocation pile-ups on which it is based have never been observed in materials subjected to creep.

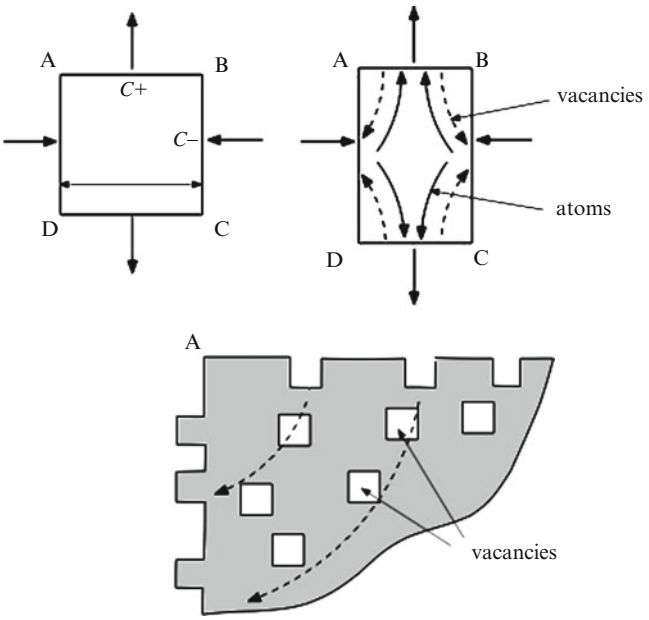


Fig. 4.28 Herring-Nabarro model: transport of vacancies from faces in tension (AB, CD) to faces in compression (AD, BC)

4.3.2.2 Diffusion Creep

At high temperatures ($T > 0.7\ T_m$) and low stresses the relation between deformation rate and applied stress is often found to be linear, as indicated in the introduction to this chapter. The deformation then results not from movement of the dislocations but from transport of material by diffusion, either across the grains (Herring-Nabarro model) or along the grain boundaries (Coble model). The activation energy for the Coble model is less than that for volume diffusion; consequently Coble creep will occur at lower temperatures than will Herring-Nabarro creep. These purely diffusive modes of transport are very important for the study of the strain compatibility between the grains of a polycrystal and intergranular slip; and further, they are the origin of the deformation of super-plastic alloys. We now consider the two in turn.

(a) *Herring-Nabarro creep*
Herring (1947, 1952) suggested that since in a stress field that is not purely hydrostatic the concentration of vacancies will differ on surfaces having different orientations, there will be a concentration gradient that can generate a flow of vacancies and consequently a flow of material in the opposite direction. We now give a simple form of the model based on this and later improved by Herring.

Consider first a cubic single crystal, of side d, subjected to shear (Fig. 4.28); we assume that it contains no dislocations and that the free surfaces are therefore the only sources and sinks of vacancies. Creating a vacancy on a surface in compression

(BC) is equivalent to moving an atom from this surface to the interior; if the atom is represented by a small cube of side b, the atomic volume Ω can be taken as approximately b^3. To leave the surface the atom must do work against the stress σ, and the energy thus dissipated will be $\sigma\ b^2$ x $b = \sigma\ b^3$, which is equivalent to saying that the energy of creation of a vacancy on BC is increased by $\sigma\ b^3$.

The corresponding atomic fractions will be:

– on the faces in tension, AB, DC:

$$n^+ = \exp\left(-\frac{\Delta G_f - \sigma b^3}{\mathrm{k}T}\right) = n_0 \exp\frac{\sigma b^3}{\mathrm{k}T} \tag{4.60a}$$

– on the faces in compression, AD, BC

$$n^- = \exp\left(-\frac{\Delta G_\mathrm{f} + \sigma b^3}{\mathrm{k}T}\right) = n_0 \exp\left(-\frac{\sigma b^3}{\mathrm{k}T}\right) \tag{4.60b}$$

where

$$n_0 = \exp\left(-\Delta G_\mathrm{f}/\mathrm{k}T\right) \tag{4.61}$$

is the atomic fraction in thermal equilibrium in the absence of applied stress and ΔG_f is the Gibbs free energy of creation of vacancies, of the order of the sublimation energy of the solid.

This results in a gradient in the concentration of vacancies that causes a flow of vacancies across the body of the grains and of atoms in the opposite direction. If the faces are perfect sources and sinks for vacancies then, in the steady state, the number transported across a face in tension, whose area is d^2, will be

$$\phi = -J d^2 \tag{4.62}$$

where, from Fick's law:

$$J = -D_\mathrm{L}\mathrm{grad}C \tag{4.63}$$

with D_L the diffusion coefficient for the vacancies and grad$C$ denoting the concentration gradient across the faces. We can write:

$$\mathrm{grad}C = \alpha\left(C^+ - C^-\right)/d \tag{4.64}$$

where α is a numerical factor, close to 1.

The concentration C is related to the atomic fraction n by

$$C = n/b^3 \tag{4.65}$$

and the diffusion coefficient $\mathrm{D_v}$ to $\mathrm{D_L}$ by:

$$D_\mathrm{v} = D_\mathrm{L} n_0 = D_\mathrm{L} C_0 b^3 \tag{4.66}$$

from which relations we find:

$$\phi = \frac{2\alpha D_{\mathrm{v}} d}{b^3} \sinh \frac{\sigma b^3}{\mathrm{k}T} \tag{4.67}$$

The escape of an atom from the face AB corresponds to an elementary extension, which can be calculated by regarding the atom, of volume b^3, as spread over the surface, of area d^2:

$$\varepsilon = \left(b^3 \,/ d^2\right)(1 \,/ d\,) = b^3 \,/ d^3 \tag{4.68}$$

Thus the deformation rate corresponding to the flux Φ is:

$$\dot{\varepsilon}^{\mathrm{v}} = \phi \frac{b^3}{d^3} = \frac{2\alpha D_{\mathrm{v}}}{d^2} \sinh \frac{\sigma b^3}{\mathrm{k}T} \tag{4.69}$$

which for small values of σ becomes:

$$\dot{\varepsilon}^{\mathrm{v}} = \frac{2\alpha D_{\mathrm{v}}}{d^2} \frac{\sigma b^3}{\mathrm{k}T} \tag{4.70}$$

Extending this simplified calculation to the case of a polycrystal presents a number of problems: we have to take into account first the fact that the grain boundaries can act as perfect sources and sinks, and second the actual shape of the grains. However, all cases lead to expressions of the form:

$$\dot{\varepsilon}^{\mathrm{v}} = B \frac{D_{\mathrm{v}} \sigma \Omega}{d^2 \mathrm{k}T} \tag{4.71}$$

where the factor B is a function of the geometry and Ω ($\approx b^3$) is the atomic volume.

Herring gives the value $B = 16$ for spherical grains.

The most important results of this model appear to be the following:

(i) diffusional creep can be expressed by a law of Newtonian viscous flow type
(ii) the rate of this creep increases with increasing temperature ($D_{\mathrm{v}} = D_0 \exp(-Q_0 \,/ \mathrm{k}T\,)$) and with decreasing grain size (like $1/d^2$).

(b) *Coble creep.* Coble (1963) has proposed a model for creep in polycrystals in which the rate is controlled by diffusion, not through the grain bodies but along the grain boundaries; this can be faster, since the intergranular diffusion energy is only about half that of volume diffusion. The bases for the model are the same as for Herring-Nabarro (Fig. 4.28). We give a simplified treatment here, referring the reader to exercises in Volume III for a fuller treatment.

Consider a grain in the form of a cube of side d. The vacancies and the atoms flow along the boundaries and if δ_{GB} is the thickness of the grain boundary the flux of vacancies is:

$$\phi = \alpha D_L \frac{\Delta C}{d} (4\delta_{GB} d) \tag{4.72}$$

Introducing the intergranular diffusion coefficient $D_{GB} = D_L C_0 b^3$ and substituting the expression for ΔC deduced from Eqs. 4.60a and 4.60b:

$$\Delta C = 2C_0 \sinh \frac{\sigma b^3}{kT} \simeq 2C_0 \frac{\sigma b^3}{kT} \tag{4.73}$$

we find:

$$\phi = 8\alpha \delta_{GB} D_{GB} \frac{\sigma}{kT} \tag{4.74}$$

Proceeding as before we find for the strain rate:

$$\dot{\varepsilon}^v = 8\alpha \delta_{GB} D_{GB} \frac{\sigma b^3}{kT d^3} \tag{4.75}$$

In general, this can be written:

$$\dot{\varepsilon}^v = B' \delta_{GB} D_{GB} \frac{\sigma \Omega}{kT d^3} \tag{4.76}$$

where B' is a constant and δ_{GB}, the effective width of the boundaries for diffusion, is of the order of $10b$ for metals. The difference between the Herring-Nabarro and Coble models lies in the effects of the grain size and of the activation energy. Experiments with copper have shown creep rates varying like $1/d^3$ at low temperature (550°C) and like $1/d^2$ at higher temperatures (840°C).

(c) *Compatibility of intergranular deformation: sliding at grain boundaries*

The Herring-Nabarro models are based on simplifying assumptions concerning the shape of the grains and the homogeneity of their deformation. However, these may not hold to better than a first approximation, in particular because a set of perfectly spherical grains is not compact. Further, while the deformation is occurring the grains must remain in contact so long as there is no intergranular damage. This compatibility condition for the deformation accounts for the development of intergranular-type stresses, as in plasticity, which at high temperatures can be largely relieved by viscoplasticity, with the grain boundaries sliding with respect to each other.

The importance of intergranular sliding as a form of viscoplastic deformation is made clear by the displacement of reference marks, as shown in Fig. 4.29. If we define this intergranular deformation ε_{GB} as $\Delta u_{GB}/d$, where Δu_{GB} is the

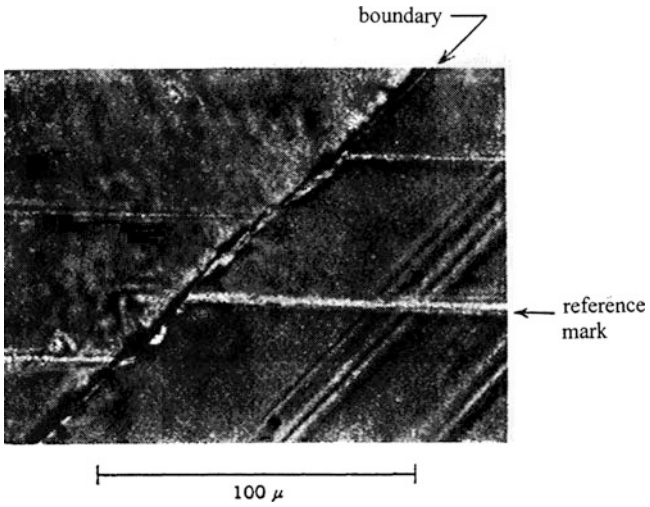


Fig. 4.29 Intergranular slide shown by displacement of reference marks in a bicrystal of lead (Garofalo 1965)

displacementof a reference mark and d is the grain size (an approximation that in particular does not take into account the orientation of the boundary plane with respect to the plane of observation) we can use as a measure of the importance of this deformation mode the ratio:

$$\lambda = \varepsilon_{GB} / \varepsilon_T \tag{4.77}$$

where ε_T is the total deformation

The value of λ varies greatly with applied stress and less rapidly with temperature and grain size. Fig. 4.30 gives some results for various materials reported by Garofalo (1970); these can be put into two groups, the first consisting of aluminium and its alloys, β-brass and tin, the second of copper, copper-beryllium, α-iron and stainless steel. In the first, λ decreases very quickly as the stress, and consequently the strain rate, increases; the contribution of the intergranular slip to the total deformation falls fairly quickly to below 5%. In the second the rate of decrease is much less. In both groups the deformations combine to form subgrains, but again much more quickly in the first than in the second. These results suggest that there is a relation between intergranular sliding and the changes to the microstructure associated with intergranular deformation.

Intergranular sliding and diffusional creep are closely related (Raj and Ashby 1971). The resultant deformation can be described as due either to diffusional creep of the grains accommodated by slip at the boundaries, or reciprocally to slip at the boundaries accommodated by diffusion creep of the grains. Figure 4.31 gives a simple illustration of this.

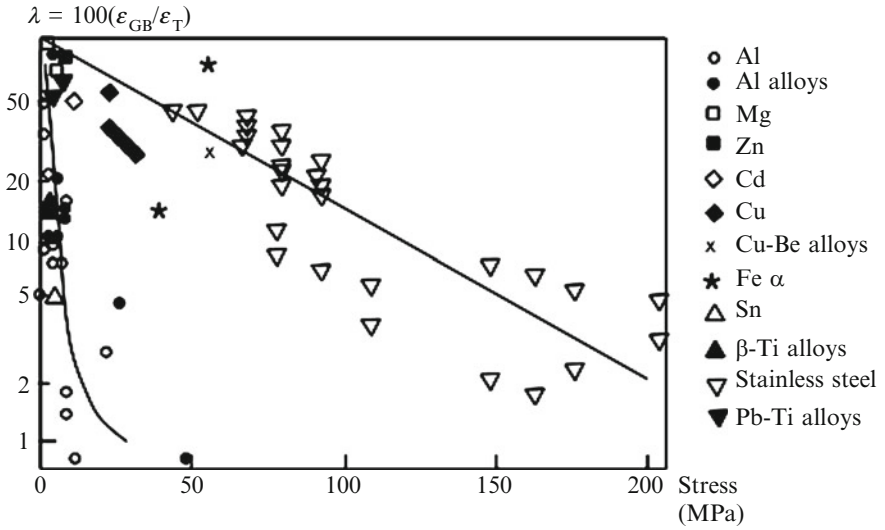


Fig. 4.30 Effect of applied stress in creep tests on the ratio of boundary slide deformation (ε_{GB}) to total deformation (ε_T) for various metals and alloys (Garofalo 1965)

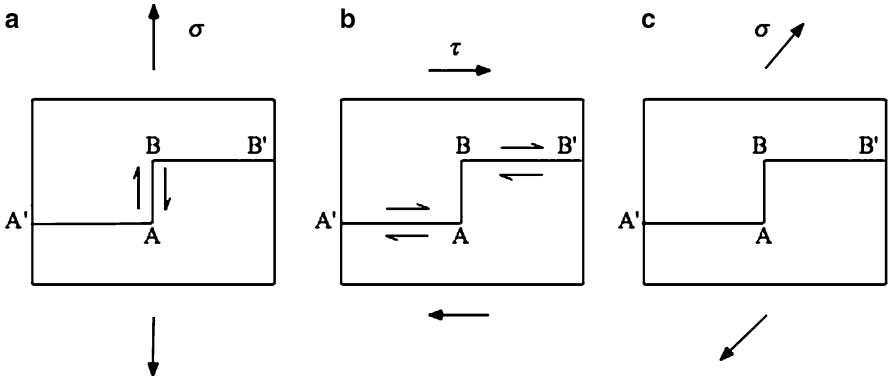


Fig. 4.31 Sketch of boundary sliding in a bicrystal, enabling strain compatibility to be achieved

In Fig. 4.31a the tensile stress favours transport of material over the surfaces AA′, BB′, which must be accommodated by slip along the boundary AB. In Fig. 4.31b the shear stress favours intergranular sliding which must be accommodated by transport of material over AB. The general situation, as in Fig. 4.31c, will be intermediate between these extremes, necessitating both sliding and diffusion: both contribute to the deformation and each provides the accommodation required by the other.

Raj and Ashby (1971) have made a detailed study of a non-planar boundary (Fig. 4.32), considering the case in which the incompatibilities (or internal stresses) created by the slip are accommodated either elastically or by diffusion. They showed

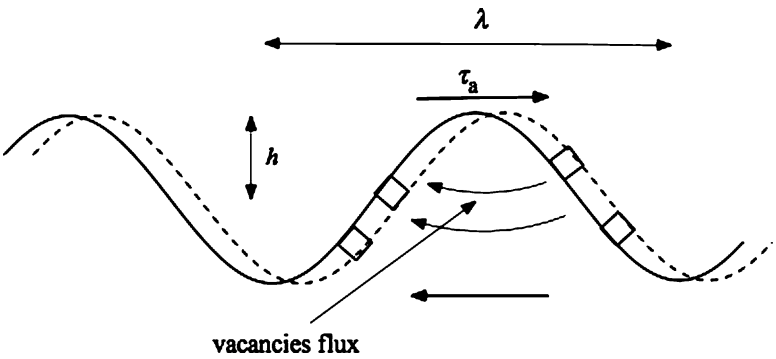


Fig. 4.32 Diagram showing boundary glide with accommodation by diffusion. The boundary is displaced to the right

that the way in which this accommodation is made controls the strain rate. In this work they considered the polycrystal as a two-dimensional compact set of hexagons.

If diffusion is assumed to occur both in the bulk and along the boundary, and that the latter can be represented by two sine waves of amplitude $h = \delta_{GB}/2$ and wave length λ, the strain rate is found to be:

$$\dot{\varepsilon}^{v} = B'' \frac{\sigma \Omega}{kT} \frac{D_v}{d^2} \left(1 + \frac{\pi \delta_{GB}}{\lambda} \frac{D_{GB}}{D_v} \right) \tag{4.78}$$

where B'' is a numerical constant, of value about 40.

The form of this result will be seen to combine volume diffusion creep (Herring-Nabarro) with boundary diffusion creep (Coble): this is because the hexagonal arrangement gives $\lambda = d$.

4.3.2.3 Creep Diagrams

It is helpful to gather together all the experimental information and all the theoretical models concerning viscoplastic behaviour and present this in the form of maps or charts, as did Ashby for a number of polycrystalline materials (see Frost and Ashby 1982). Figure 4.33 shows one such presentation, for pure polycrystalline nickel with a grain size of 100 μm. In this map the dimensionless stress variable is the shear stress related to the shear modulus; it is plotted as ordinate against the absolute temperature related to the melting temperature $T/T_{m,}$. The various regions that we have described are identified. One such region is that defined by high stresses and low temperatures, corresponding to low-temperature viscoplasticity, which in the case of pure nickel, an FCC material, is of small dimension. At the other extreme of low stress and high temperature are two regions corresponding to diffusional creep: Coble creep in the lower part of the temperature range

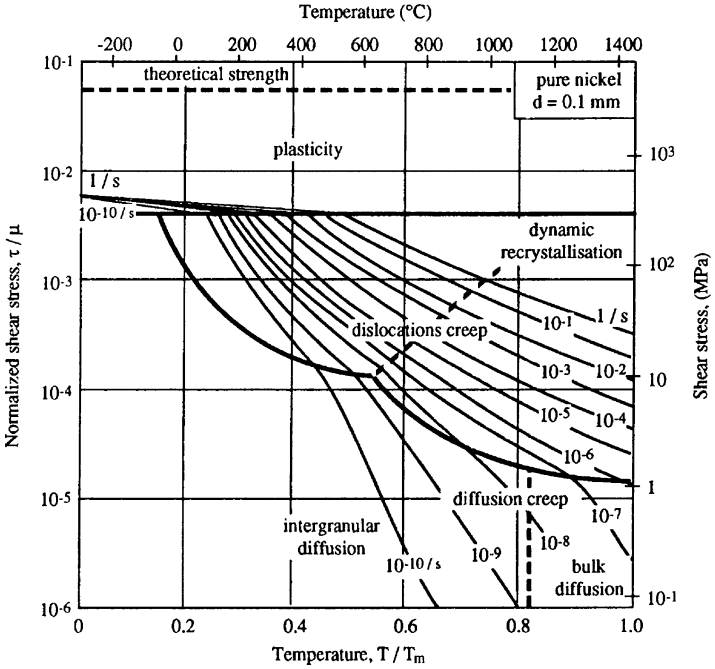


Fig. 4.33 Stress-temperature map for pure nickel of grain size 100 μm (Adapted from Frost and Ashby 1982)

($T/T_m < 0.8$), Herring-Nabarro in the upper part ($T/T_m > 0.8$). The rest of the map corresponds to dislocation creep.

At high stresses and temperatures the map reported in Fig. 4.33 shows a region of dynamic recrystallisation, which we have not considered in this chapter; it was mentioned in the general introduction of this chapter, and in Sect. 3.4.2.2i. In the condition shown in Fig. 4.33 nickel recrystallises as it is being deformed, hence the term "dynamic".

In addition to showing the regions in which the different mechanisms are dominant the maps have curves of constant shear strain rate, where this is defined by $\dot{\gamma} = \left(2\varepsilon_{ij}\varepsilon_{ij}\right)^{1/2}$. Care must be taken when using these to avoid attributing too great a precision to them, since the experimental observations and the models derived from them are necessarily not perfect; but they are very useful for several purposes, for example:

- finding the order of magnitude of the strain rate for a given applied stress and temperature
- identifying the type of mechanism that will be operating in given conditions and, consequently, the type of law that will govern the creep behaviour. This can remove the need to make risky extrapolations from data obtained under

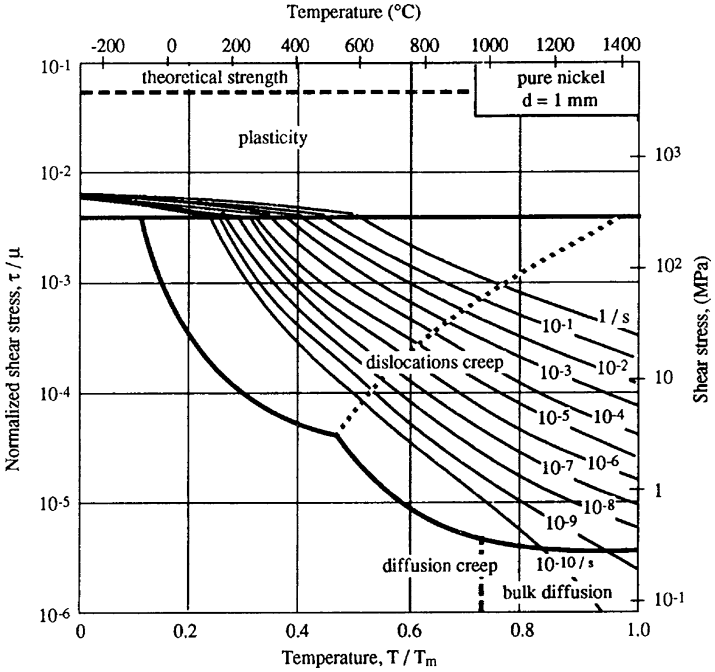


Fig. 4.34 Deformation map for pure nickel of grain size 1 mm. (Adapted from Frost and Ashby 1982)

conditions in which one mechanism is operating (dislocation creep, say) to conditions in which a different one (diffusional creep, say) dominates.

These deformation diagrams have been especially well determined for pure metals. Thus comparison of Figs. 4.33 and 4.34 will show the effect of grain size on the behaviour of pure nickel: this makes it very clear that whilst at low temperatures reducing the grain size will increase the resistance to deformation, the reverse is true at high temperatures.

Figures 4.35 and 4.36 are for pure tungsten, a BCC metal very important for uses at high temperatures. Qualitatively, the diagrams for this metal and for nickel are quite similar. The main difference is at low temperatures ($T/T_m < 0.15$), where the yield strength increases much faster with decreasing temperature for a BCC metal than for FCC, because of the lattice (Peierls) forces. At high ($T/T_m > 0.5$) or even moderate temperatures the resistance of BCC metals such as tungsten is lower than that of FCC metals such as nickel, the difference arising from the faster diffusion in the less dense BCC structure.

Similar diagrams are available for a number of metallic alloys of great practical importance. Good examples are Figs. 4.37 and 4.38 for austenitic stainless steels. The 304 and 316 steels are of essentially the same composition, 15% Cr and 10–12% Ni, differing in that 316 has also about 2.5% Mo, the effect of which is

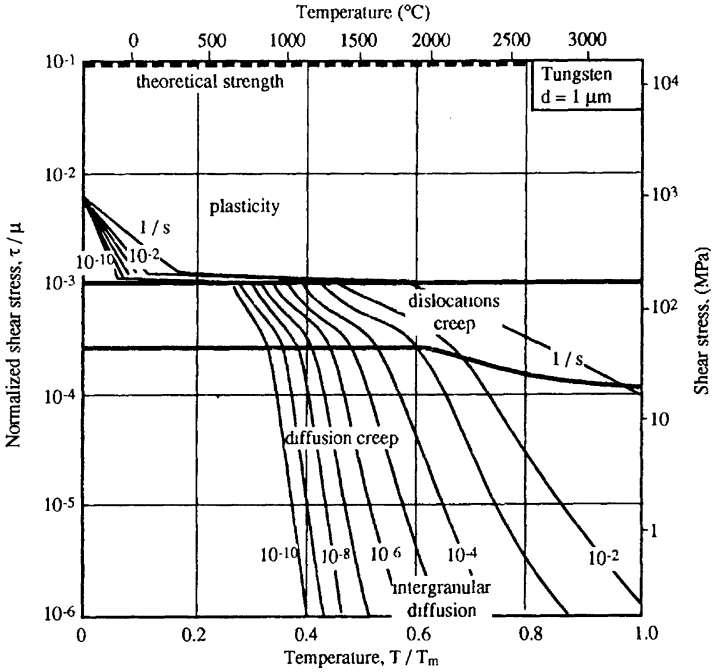


Fig. 4.35 Deformation map for pure tungsten of grain size 1 μm (Adapted from Frost and Ashby 1982)

to reduce the creep rate in the dislocation creep region by a factor of over 10. Both materials are single phase FCC.

Frost and Ashby (1982) gave data for materials other than metal alloys, including oxides of various types, and ice.

Similar deformation maps have been established for zirconium-based alloys, which are currently used as cladding tubes in the fuel assembly of nuclear light water reactors. These alloys have a low-temperature (α, HCP) phase and a high-temperature (β, BCC) phase. Although the normal operating temperature is lower than 400°C (*i.e.*, well within the α phase domain), one must ensure integrity and excess ballooning of cladding tubes up to high temperature and for high (internal pressure) stresses. In particular the behaviour of these materials must be investigated when the temperature is in the two phases ($\alpha + \beta$) domain.

Steady-state creep rate versus stress and temperature data are gathered in the map of Fig. 4.39. Two typical alloys are considered: Zircaloy 4 (Zr, Fe, Sn) (Zy – 4) and Zr – 1%NbO alloy. Two creep regimes were found in the near-α domain ($M \simeq 1$ and $\simeq 4$). In the β temperature range only one regime corresponding to the higher exponent ($M \simeq 4$) was found. Only one regime corresponding to the lower stress exponent (~ 1.4) was found for the ($\alpha + \beta$) temperature range. More details can be found elsewhere (Kaddour et al. 2004, 2011). Surprisingly, for applied stresses

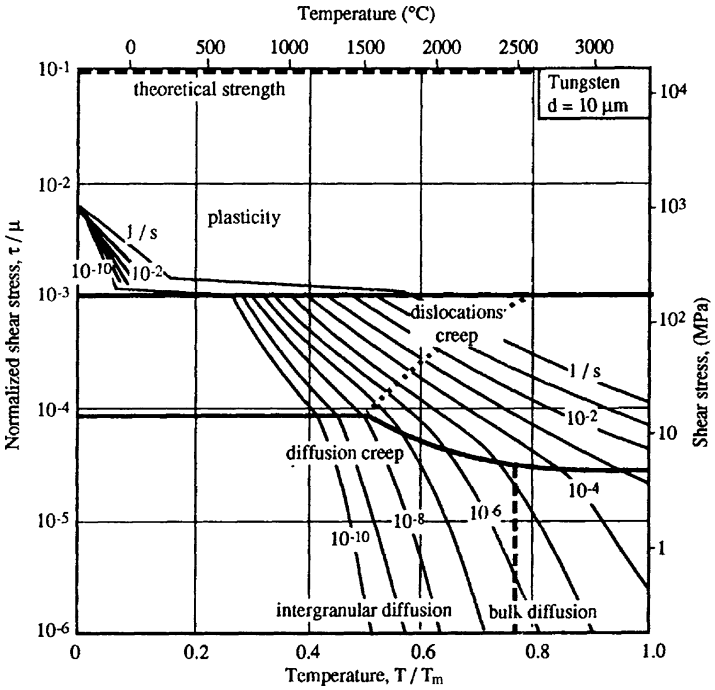


Fig. 4.36 Deformation map for pure tungsten of grain size 10 μm (Adapted from Frost and Ashby 1982)

lower than about 4 MPa, the two-phase materials show a higher creep strain rate than does the 100% β material at higher temperature. This behaviour is related to the most probable underlying mechanisms, which are high diffusion kinetics along α/β inter-phase interfaces and a strong grain size effect.

Note: In contrast to conventional "steady state" approaches, the idea that most normal creep curves should be envisaged only in terms of decaying primary and accelerating tertiary stages is an essential feature of the θ projection concept introduced by Evans and Wilshire (1985, 1993) and revisited more recently by Williams et al. (2010). The creep strain is written as:

$$\varepsilon = \theta_1 \left[1 - \exp\left(-\theta_2 t\right)\right] + \theta_3 \left[\exp\left(\theta_4 t\right) - 1\right] \tag{4.79}$$

where θ_1 and θ_3 quantify the primary and tertiary strains, while θ_2 and θ_4 are rate parameters governing the curvatures of the primary and tertiary components.

The systematic variations in creep curve shape have been quantified through plots of log θ_1 against σ/R_p, so that:

$$\theta_1 = G_1 \exp H_1 \left(\sigma \,/\, R_p\right) \tag{4.80a}$$

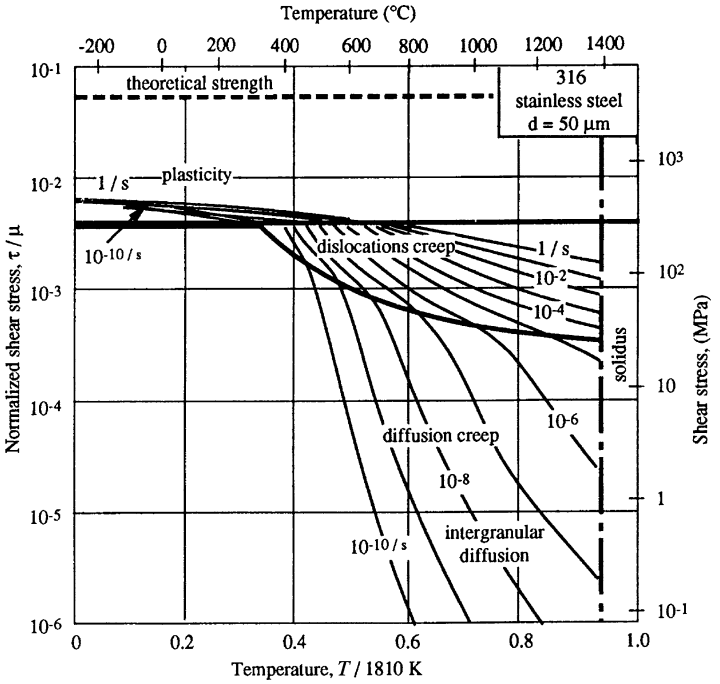


Fig. 4.37 Deformation map for austenitic stainless steel 316 (American standard) (Z05CND17-13, Mo content 2.5%), grain size 50 μm (Adapted from Frost and Jones 1982)

$$\theta_2 = G_2 \exp -\left[Q_2 - H_2\left(\sigma \,/\, R_\mathrm{p}\right)\right]/\mathrm{k}T \tag{4.80b}$$

$$\theta_3 = G_3 \exp H_3\left(\sigma \,/\, R_\mathrm{p}\right) \tag{4.80c}$$

$$\theta_4 = G_4 \exp -\left[Q_4 - H_4\left(\sigma \,/\, R_\mathrm{p}\right)\right]/\mathrm{k}T \tag{4.80d}$$

where G_i and H_i are constants, while Q_2 and Q_4 are the activation energies associated with θ_2 and θ_4, respectively.

4.3.2.4 Super-Plasticity

(a) *Introduction*

A material is said to be *super-plastic* if it can sustain elongation to fracture without breaking - a rather loose definition - in that the "elongation to fracture" can vary considerably ($\varepsilon_\mathrm{R} = 100$–$1{,}000\%$). It is a very familiar property in the case of a number of non-crystalline materials – in glass-blowing, for example, or certain

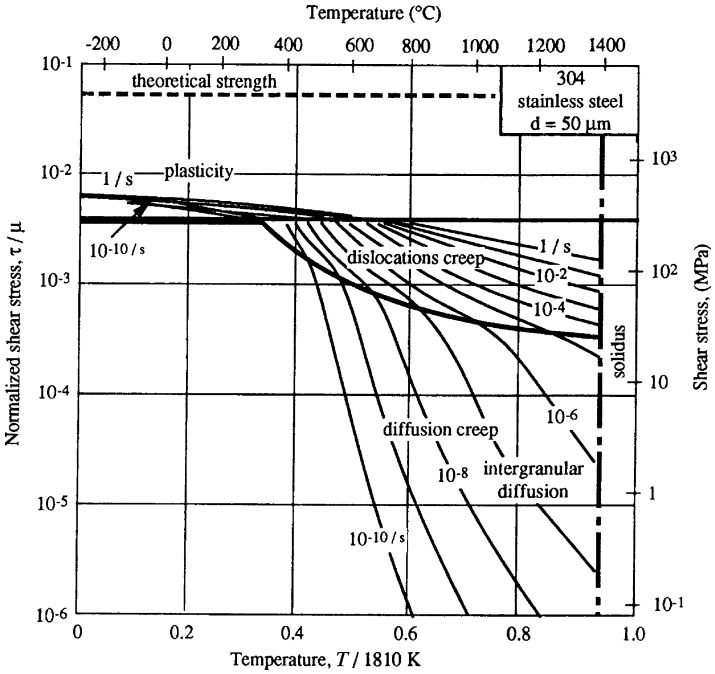


Fig. 4.38 Deformation map for austenitic stainless steel 304 (American standard) (Z05CND17-13, no Mo content), grain size 50 µm (Adapted from Frost and Ashby 1982)

resins; and it is seen in metals and metallic alloys (Al-Zn, Pb-Sn, Bi-Sn, titanium alloys, etc.) which have a very fine microstructure and are in general in multi-phase conditions when deformed under certain conditions of temperature and strain rate.

For a long time, super-plasticity remained a laboratory curiosity without any industrial applications. During the past 20 years, however, it has found practical uses, such as shaping certain alloys that are difficult to work (isothermal forming of some nickel-based superalloys manufactured by powder metallurgy), producing complex shapes (hemispheres in titanium alloys) or direct fabrication of assemblies (diffusion welding of aluminium- or titanium-alloy sheets, or forming hollow turbine blades with honeycomb reinforcement by super-plastic inflation.).

(b) *Phenomenological description*

We recall that for materials deformed at low temperature the stability of the deformation of a tensile test piece is controlled by the rate of work-hardening, characterised by the coefficient $n = \partial \log \sigma / \partial \log \varepsilon$. But in the region of the temperature-strain rate diagram where the material is super-plastic the value of this coefficient is very small, and stability results from the material great sensitivity to the strain rate. This can be characterised by another coefficient, $m = \partial \log \sigma / \partial \log \dot{\varepsilon}$. The variation of this coefficient with strain rate is displayed in Fig. 4.40 pertaining to a TA6V titanium alloy (Vairis 2008). It can be seen that it is small (<0.2) at high rates, and

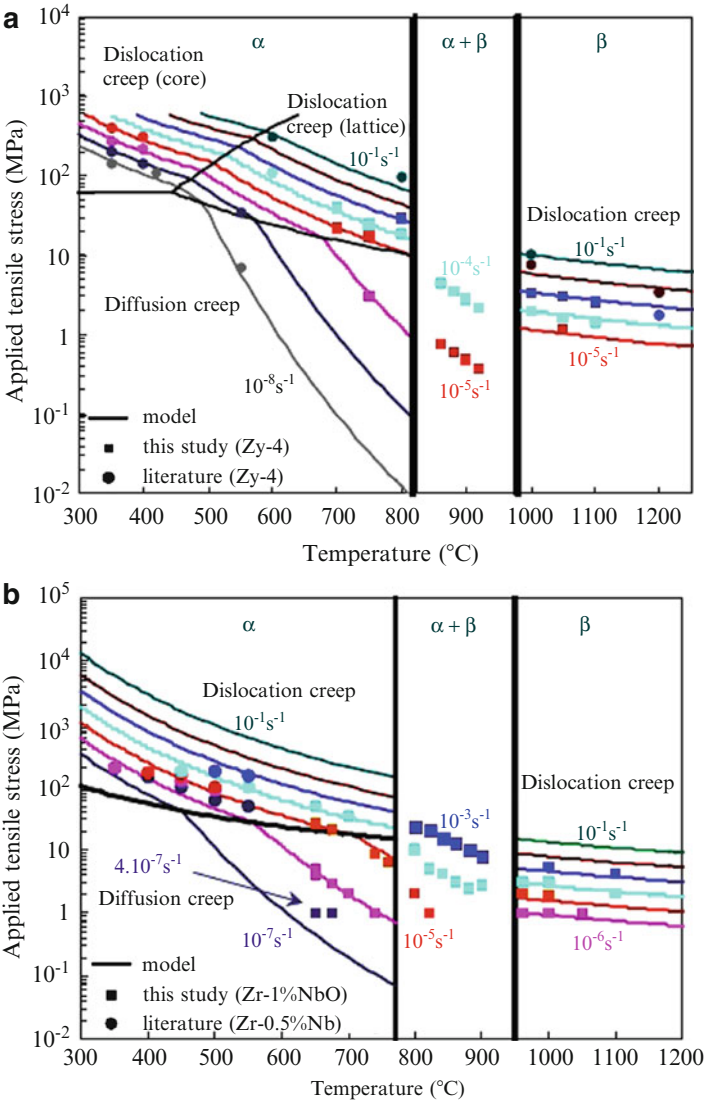


Fig. 4.39 Deformation mechanism maps for (**a**) Zy 4 and (**b**) Zy 1%NbO under uniaxial tension. Constant strain rate lines are plotted (in s^{-1}) for every decade (from Kaddour et al. 2004, 2011)

reaches its greatest values, about 0.5, at rates between 10^{-3} and 10^{-2}, where the behaviour is super-plastic. At lower strain rates, the exponent m would again have low values.

This sensitivity of super-plasticity to deformation rate can be explained qualitatively as follows. If the deformation becomes localised in a neck the rate increases locally, resulting in an increase in the stress. This increase is greater, the greater the

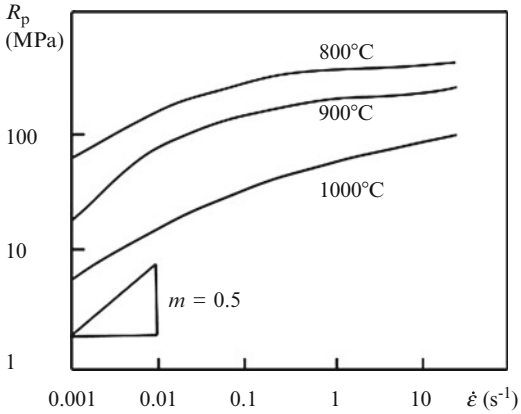


Fig. 4.40 Variation of the yield strength as a function of strain rate for a TA6V titanium alloy at various temperatures exhibiting a super-plastic region at strain rates between 10^{-3} and 10^{-2} (Adapted from Vairis 2008)

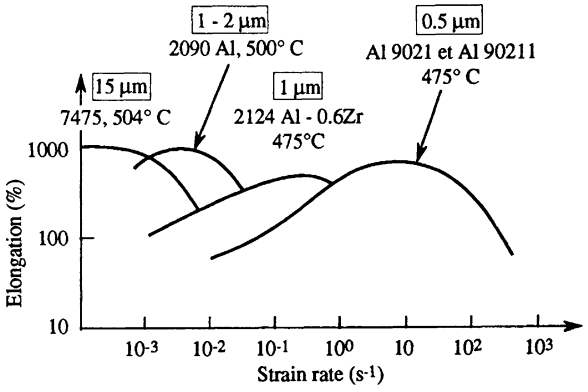


Fig. 4.41 Effect of grain size on strain rate for four alloys, showing that in alloys of the 9xxx series with very small grain sizes (<1 µm), super-plasticity can be achieved at relatively high rates (≈10 s^{-1}).

value of m; and its effect is to slow the development of the necking, resulting in the production of very diffuse necking and large elongations before rupture occurs.

(c) *Microstructure parameters and super-plasticity in metallic alloys.*
Super-plastic behaviour occurs only at high temperatures (>0.5 T_m), and usually at low strain rates (<10^{-3} s^{-1}). The activation energy that controls the deformation is small, in general less than the bulk diffusion energy and close to that corresponding to intergranular diffusion Q_{GB} (approx. ½Q_v) (see the tables of values at the end of this volume.)

The microstructure of a super-plastic material must fulfil certain conditions. Above all, the grain size must be very small. As Fig. 4.41 shows for a number of aluminium alloys, an increase in grain size shifts the super-plastic region very much towards small strain rates. Observations have shown that the variation of strain rate with grain size can be expressed as:

$$\dot{\varepsilon}^v = d^{-\beta} \tag{4.81}$$

with $\beta =$ between 2 and 3.

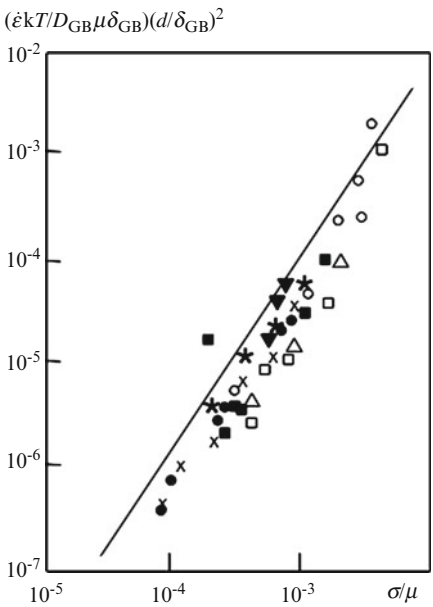


Fig. 4.42 Variation of the reduced strain rate with applied stress for the following super-plastic alloys: Pb 40%In (64°C), Δ: $d = 80$ μm; □: $d = 50$ μm; ○: $d = 200$ μm – Pb 2.5%Tl (21°C), ●: $d = 28$ μm; ▼: $d = 83$ μm – Sn 1%Bi (22°C), ■: $d = 5$ μm – Zn 22%Al (250°C), x: $d = 2.5$ μm – Pb 38%Sn (20°C), *: $d = 2.46$ μm (See also Baudelet and Suéry 1987)

Figure 4.42, from V.N. Perevezentsev et al. (1992), gathers together the effects of grain size, temperature, intergranular diffusion coefficient and deformation rate on applied stress for a number of materials. These results show that the rate-sensitivity coefficient m is close to 0.5 and that to a first approximation the rate is inversely proportional to the square of the grain size. Generally, to obtain such small grain sizes (<1 μm), which must remain stable during super-plastic deformation, alloys consisting of approximately equal volumes of two phases (eutectic or eutectoid alloys) must be used. Each phase will inhibit the growth of grain size in the other.

In some cases (high-resistance aluminium alloys) super-plastic microstructure is associated with a very special grain structure, characterised not only by small size but also by the presence of a precipitate-free zone around the boundaries, which is much softer than the material of the body of the grains. Using this, it has been found possible to develop super-plastic aluminium alloys by controlling the intergranular diffusion of gallium and forming a low-melting solid solution of Al-Ga in the neighbourhood of the boundaries.

After super-plastic deformation the grains remain equiaxial; sometimes their size increases slightly, and often there is no deformation at all in the bulk. Examination of the surface shows that the grains slide with respect to one another, and that this intergranular sliding can account for over half of the total deformation. This agrees with the data given in Fig. 4.30, which show that sliding at grain boundaries becomes important at low stress.

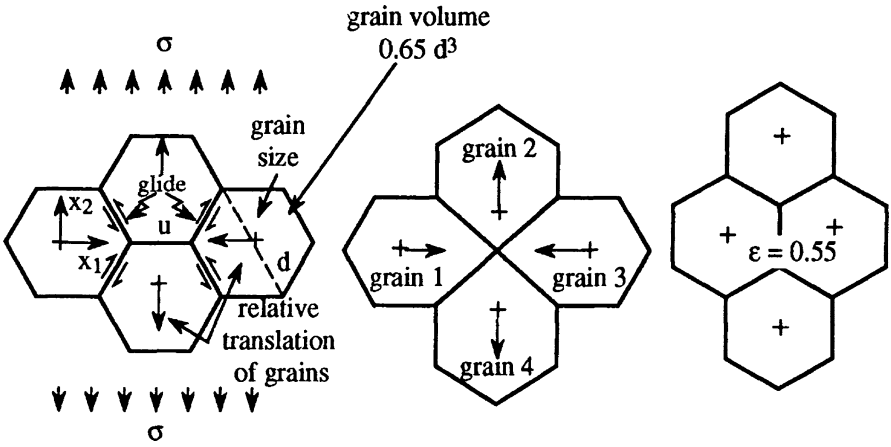


Fig. 4.43 Elementary mechanisms of deformation in a super-plastic polycrystalline material (Ashby and Verral 1973)

(d) *Physical models.*
Most models take as starting point the following observations:

- The value of the rate-sensitivity coefficient m increases with increasing rate, up to some critical rate which itself decreases with increasing grain size.
- Intergranular sliding plays a dominant role. If two grains slide with respect to each other along their common boundary, cavities should form at the triple points; but in fact such cavities appear only after some delay, or not at all. This suggests that grain boundary sliding must be accompanied by mechanisms such as migration of the boundaries, deformation of neighbouring grains (both of these to only a limited extent) or diffusion of matter either in bulk or along the boundaries.

Of the models that attribute super-plastic deformation to boundary sliding accommodated by diffusional creep in the grains the most elaborate is that due to Ashby and Verral (1973). These authors have shown how the deformation can occur without the need for any deformation of the grains but as a result of sliding along the boundaries only (see Fig. 4.43). Taking as the elementary event the change of shape associated with a set of four grains, they showed that the intermediate position for which the area of the boundary is maximum is a barrier. The grains slide over one another and this sliding is accommodated by diffusional creep, which brings about transport of material either through the volume (diffusion coefficient D_v) or along the boundaries (D_{GB}). Assuming that the grain boundaries are perfect sources and sinks of vacancies, they have established the following relation:

$$\dot{\varepsilon}^{\mathrm{v}} = 100\frac{\Omega}{\mathrm{k}Td^2}\left(\sigma - \frac{0.72\gamma_{\mathrm{GB}}}{d}\right)D_{\mathrm{v}}\left(1 + \frac{3.3\delta_{\mathrm{GB}}D_{\mathrm{GB}}}{dD_{\mathrm{v}}}\right) \tag{4.82}$$

where γ_{GB} is the energy of the grain boundaries.

As would be expected, this has the same form as Eq. 4.78. However, it involves a threshold stress in the term γ_{GB}/d, which could explain why in a logarithmic plot of stress against deformation rate there is often a change of slope in the region of small values of stress.

4.3.3 Creep in Ceramic Materials

4.3.3.1 Intergranular Glassy Phases in Ceramic Materials

Two specific features characterise the creep behaviour of ceramics compared to that of metals:

- creep strains in these refractory materials are much lower than those usually measured in metals. They are of the order of a few per thousand and are associated with extremely low strain rates;
- intergranular phases, in particular in technical (thermo-mechanical) ceramic materials, play a key role (Boussuge, 2009, Cannon and Langdon, 1983, 1988; Wilkinson, 1998).

Examinations of ceramic microstructures using the techniques of high-resolution TEM (see Chap. 1) have shown that these materials contain a thin intergranular (generally siliceous) glass phase. The thickness of the intergranular glass phase can be quite small, being ≈0.5 to ≈0.2 nm in some silicon nitride and zirconium ceramics and upward of ≈5 nm in some alumina-based materials (Clarke1987). In the majority of ceramic materials the intergranular phase is continuous throughout the microstructure and, significantly, located at the grain junctions.

The presence of a remnant intergranular glass phase in ceramics can be due to a number of different processes. In many ceramics, this phase results from the liquid-phase sintering process used during sintering (*e.g.*, Silicon nitride alloys, alumina substrate materials).

In others, such intergranular phases are present because the materials are prepared by the controlled but incomplete crystallisation of a glass (glass-ceramics). A third, but, practically important category, is that in which the glassy phase forms from the impurities present in the materials.

The intergranular phase wets the boundaries, meaning that it is energetically favourable for the phase to exist at the grain boundaries. The intergranular phase acts as a liquid phase and adopts an equilibrium thickness at high temperatures. This phase can sustain normal stresses because of their extremely small thicknesses.

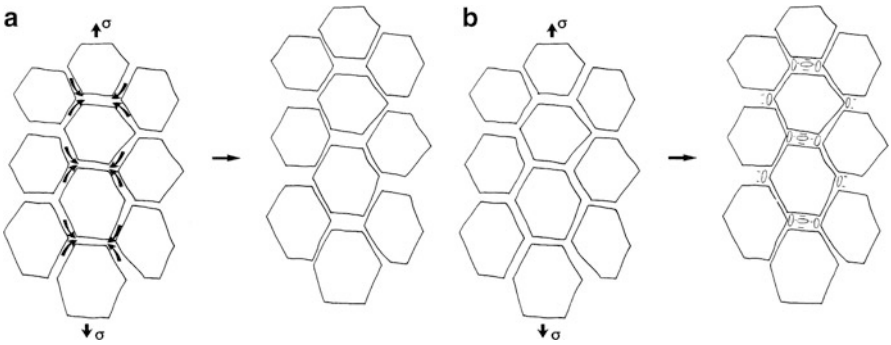


Fig. 4.44 Ceramic materials: creep mechanisms without any deformation of the grains; (**a**) plastic flow of the intergranular phase; (**b**) intergranular diffusion and damage of the intergranular phase

4.3.3.2 Creep Mechanisms in Ceramic Materials

Two main types of mechanisms can be invoked to explain the creep behaviour of ceramics:

– those operating within the grains (dislocation movement; transgranular diffusion), like in metals;
– those acting along the grain boundaries (Muto and Sakai1998).

The transgranular mechanisms operate in single crystalline ceramics and in a number of polycrystalline ceramics, (in particles, in oxides) under high compressive stresses. However these mechanisms are limited because of the covalent nature of the bindings in ceramics materials. This explains why intergranular mechanisms are prevalent. Several cases must be distinguished depending on the origin of the deformation and the transport phenomena:

- Grain boundary sliding with undistorted grains (Fig. 4.44). Several types of accommodation can be envisaged depending on the wetability and the viscosity of the intergranular phase:
  - Newtonian plastic flow of the intergranular phase (Fig. 4.44a). This mode of deformation is limited when the amount of intergranular phase is small, which is usually the case in most thermo-mechanical ceramics.
  - Nucleation of cavities or cracks at grain boundaries (Fig. 4.44b). This form of damage occurs significantly in technical ceramics which contain an intergranular phase which is poorly or not recrystallised.
- Grain boundary sliding accompanied with grain deformation (Fig. 4.45). Three main types of deformation mechanisms can operate:

– Newtonian plastic flow from the grain boundaries under compression to the grain boundaries under tension, with dissolution – re-precipitation mechanisms

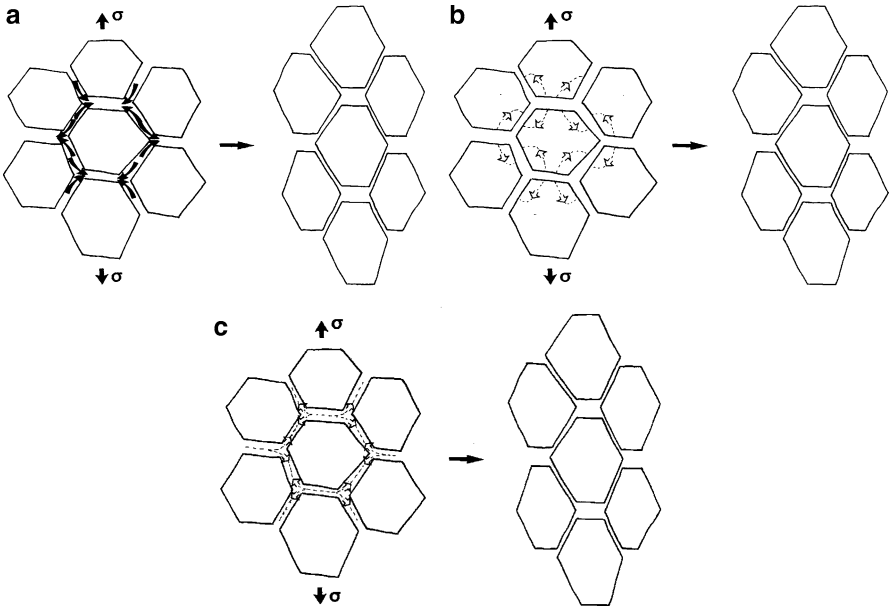


Fig. 4.45 Ceramic materials: creep mechanisms involving deformation of the grains; (**a**) viscoplastic flow; (**b**) transgranular diffusion; (**c**) intergranular diffusion

(Fig. 4.45a). This mechanism is facilitated in the presence of intergranular phases of low viscosity and reacting with the grains,

– diffusional flow, as in metals, according to the Herring-Nabarro mechanism (Fig. 4.45b),
– diffusion flow along the grain boundaries, according to the Coble creep mechanism (Fig. 4.45c).

One example of a deformation mechanism map for α alumina (Al_2O_3) is given in Fig. 4.46a, b (Frost and Ashby 1982). These figures clearly show that, at elevated temperatures (>1000°C) creep deformation is dominated by boundary diffusional flow. In these conditions the strain rate varies linearly with the applied stress (Fig. 4.46b).

4.3.4 Creep in Polymers

Creep deformation has been much less investigated in polymers than in metals and ceramics. In polymeric materials it is not always possible to distinguish the viscoelastic behaviour (see Chap. 5 of this volume) from the creep behaviour. Creep strain, by definition, is not recovered after unloading contrarily to viscoelastic strain. There are still very few detailed studies devoted to the description and the

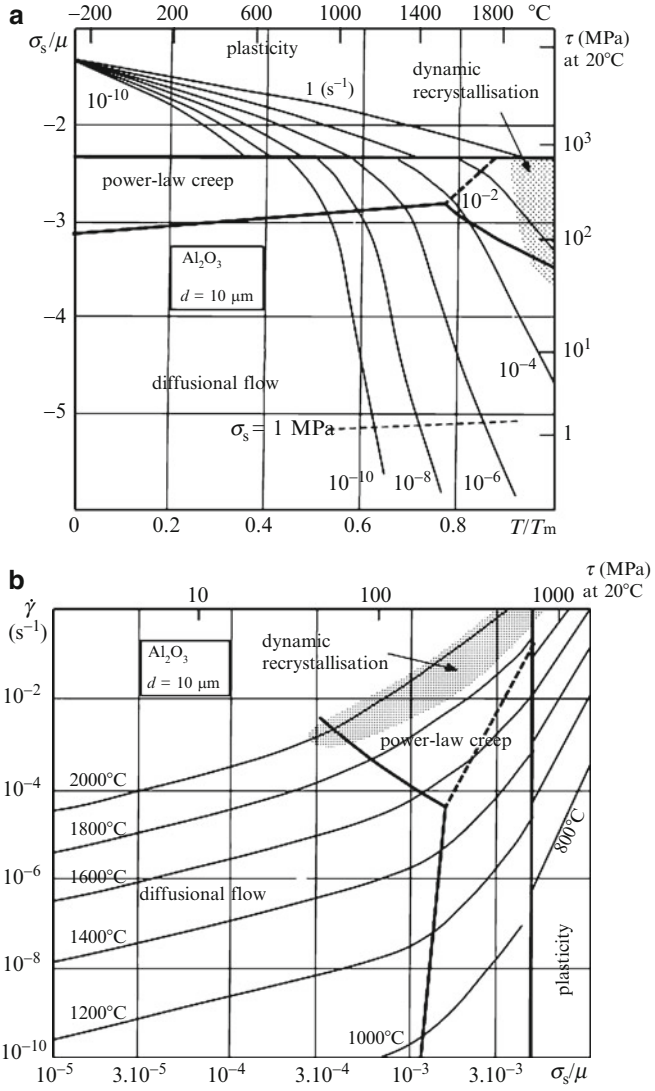


Fig. 4.46 (**a**) stress/temperature map for Al_2O_3 with a grain size of 10 μm (**b**) stress-rate/stress map for Al_2O_3 with a grain size of 10 μm (data labelled with temperature in °C) (Adapted from Frost and Ashby 1982)

understanding of the creep behaviour of polymers. However, recently, a number of studies on semi-crystalline polymers, such as Polyethylene (PE) and Polyamide (PA6) materials have been reported (Regrain et al. 2009; Ben Hadj Hamouda et al. 2007).

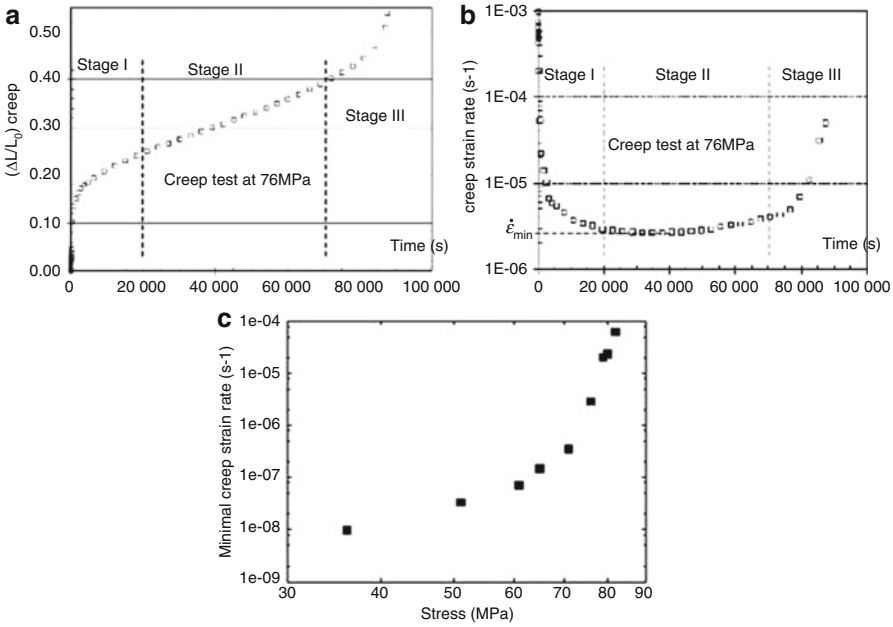


Fig. 4.47 Polyamide 6 (PA6) polymer. Creep behaviour: (**a**) creep strain versus time at an applied stress of 76 MPa; (**b**) creep strain rate versus time; (**c**) creep strain rate as a function of applied stress (Regrain et al. 2009)

In these semi-crystalline polymers, all chains have not the same mobility. In the crystalline phase, chains are well ordered and tied by both covalent and hydrogen bonds. These chains have thus few degrees of freedom. In the amorphous phase, chains are not organised and can slip past each other. Chains in the amorphous phase are much more mobile than those in the crystallites. Under an applied load, both kinds of chains do not have the same behaviour. To take the contrasting behaviour of each phase into account, it is necessary to have in mind that the stress actually experienced by each phase can be different.

The typical behaviour of polyamide 6 (PA6) submitted to a tensile stress of 76 MPa at 25°C and 50% humidity is reproduced in Fig. 4.47. This stress is lower than the plateau stress (∼85–90 MPa) observed during a tensile test carried out at different strain rates (2.5×10^{-3} s^{-1} – 2.6×10^{-2} s^{-1}). Figure 4.47a shows the three typical stages observed in metallic alloys, with the existence of a minimum "stationary" creep strain rate (Fig. 4.47b) (Regrain et al. 2009). Tests performed at different applied nominal stresses showed that the minimum strain rate cannot be described as a single power law of the applied stress. The examination of the results reported in Fig. 4.47c indicates that at low applied stress (≤70 MPa), the slope of the $\dot{\varepsilon}^{v} - \sigma$ curve is close to 4.20, while at higher stresses, the slope is much higher, of the order or 37.5.

This variation of the stress exponent of strain rate with the applied stress is not presently understood in terms of the elementary deformation modes of the two

phases present in semi-crystalline polymers. A constitutive multiphase model based on the localisation of the applied stress in each phase and the homogenisation of the local creep strains has been developed to describe the behaviour of PA6 material. This model takes into account the degree of crystallinity, which can evolve during deformation.

Tertiary creep is associated with the formation of a neck. A recent study has shown that tertiary creep in pre-notched specimens does not evolve continuously with time. A decrease of the elongation rate is observed, denoting a rheo-hardening effect of the material (PA6), ended by final rupture (Saï et al. 2011).

4.4 Mechanical Models of Macroscopic Viscoplasticity

The rheological formulation of viscoplastic behaviour can be modelled on that of plastic behaviour, using the same basic elements. As in the case of plasticity, in the present state of knowledge and of the computational methods available it would be unreasonable to perform a complete shift from the microscopic scale to that of the representative volume element (RVE), because of the many intermediate stages and the ensuing heterogeneities. In general therefore we cannot hope that models developed on the scale of dislocations will give quantitative information concerning materials having a complex microstructure: their role is rather to enlighten the user on microscopic-scale mechanisms and to define the limits within which macroscopic mechanical models will be valid. It is these latter that will be used as a last resort in structural computations; in general they rest on a broad *phenomenological* base, but the structure and critical variables of the most powerful of them is suggested by the local deformation mechanisms. Thus it is a question again of determining:

- the form of the elastic domain at the instant considered,
- the expression for viscoplastic flow,
- the evolution of the hardening variables.

The elastic domain is defined with the aid of a scalar-valued *load function* f. This depends on the stress tensor and the work-hardening variables in such a way that it provides a means for representing the initial form of this region and its possible changes during loading. The viscoplastic flow is defined by reference to this function in the classic case of associated viscoplasticity; and similarly for the work-hardening if the model is *standard*; a new element enters when the mechanisms of recovery over time or of ageing have to be taken into account. The results obtained in this chapter relate only to *isothermal conditions*.

4.4.1 Viscoplastic Potential for a Single Crystal

Our study of the microscopic mechanisms has shown that viscoplastic flow can be a consequence of slip or of diffusion. The second mechanism, which implies transport

of material, does not come within the framework of classical models. The treatment of deformation by slip that we shall now give will be limited to displacements in the plane without any consideration of the mechanisms of climb, which involve stress components normal to the slip plane and therefore violate the assumption of normality. The success of models of this type rests on the fact that although slip is not always the dominant mechanism it is often responsible for the major part of the deformation. This is certainly true at low temperatures, but for dislocation creep it holds also at high temperatures to the extent that the climb process displaces dislocations only over short distances compared to the "long" distances covered by the dislocations.

Under the assumption that the material obeys the Schmid law, the rate of slip $\dot{\gamma}_r$ on a given system (r) will depend only on the resolved shear $\tau_r = \underline{m}^{(r)}.\underline{\underline{\sigma}}.\underline{n}^{(r)}$, the work-hardening variables, here represented conventionally by a vector of components Y_m, and possibly the temperature T. This allows us to define the rate of viscoplastic deformation, in the case when several slip systems are active, by a function Φ, non-decreasing with τ_r:

$$\dot{\gamma}_r = \Phi\left(\tau_r, Y_m, T\right) \quad \underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} = \sum_r \dot{\gamma}_r (\underline{m} \otimes \underline{n})^{\mathrm{S}_r} \tag{4.83}$$

Applying a virtual variation $\mathrm{d}\underline{\underline{\sigma}}$ to the existing stress state represented by the tensor $\underline{\underline{\sigma}}$, we have:

$$\underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} : \mathrm{d}\underline{\underline{\sigma}} = \sum_r \Phi\left(\tau_r, Y_m, T\right) \mathrm{d}\tau_r \tag{4.84}$$

The viscoplastic potential is then defined as a scalar-valued function of $\underline{\underline{\sigma}}$,$Y_m$ and T by:

$$\Omega\left(\underline{\underline{\sigma}}, Y_m, T\right) = \sum_r \int \Phi\left(\tau_r, Y_m, T\right) \mathrm{d}\tau_r \tag{4.85}$$

This function enables us to calculate the macroscopic viscoplastic strain rate corresponding to any given state of stress and hardening; the rate vector is orthogonal to the equipotential surface through the point corresponding to that state:

$$\underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} = \partial\Omega \Big/ \partial\underline{\underline{\sigma}} \tag{4.86}$$

While no system is active, Ω remains constant. Without loss of generality it can be given the value zero inside and on the boundary of the elastic domain, and the surfaces $\Omega =$ constant are located inside one another and are convex. It is easy to show that Ω is a convex function of $\underline{\underline{\sigma}}$; the proof is based on an inequality relating to the slip systems, into which the stress and viscoplastic strain tensors are substituted, thus:

$$\left[\Phi\left(\tau_r, Y_m, T\right) - \Phi\left({\tau_r}^0, Y_m, T\right)\right]\left(\tau_r - {\tau_r}^0\right) \geq 0 \tag{4.87}$$

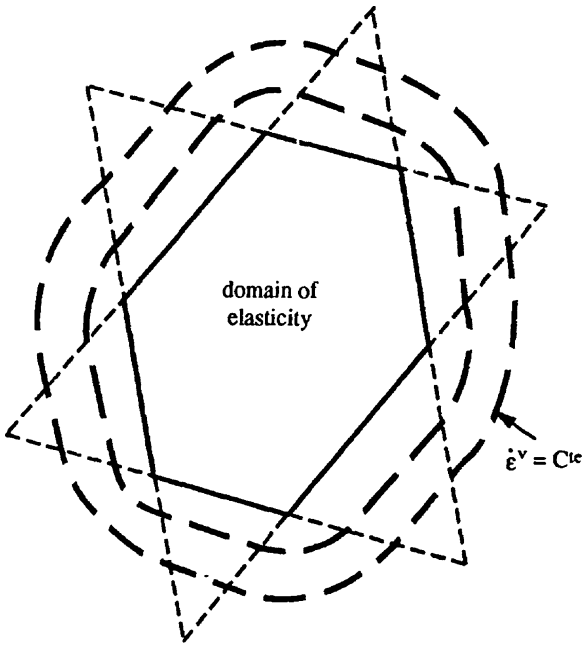


Fig. 4.48 Viscoplastic equi-potentials for a single crystal

(expressing the condition that the dissipated energy must be positive), giving:

$$\left[\frac{\partial \Omega}{\partial \underline{\underline{\sigma}}}\left(\underline{\underline{\sigma}}\right) - \frac{\partial \Omega}{\partial \underline{\underline{\sigma}}}\left(\underline{\underline{\sigma}}^0\right)\right] : \left(\underline{\underline{\sigma}} - \underline{\underline{\sigma}}^0\right) \geq 0 \tag{4.88}$$

from which, between two points $\underline{\underline{\sigma}}^0, \underline{\underline{\sigma}}^1$

$$\Omega\left(\underline{\underline{\sigma}}^1\right) - \Omega\left(\underline{\underline{\sigma}}^0\right) \geq \left(\underline{\underline{\sigma}}^1 - \underline{\underline{\sigma}}^0\right) : \frac{\partial \Omega}{\partial \underline{\underline{\sigma}}}\left(\underline{\underline{\sigma}}^0\right) \tag{4.89}$$

The convexity is not strict, since the resolved shear depends only on the deviatoric part of the stress tensor; but if Φ is a strictly increasing function of τ_r, Ω is a strictly convex function of the deviator of $\underline{\underline{\sigma}}$. On the other hand, it is found that an equi-potential surface is defined piecewise. If only one system is active the surface consists of hyper-planes, since, as for the elastic domain in the case of time-independent plasticity, the constancy of Ω implies that of τ_r also. Between two adjacent hyper-planes (Fig. 4.48), for example if two systems r and s are active, it is defined, for given temperature and work-hardening state, by:

$$\int \Phi\left(\tau_r, Y_m, T\right) \mathrm{d}\tau_r + \int \Phi\left(\tau_s, Y_m, T\right) \mathrm{d}\tau_s = \text{constant} \tag{4.90}$$

4.4.2 Viscoplastic Potential for a Polycrystal

Treating a polycrystal as an assembly of single crystals gives a "macroscopic" viscoplastic potential, which is the sum of the single-crystal equipotentials for the separate grains, weighted by their volume fraction. Thus for N grains:

$$\Omega\left(\underline{\underline{\sigma}}, {Y_m}^1, \ldots, {Y_m}^N, \underline{\underline{\sigma}}^1_{\text{loc}}, \ldots, \underline{\underline{\sigma}}^N_{\text{loc}}, T\right) = \sum_i^N \frac{V_i}{V}\Omega_i \tag{4.91}$$

where each potential depends on the local stress $\underline{\underline{\sigma}}^i_{\text{loc}}$ and the work-hardening variables for the individual grain Y^i_m, as well as on the mean stress on the aggregate. This shows that the convexity properties continue to hold and that equi-potentials can again be defined by their normals to the direction of viscoplastic flow.

The phenomenological approach takes this route but reduces the number of variables used to define the state of the material: the local stresses are not evaluated explicitly and for the work-hardening only global variables relating to the "average" material are taken. The previous formalism is retained and, following once again the example of plasticity, generalised standard viscoplastic materials are defined, using the convex potentials for each of the variables $\underline{\underline{\sigma}}$ and Y_m and putting:

$$\dot{\varepsilon}^{\text{v}} = \frac{\partial\Omega}{\partial\underline{\underline{\sigma}}} \qquad \dot{\alpha}_m = \frac{\partial\Omega}{\partial Y_m} \tag{4.92}$$

For all models of this type the value of the mechanical dissipation for any given state of stress and work-hardening exceed that of Ω; this ensures that it is positive, and therefore that the second law of thermodynamics is not violated.

If recovery is to be taken into account this requires a rather different viscoplastic potential, with two parts. The first, Ω_{v}, corresponds to what we have just developed; the second, Ω_{r}, to the effects of recovery on the internal variables, and can exist even if the representative point for the stress lies inside the elastic domain. In general, this second potential depends only on the internal variables, thus:

$$\Omega = \Omega_{\text{v}}\left(\underline{\underline{\sigma}}, Y_m, T\right) + \Omega_{\text{r}}\left(Y_m, T\right) \tag{4.93}$$

4.4.3 Time-Independent Plasticity and Viscoplasticity Compared

Time-independent plasticity can be studied by starting with a viscoplasticity law and assuming that the equipotentials defined by the function Ω are very closely packed. The inelastic strain rate will then vary very quickly with the stress, and this will tend to give instantaneous inelastic deformations. By going to a limit, Ω can be replaced

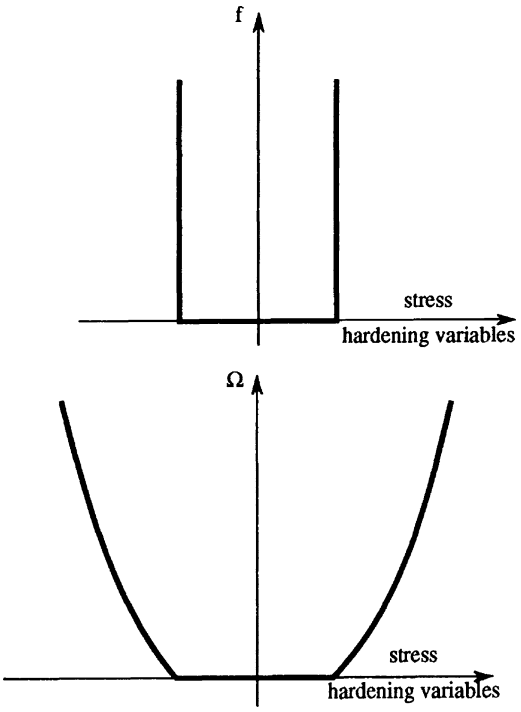


Fig. 4.49 Plastic and viscoplastic potentials; (**a**) plastic: f is the indicator function for the elastic domain; (**b**) viscoplastic: Ω is the viscoplastic potential

by f, the *indicator function* for the elastic domain, represented by a cylinder; this is shown in Fig. 4.49a, for comparison with the viscoplastic potential in Fig. 4.49b. In this case it is clear that whilst the direction of flow is still specified, the same is not true for the magnitude, all the sections being projected onto the surface $f = 0$. Thus the deformation rate is specified to within a scalar factor, the plastic factor, whose value is determined by the condition that the working point must remain on the surface $f = 0$:

$$\underline{\underline{\dot{\varepsilon}}}^{p} = \dot{\lambda} \frac{\partial f}{\partial \underline{\underline{\sigma}}} \tag{4.94}$$

Thus the essential difference between plastic and viscoplastic theory is that the latter allows freedom of choice for the rate of inelastic deformation, whilst in the former a further equation has to be satisfied, with the result that this rate is determined once the shape of the elastic region has been chosen.

4.4.4 Specific Constitutive Equations

The study of physical models shows that viscoplastic behaviour occurs in various conditions, which explain why these models can have very different forms. They are

all characterised by an expression that gives the inelastic strain rate as a function of the work-hardening and the current stress; this relation forbids instantaneous plastic strain and separates the strain into an elastic and an inelastic part. However, by passage to a limit, models can be constructed for which the strain rate is practically independent of time, and also rigid-viscoplastic models. In the following paragraphs we first give models that do not involve a threshold and then models that do, finishing with some comments on the use of viscoplastic models for numerical work.

4.4.4.1 Laws Not Involving Work-Hardening

Norton's law (4.1c) is obtained by choosing for the viscoplastic potential a power law in von Mises equivalent stress, with two coefficients K, M that characterise the material:

$$\Omega = \frac{K}{M+1} < \frac{\bar{\sigma}}{K} >^{M+1} \tag{4.95}$$

where $< H > = \max(0,H)$ and

$$\bar{\sigma} = \left[(3/2)\, \underline{\underline{s}} : \underline{\underline{s}} \right]^{1/2} \tag{4.96}$$

The viscoplastic flow is then given by:

$$\underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} = \frac{\partial \Omega}{\partial \underline{\underline{\sigma}}} = \frac{3}{2} \left(\frac{\bar{\sigma}}{K} \right)^{M} \frac{\underline{\underline{s}}}{\bar{\sigma}} \tag{4.97}$$

Under uniaxial loading, when the only non-zero component of the stress tensor is $\sigma_{11} = \sigma$, this simplifies to:

$$\begin{aligned} \dot{\varepsilon}^{\mathrm{v}}_{11} &= (\bar{\sigma}/K)^{M} \\ \dot{\varepsilon}^{\mathrm{v}}_{22} &= \dot{\varepsilon}^{\mathrm{v}}_{33} = -(1/2)\, \dot{\varepsilon}^{\mathrm{v}}_{11} \end{aligned} \tag{4.98}$$

There is no work-hardening and the rate of viscoplastic deformation is determined fully by the secondary creep. M is the *Norton exponent* of (4.1c), equal to the slope of the line in the log-log plot of strain rate against stress, as in Fig. 4.3.

Other threshold-free laws can be envisaged, in particular when the power law is a poor approximation at high stresses. Provided that the equipotentials are given in terms of a von Mises equivalent stress, all these laws give an expression of the type:

$$\underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} = \frac{3}{2} \dot{\bar{\varepsilon}}^{\mathrm{v}} \frac{\underline{\underline{s}}}{\bar{\sigma}} \tag{4.99}$$

where $\dot{\bar{\varepsilon}}^{\mathrm{v}} = \left[(2/3)\,\underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} : \underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}}\right]^{1/2}$ is the absolute rate of viscoplastic strain rate under uniaxial loading. The most currently used formulae involve products of exponentials and powers, sometimes a hyperbolic sine. Temperature may enter explicitly through activation terms in exp $(-Q/\mathrm{k}T)$:

$$\dot{\bar{\varepsilon}}^{\mathrm{v}} = \dot{\varepsilon}_0(\bar{\sigma}/K)^q \exp\left[\mathrm{F}(\bar{\sigma}, T)\right] \tag{4.100}$$

Thus the Sellars-Teggart (1966) law is:

$$\dot{\bar{\varepsilon}}^{\mathrm{v}} = \dot{\varepsilon}_0(\sinh K\bar{\sigma})^q \exp\left(-Q/\mathrm{k}T\right) \tag{4.101}$$

4.4.4.2 Laws with Multiplicative Work-Hardening

All the preceding laws state that under any non-zero constant loading the flow rate of the material will be non-zero and constant: they reduce the elastic domain to a single point. Hardening of the material can be represented while still retaining this property by manipulating the spacing of the equi-potentials and thus introducing a work-hardening variable into the expression for the deformation rate. Although not mandatory, it has become standard to use for this the cumulative viscoplastic deformation v, defined as:

$$v(t) = \int_0^t \dot{\bar{\varepsilon}}^{\mathrm{v}}(\tau)\,\mathrm{d}\tau \tag{4.102}$$

In expressions of the type of (4.95) the constant K, which has the dimensions of a stress, now depends on v and is called *drag stress*, expressing the resistance to deformation due to the rate term. It has become standard also to represent the hardening by a power law in v, $K = K_0 v^n$; this leads to a new expression for v, which can be inverted to give $\bar{\sigma}$:

$$\dot{v} = <\bar{\sigma}/K_0>^{1/m} v^{-n/m}$$
$$\bar{\sigma} = K_0 \dot{v}^m v^n \tag{4.103a, b}$$

with $m, n > 0$. Equation 4.103a gives the strain-hardening rule and shows that the strain rate is infinite at the start of the loading and tends to zero as the strain increases. For the case of creep in simple tension under a uniaxial stress σ_0 the relation can be integrated to give:

$$\varepsilon^{\mathrm{v}} = \left[\frac{n+m}{m}\left(\frac{\sigma_0}{K_0}\right)^{1/m} t\right]^{\frac{m}{m+n}} \tag{4.104}$$

4.4.4.3 Laws with Isotropic Additive Work-Hardening

There are several ways in which a threshold might be introduced:

– through the existence of an elastic domain of constant size throughout the course of the deformation. There is no work-hardening and the relation is of the type of (4.6), in which the critical stress remains always equal to the initial elastic limit σ_0:

$$\dot{\nu} = \left(\frac{\bar{\sigma} - \bar{\sigma}_0}{K} \right)^M \tag{4.105}$$

– through the existence of an elastic region whose size increases in the course of the deformation. As in plasticity, there is isotropic additive work-hardening and the work-hardening variable depends on the cumulative viscoplastic deformation, for example exponentially:

$$\dot{\nu} = \left(\frac{\sigma - R - \sigma_0}{K} \right)^M$$

$$R = Q\,[1 - \exp(-b\nu)] \tag{4.106–4.107}$$

R and Q have the dimensions of stress. R represents a friction stress, whose rheological image is a friction element; it corresponds to the applied stress necessary to produce inelastic deformation. To the extent that it is independent of the direction of the load its existence is normally associated, on the microscopic scale, with local obstacles to the movement of the dislocations, such as interactions between dislocations and particles or between dislocations and precipitates.

4.4.4.4 Laws with Kinematic Additive Work-Hardening

Kinematic work-hardening displaces the centre of the elastic domain in the stress space, corresponding to a change of macroscopic "neutral state" for the material. As we discussed in connexion with physical models (Sect. 4.3.2), this is related to heterogeneities on the microscopic scale. These give rise to non-uniform local stress fields over "long" distances (long range stresses), so that the shear force on the dislocations can be determined from the local "effective" stress, that is, the difference between the applied stress and the "back-stress", which exerts a counteracting force tending to produce a deformation in the sense opposite to that of the load. Under a uniaxial load, with a constant threshold and a back-stress σ_i, the critical stress of (4.6) is:

$$\sigma_s = \sigma_i + \sigma_0 \text{ (simple tension)} \tag{4.108a}$$

$$\sigma_s = \sigma_i - \sigma_0 \text{ (simple compression)} \tag{4.108b}$$

In both metallic alloys and composite materials there are many sources of heterogeneities: phases, grains, precipitates, dislocation cells, ... and to each of these corresponds a procedure for localising the stresses and the deformations. These redistributions are very complex and are still not fully characterised, but in general they are all governed by local and macroscopic deformations. Thus the corresponding phenomenological laws also are governed by the deformation; the simplest law for work-hardening, that due to Prager, is:

$$\underline{\underline{X}} = \frac{2}{3} C \underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} \tag{4.109a}$$

which gives an elastic domain and a viscoplastic potential depending on both $\underline{\underline{\sigma}}$ and $\underline{\underline{X}}$:

$$\mathrm{f} = \overline{(\underline{\underline{\sigma}} - \underline{\underline{X}})} - \sigma_0 \tag{4.109b}$$

In viscoplasticity as in plasticity more elaborate work-hardening laws can of course be considered. For nonlinear kinematic work-hardening in particular, a fading memory term can be included in the expression for strain rate, whose direction can be given either by the present value of the back-stress or by the instantaneous viscoplastic strain rate, thus:

$$\underline{\underline{\dot{X}}} = \frac{2}{3} C \underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} - \gamma \underline{\underline{X}} \dot{v}$$

$$\underline{\underline{\dot{X}}} = \frac{2}{3} C \underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} - \gamma \left(\underline{\underline{X}} : \frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} \right) \underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} \tag{4.110a, b}$$

Equations 4.109 and 4.110a were discussed in detail, in both uniaxial and multiaxial formulation, in Sect. 3.5.2.4. Equation 4.110b differs from (4.110a) only in multiaxial loading, when it has different properties concerning the ratchetting effect.

4.4.4.5 Laws with Recovery

Recovery, which results in a softening of the material, is often considered as a decrease of internal stresses; it is thus represented by the addition to the viscoplastic potential of a term originating in a recovery potential Ω_{r}, as in (4.93), modifying the definition of the variable $\underline{\underline{X}}$ to:

$$\underline{\underline{\dot{X}}} = \frac{2}{3} C \underline{\underline{\dot{\varepsilon}}}^{\mathrm{v}} - \left(\frac{Q X_{\mathrm{II}}}{L} \right)^{l} \frac{\underline{\underline{X}}}{X_{\mathrm{II}}} \tag{4.111}$$

where $X_{\mathrm{II}} = \left(\frac{2}{3} \underline{\underline{X}} : \underline{\underline{X}} \right)^{1/2}$.

If l is given the value 1 this corresponds exactly to the approach given in Sect. 4.3.2.1 for dislocation creep.

Other forms of recovery can be envisaged, in particular those bearing on the isotropic variable in the case of additive work-hardening.

4.4.4.6 Using the Models

The work-hardening laws we have discussed should be understood as constituting a hierarchical model, in which one can take more or fewer of the elements according to the material, the temperature range and the particular application (for example, whether monotonic or cyclic loading). It should never be forgotten that the more complex the model the more tests will be needed to specify it. A fundamental requirement for characterising viscous stress is a set of test results that covers a wide range of deformation rates. Further, it is not always certain that the available models give a correct interpretation of the results of both creep tests and relaxation tests at the same time, showing that even with monotonic loading and steadily increasing viscoplastic deformation, viscosity will make the mechanical responses more complex.

Thus there are difficulties in specifying the coefficients that represent the properties of real materials. Further, a viscoplastic formulation is often used in preference to a plastic formulation for reasons of computational simplicity. It has been shown that the viscoplastic formulation "regularises" the plasticity, leading to more regular numerical solutions: two classical examples concern rigid-viscoplastic models and threshold-free models with power functions in which the exponents can be very large.

Rigid-viscoplastic models may or may not include a threshold term; they can express for example the model of (4.103b) in terms of total deformation, assuming that this takes place without change of volume. At any instant $\dot{v}$ is given by: $[(2/3)\,\underline{\underline{\dot{\varepsilon}}}(\tau) : \underline{\underline{\dot{\varepsilon}}}(\tau)]^{1/2}$, with $\mathrm{Tr}\,(\underline{\underline{\dot{\varepsilon}}}) = 0$, and v is the integral of this with respect to time. Given the data for the present state, represented by v and the deformation rate, (4.103b) enables the stress state to be determined. This formulation is particularly well adapted to finite-element calculations ($\underline{\underline{\dot{\varepsilon}}}$ given) in cases when the elastic deformation is negligible, such as metal forming. *Bingham's law* for viscous fluids can be obtained by the same procedure, starting from (4.105) and treating this as a limiting case of viscoplasticity.

The viscoplastic formulation leads to calculations in which the behaviour is practically independent of the deformation rate. The standard method given in literature is to choose a Norton-type law, or multiplicative work-hardening, with a power law with a very high exponent, of the order of 200. This is the current approach for single crystals of Ni-based superalloys in particular. A numerical

cut-off is often associated with the method, with the reintroduction of an elastic region, the approach by infinitesimal deformation rates not affecting the results:

$$\dot{\gamma} = 0 \text{ if } \tau \leq \tau_c$$
$$\dot{\gamma} = (\tau / \tau_c)^M \text{ otherwise} \qquad (4.112a, b)$$

τ_c is a work-hardening variable, a function of the slip in the system to which it relates, and possibly of slips in other systems also.

4.4.5 *Simultaneous Treatment of Plasticity and Viscoplasticity*

In this section we have often introduced plasticity with the aim of either comparing or contrasting the plastic formulation to the viscoplastic one: the first leads to instantaneous and the second to delayed responses. A viscoplastic model, referred to as "unified modelling", can treat both types of responses. However, with some materials it can be useful to maintain the distinction between the two regimes, one of fast deformation and the other of slow. This is especially the case for:

- materials with considerable capacities for both work-hardening and creep, such as austenitic stainless steels, and which are inversely sensitive to the loading rate load in certain temperature ranges. Here "inverse" means that the stress necessary for inelastic flow decreases as the loading rate increases;
- materials in which instantaneous deformation (plasticity) and creep (viscoplasticity) occur in distinct stress regions; for example, nickel-based alloys such as Inco 78 (18Cr, 18Fe, 3Mo, 1Ti, 5.2Nb, Bal. Ni).

Thus in the unified models two thresholds, plastic and viscoplastic respectively, co-exist and the total deformation is the sum of three parts: elastic, plastic and viscoplastic:

$$\underline{\underline{\varepsilon}} = \underline{\underline{\varepsilon}}^{\mathrm{e}} + \underline{\underline{\varepsilon}}^{\mathrm{p}} + \underline{\underline{\varepsilon}}^{\mathrm{v}} \qquad (4.113a)$$

$$\dot{\underline{\underline{\varepsilon}}}^{\mathrm{p}} = \dot{\lambda} \frac{\partial \mathrm{f}}{\partial \underline{\underline{\sigma}}} \qquad (4.113b)$$

$$\dot{\underline{\underline{\varepsilon}}}^{\mathrm{v}} = \frac{\partial \Omega}{\partial \underline{\underline{\sigma}}} \qquad (4.113c)$$

The models most in current use are limited to a Norton law for viscoplasticity and linear isotropic or kinematic work-hardening for plasticity. Experience has shown, however, that some coupling between the plastic and viscoplastic laws must be introduced. The resulting models, which are outside the scope of this chapter, enable plasticity and viscoplasticity to be identified in the results of basic tests, which must

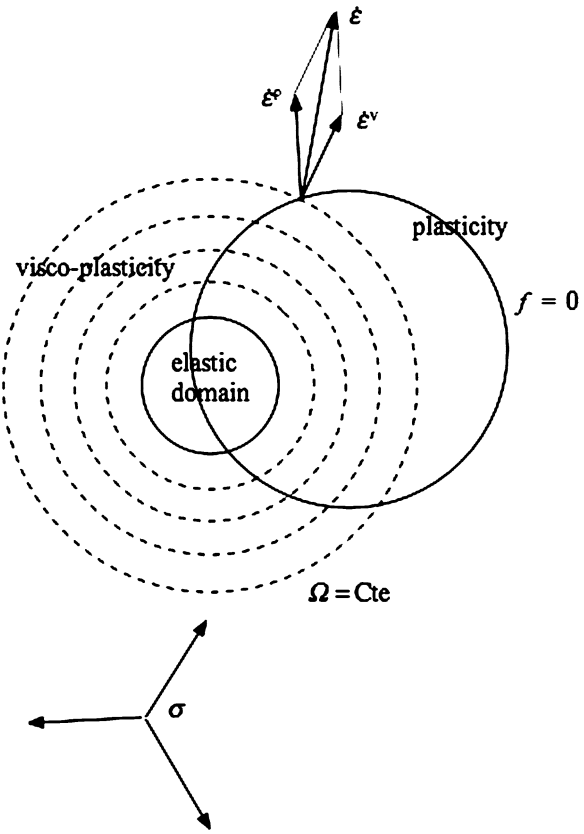


Fig. 4.50 Construction of inelastic strain rate for a model covering plasticity and viscoplasticity

then be supplemented by tests that show the plasticity-creep coupling. Figure 4.50 shows the construction of the inelastic deformation for two models with kinematic and isotropic work-hardening in plasticity and viscoplasticity.

Detailed description of macroscopic viscoplastic constitutive equations and of viscoplastic structural design can be found in text-books (see *e.g.* Lemaitre and Chaboche 1990; Besson et al. 2009).

4.5 Methods for Reinforcing Against Creep

We end the chapter with a few indications that may be useful in the development of materials that must resist creep under loading at high temperatures.

It will be seen from what we have shown that two approaches are possible:

– reducing the diffusional transport of material,
– blocking the slip or climb movements of the dislocations.

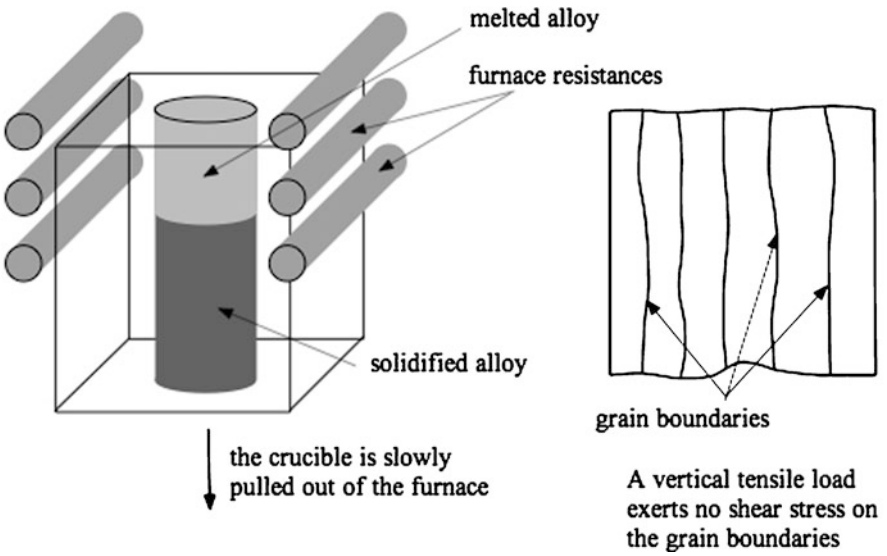


Fig. 4.51 Directional solidification of turbine blades, giving elongated grains or even single crystals (Adapted from Ashby and Jones 1980)

We study these approaches, considering first methods concerning diffusion and then two mechanisms for blocking dislocations, one by solid solution and the other by precipitation.

4.5.1 Reinforcement by Reducing Diffusion

This is the appropriate choice of method when working with materials of high melting point, such as refractory metals (Ta, Mo, W, Nb) or ceramics (Al_2O_3, ZrO_2, Si_3N_4, etc.). Such materials, unfortunately, are brittle and difficult to work.

As we have seen, important contributors to creep resistance at high temperature are grain size and intergranular slip. If the load is essentially a uniaxial tension a good way to reduce slip at the boundaries is to orient them all in the direction of the load, so that no shear stress is acting on them. This can be achieved by what is called *directional solidification* (DS), used in the manufacture of gas-turbine blades by casting. Figure 4.51 shows the principle. An even more radical solution to the problem, now being developed widely, is to use this process to eliminate the grain boundaries completely and produce single-crystal blades.

Figure 4.52 shows that up to 1960 turbine inlet temperatures were virtually the same as the metal temperatures. After this date methods for cooling the blades – internally by circulating air through channels running the length of the blade, or externally by creating a cool boundary layer over the surface – enabled the inlet

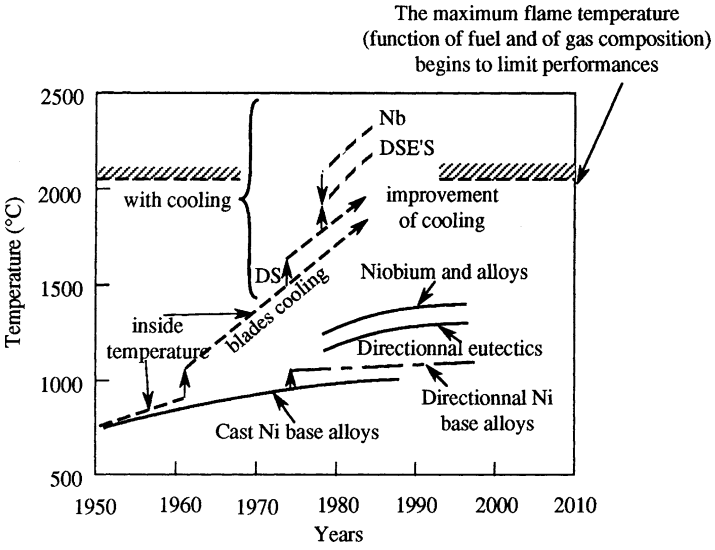


Fig. 4.52 Evolution of working temperature for turbine blades, and materials used for their manufacture (Adapted from Ashby and Jones 1980)

temperature to be raised without changing the material. The figure shows also the progress made possible by the use of DS materials and replacement of nickel-based alloys by niobium alloys or composite materials; and indicates possible further improvements that may result from studies of mixtures of conventional nickel-based alloys and directional eutectics (see also Chap. 1, Fig. 1.2).

4.5.2 Creep of Solid Solutions

Substantial reductions in the creep rate can be achieved by adding elements in solid solution. We have already mentioned the austenitic stainless steel with basic percentage composition 18Cr-12Ni, for which the rate was greatly reduced by adding 2.5% of molybdenum. Another example, of niobium alloys, is given in Fig. 4.53. For this the reinforcement resulting from the addition of various elements in solid solution (Mo, W, Re) is expressed by a parameter S_B defined by:

$$S_B = \left(\sigma_{alloy} - \sigma_{Nb}\right)/\sigma_{Nb} \tag{4.114}$$

where σ_{alloy} and σ_{Nb} are the stresses that have to be applied to the alloy and to niobium respectively to give a creep rate of 10^{-5} s^{-1}. It is important to realise that with the value $M = 3{-}5$ of the stress exponent, the values shown for S_B correspond to reductions in rate by 2–3 orders of magnitude.

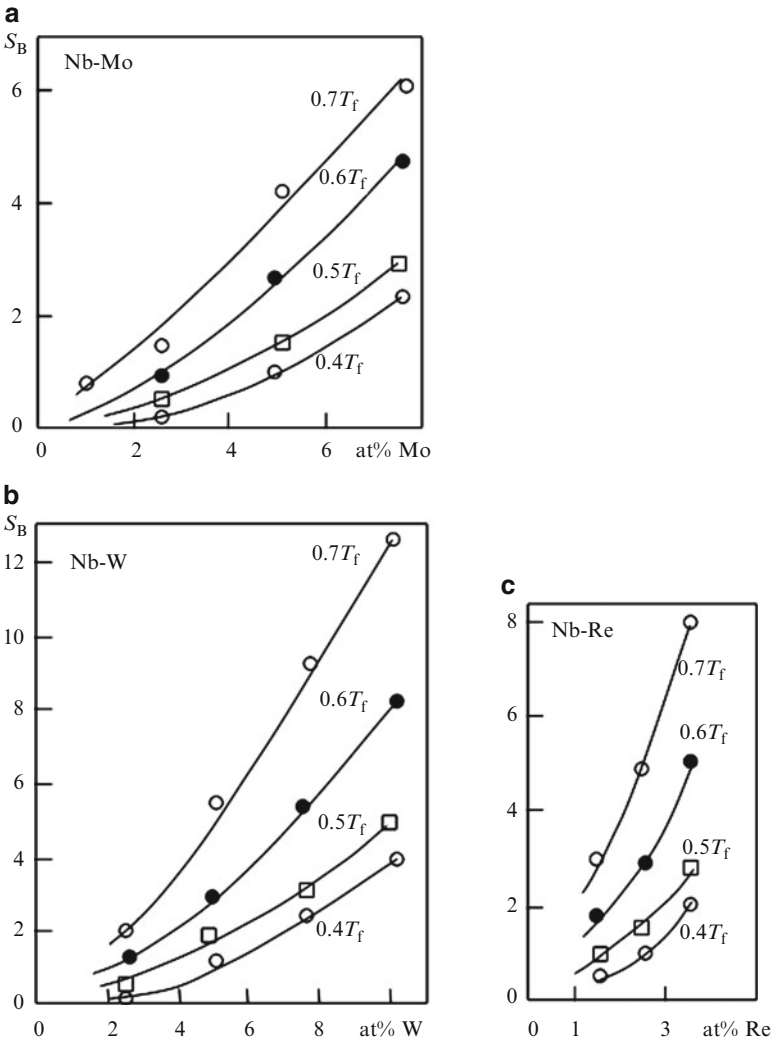


Fig. 4.53 Increased creep resistance of niobium resulting from additions of (**a**) Mo; (**b**) W; (**c**) Re

As in low-temperature plasticity, elements added in solid solution constitute obstacles to the movement of dislocations. There are three effects to consider:

– size,
– modulus,
– distortion, when the deformation field surrounding the solute atoms is not spherically symmetrical. This is especially the case for the interstitial atoms C, N, O in BCC metals (see Sect. 3.4.3.1).

Further, at high temperature the solute atoms are as mobile as those of the solvent, a situation which gives rise to what is often a rather serious difficulty in the detailed analysis of creep phenomena. It concerns essentially the choice of the diffusion coefficient, which, as we have seen, appears in the various models of creep. For dilute solutions the relevant coefficient is that of the solvent itself in the solution: stated simply, it is known that adding a "fast" solute increases the self-diffusion of the basic metal, whilst a "slow" solute reduces this. The effect is greater, the higher the solute concentration, up to a limit of around five atom %. This simple rule explains, qualitatively at least, the effect of adding an atom such as molybdenum to the matrix of an austenitic stainless steel.

With concentrated solid solutions the choice of the diffusion coefficient raises more difficult conceptual problems.

Another aspect of high-temperature behaviour of solutions and atomic mobility is the existence of "viscous drag". The main origins of this are the elastic interaction between the dislocations and the solute atoms (Cottrell effect, see Sect. 3.4.3.1d), the segregation of solute atoms on stacking faults (Suzuki effect, see Sect. 3.4.3.1f) and the destruction of short- or long-range order by the passage of a dislocation. Stated simply, this drag effect can be represented by the fact that the dislocations are submitted not only to the applied stress field but also to a force due to their interaction with the segregated atoms.

4.5.3 Creep of Alloys Reinforced by Particles

Another, and very effective, way to impede the movement of dislocations is to incorporate particles into the material. This can be done either by precipitation (as with nickel-based superalloys and creep-resistant ferritic steels) or by powder metallurgy (for example, particles of the very stable oxides of thorium or yttrium in nickel- or aluminium-based matrices, or of silicon carbide in aluminium-based matrices.)

Taking as example the nickel-based alloys used in manufacturing turbine blades, the solutions adopted are based mainly on the excellent temperature stability of compounds of the type of Ni_3Al (γ' phase). The hardening due to these γ' precipitates underlies the development of the superalloys, the most modern having over 50% of this phase.

Figure 4.54 shows the progress made in superalloy development over more than 30 years; the average increase in resistance can be expressed as a gain of almost 10°C per year over an entire 30 year period.

Another example is that of the ferritic/martensitic 9–12%Cr steels widely used in the thermal power generation industry where operational temperatures are in the neighbourhood of 550°C. The necessity for increased operating temperatures led to the development of compositional modifications in these materials. The first

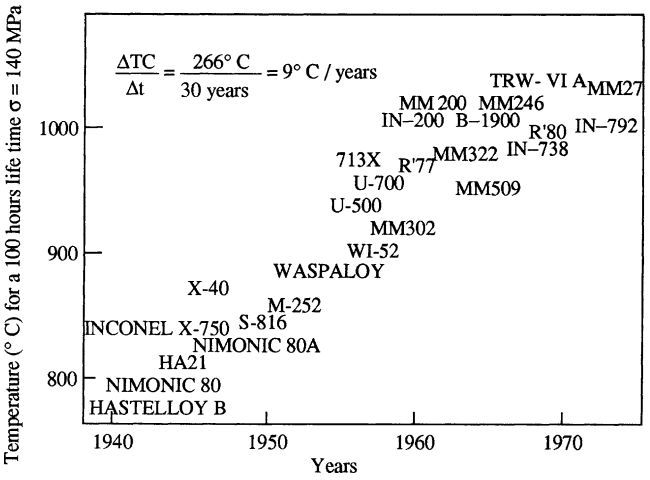


Fig. 4.54 The evolution of Ni-based superalloys

of these modifications was the addition of 1% molybdenum followed by tungsten, vanadium and nitrogen. Figure 4.55 summarises the various compositions that have been developed and gives the 600°C creep rupture stress. There is considerable hope in the most advanced compositional modifications for the fabrication of components to be used in fourth-generation fission reactors. Compared to austenitic stainless steels, these steels in addition to other advantages have superior thermal conductivity, a lower thermal expansion and less sensitivity to radiation-induced swelling, especially the ODS (oxide dispersion strengthened) versions that are produced by mechanical grinding using ball milling. The new ODS steels are strengthened by the formation of oxide nanoparticles which are extremely stable.

4.6 Exercise: Activation Energy Needed for Dislocations to By-Pass Ordered Precipitates

In nickel-based alloys hardened by precipitates of the coherent and ordered phase of the type Ni_3 (Ti, Al) (pseudo-FCC structure, Ll_2), it is found that a dislocation will shear small particles but will wrap around those above a certain critical size. The Orowan mechanism operates above this critical stress (see Sect. 3.4.3.2 and Annex 2). In this annex it is suggested to calculate the activation energies for these two modes of by-passing precipitates.

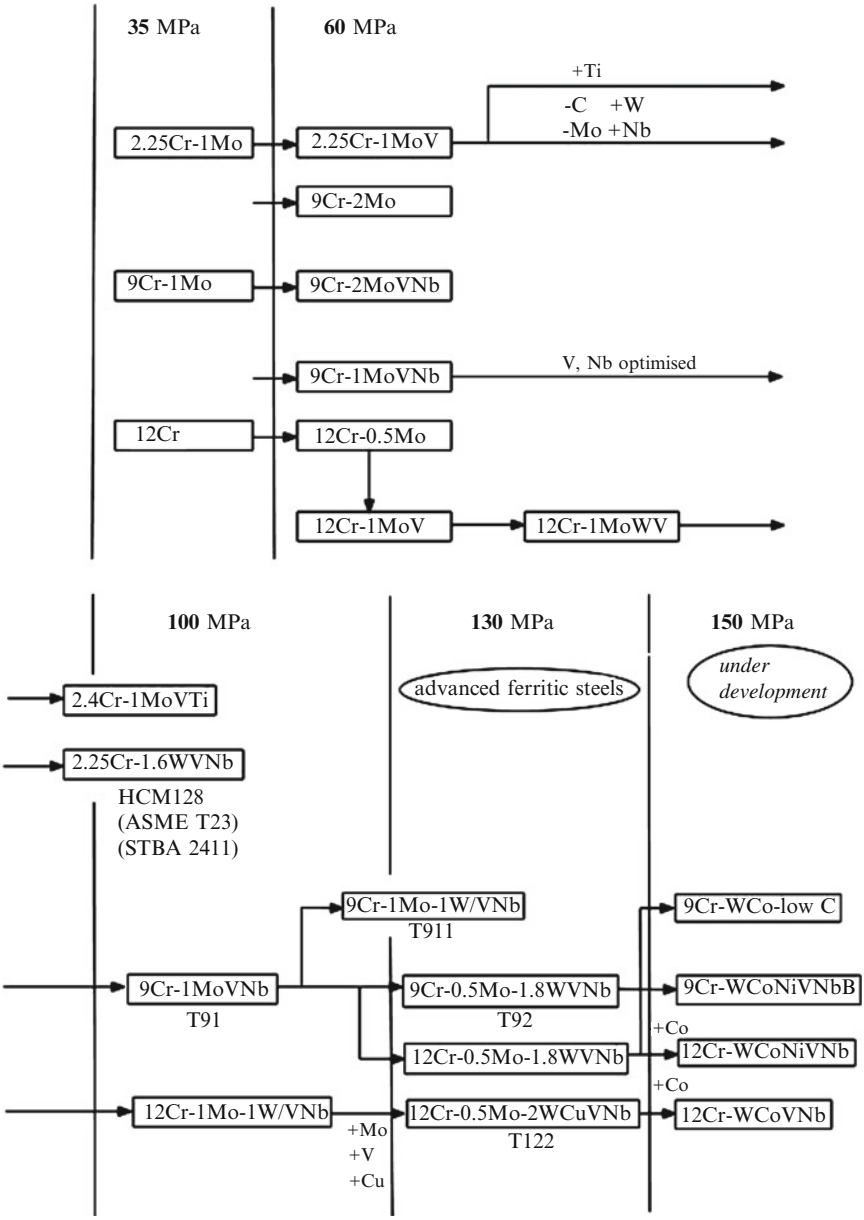


Fig. 4.55 Chromium-bearing ferritic and martensitic steels showing the evolution in chemistry and improvement in 100,000 h creep strength at 600°C (Antolovich and Pineau 2011)

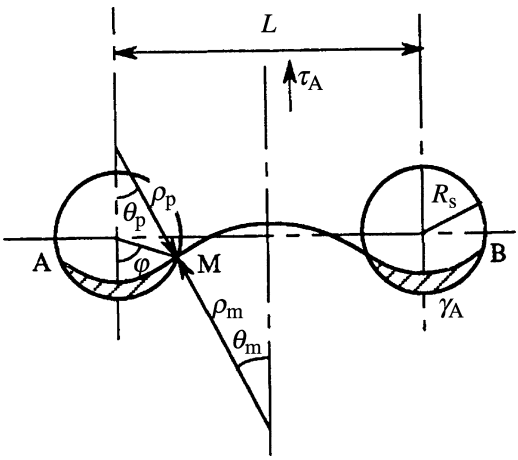


Fig. 4.56 Shearing of two precipitates

4.6.1 Shearing

Consider two particles a distance L apart in the slip plane, on which latter there is a shear stress τ_A (Fig. 4.56). The dislocation penetrates the particles, taking the position AMB; the antiphase energy γ_A which holds the dislocation line is responsible for the curvature.

4.6.1.1 By drawing the unit cell of the Ll2 structure (Annex 2), explain how this antiphase energy originates. Show that order can be restored by the passage of a second dislocation that has the same Burgers vector as the first.

4.6.1.2 Show that if R_0 is the mean radius of particles in the bulk of the material, those in the slip plane will have a mean radius $R_s = (\pi/4)R_0$

4.6.1.3 Show that the equilibrium equations for the arc of the dislocation are

$$\rho_p = t_p / (\gamma_A - b\tau_A) \tag{4.115}$$

$$\rho_m = t_m / b\tau_A \tag{4.116}$$

and that the following geometrical relations hold:

$$L/2 = \rho_m \sin\theta_m + \rho_p \sin\theta_p \tag{4.117}$$

$$\rho_p \sin\theta_p = R_s \sin\varphi \tag{4.118}$$

where t_m, t_p are the line tensions of the dislocation in the matrix and precipitates respectively. If these are equal then the junction is tangential and $\theta_m = \theta_p = \theta$.

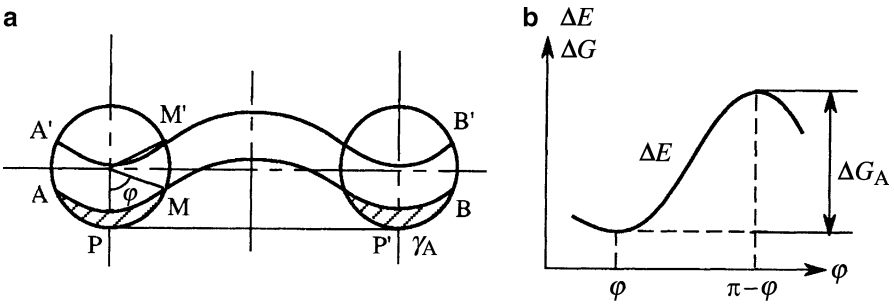


Fig. 4.57 Shearing of two particles by a dislocation: (**a**) stable and metastable positions of the dislocation; (**b**) variation of energy ΔE; ΔG_A is the activation energy

Using these relations, show that for any given value of τ_A there are two possible values φ *and* $\pi - \varphi$ that satisfy the equilibrium conditions (Fig. 4.57a). Show these on the energy diagram, as in Fig. 4.57b.

Note that in general, with $t_m \neq t_p$, Eqs. 4.115, 4.116, 4.117, and 4.118 are not sufficient to define the equilibrium: the further condition $d(\Delta E)/d\varphi = 0$ must be added, where

$$\Delta E = \left(t_m \Delta l_m + t_p \Delta l_p\right) - b\tau_A \Delta S_m - (b\tau_A - \gamma_A)\,\Delta S_p \tag{4.119}$$

4.6.1.4 Compute the energy variation ΔE during the crossing; then, starting from some reference position, for example PP' in Fig. 4.57, find:

- the increases in length Δl_m, Δl_p of the dislocation in the matrix and the precipitates respectively,
- the areas ΔS_m, ΔS_p swept by the dislocation in the matrix and the precipitates.

Substitute these results in (4.119) and, with $d(\Delta E)/d\varphi = 0$, derive the relation

$$2\left(t_p \cos\theta_p - t_m \cos\theta_m\right) + (2R_s\gamma_A \sin\varphi - bL\tau_A)\tan\varphi = 0 \tag{4.120}$$

Show that Eqs. 4.115, 4.116, 4.117, 4.118 and 4.120 together enable the complete solution of the problem to be found.

4.6.1.5 Assume throughout that $t_m = t_p$. Show that:

$$\sin\varphi = bL\tau_A / 2\gamma_A R_s \tag{4.121}$$

$$\tau_A = \tau_A{}^c\,(\varphi = \pi/2) = 2\gamma_A R_s / bL \tag{4.122}$$

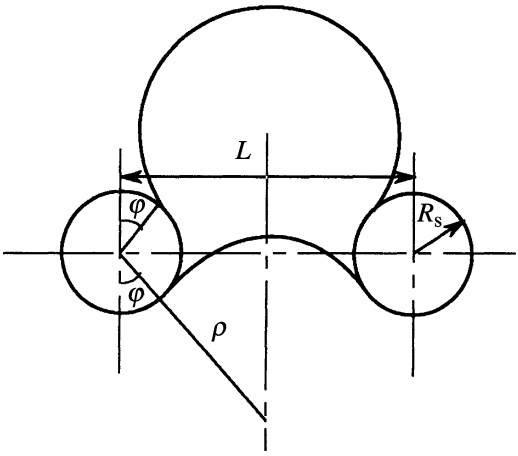


Fig. 4.58 By-passing of two precipitates by Orowan mechanism

where τ_A^c corresponds to the stress needed for crossing at absolute zero, when there is complete absence of thermal activation.

Show that the activation energy ΔG_A can be expressed in the simple form:

$$\Delta G_A = \gamma_A {R_s}^2 (\pi - 2\varphi - \sin 2\varphi) \tag{4.123}$$

and the activation area A^*

$$A^* = 2LR_s \cos\varphi = 2LR_s(1 - bL\tau_A / 2\gamma_A R_s)^{1/2} \tag{4.124}$$

4.6.1.6 In the neighborhood of 0 K the crossing energy is given approximately by (4.122). Show that here the expression for ΔG_A can be simplified to:

$$\Delta G_A = \pi\gamma_A {R_s}^2 (1 - 2bL\tau_A / \pi\gamma_A R_s) \tag{4.125}$$

4.6.1.7 Numerical application; evaluate ΔG_A for $\gamma_A = 0.2$ Jm^{-2}, $R_s = 22{\cdot}10^{-10}$ m and compare the value with the thermal activation energy at ambient temperature. Is it likely that shearing by thermal activation will be possible?

4.6.2 Orowan By-Passing

Use the same reasoning and the same methods for attacking the problems as before. Figure 4.58 gives the geometry; note the two equilibrium positions for the dislocation, as before.

4.6.2.1 At 0 K, with no thermal activation, $\varphi = \pi/2$ and the two equilibrium positions coincide. Deduce that the Orowan stress is:

$$\tau_{\mathrm{A}}^{\mathrm{c}} = \frac{2t}{b\,(L - 2R_{\mathrm{s}})} \tag{4.126}$$

4.6.2.2 Proceeding as in Sect. 4.6.1.6, show that

$$\Delta G_{\mathrm{A}} = \frac{(t + bR_{\mathrm{s}}\tau_{\mathrm{A}})}{b\tau_{\mathrm{A}}}\,(\pi - 2\varphi - \sin 2\varphi) \tag{4.127}$$

4.6.2.3 Show that high energies are needed for this mode of by-passing.

References

Antolovich SD, Pineau A (2011) High temperature fatigue. In: Bathias C, Pineau A (eds) Fatigue of materials and structures, Chapter 1, volume I: 1-130, ISTE Ltd (London), John Wiley and Sons Ltd (Hoboken, NJ)

Ashby FM, Jones DRh (1980) Engineering materials 1. An introduction to their properties and applications. Int Series on Materials Science and Technology, volume 34. Pergamon International Library

Ashby MF, Verral RA (1973) Diffusion-accommodated flow and superplasticity. Acta Metall 21:149–163

Baudelet B, Suéry M (1987) Mise en forme superplastique. Techniques de l'ingénieur, M13, Paris

Ben Hadj Hamouda H, Laiarinandrasana L, Piques R (2007) Viscoplastic behaviour of a medium density polyethylene (MDPE): Constitutive equations based on double nonlinear deformation model. Int J Plasticity 23:1307–1327

Besson J, Cailletaud G, Chaboche JL, Forest S, Bletry M (2009) Non-linear mechanics of materials. Springer

Boussuge M (2009) Verres et céramiques in Endommagement et rupture des matériaux 2. Clavel M, Bompard Ph (eds) Lavoisier Paris

Caillard D (2010a) Kinetics of dislocations in pure Fe. Part I: In-situ straining experiments at room temperature. Acta Mater 58:3493–3503

Caillard D (2010b) Kinetics of dislocations in pure Fe. Part II: In-situ straining experiments at low temperature. Acta Mater 58:3504–3515

Campbell JD, Ferguson WG (1970) The temperature and strain rate dependence of shear strength of mild steel. Phil Mag 21:63–82

Cannon WR, Langdon TG (1983) Review – creep of ceramics. Part I: mechanical characteristics. J Mater Sci 18:1–50

Cannon WR, Langdon TG (1988) Review – creep of ceramics. Part II: an examination of flow mechanisms. J Mater Sci 23:1–20

Clarke RR (1987) On the equilibrium thickness of intergranular glass phases in ceramic materials. J Am Ceram Soc 70(1):15–22

Coble RL (1963) A model for boundary diffusion-controlled creep in polycrystalline materials. J Appl Phys 34:1679–1682

Dorn JE (1957) Creep and recovery. American Society for Metals, Cleveland

Evans R, Wilshire B (1985) Creep of metals and alloys. The Institute of Metals, London

Evans R, Wilshire B (1993) Introduction to creep. The Institute of Metals, London

Friedel J (1964) Dislocations, Pergamon Press

Frost HJ, Ashby MF (1982) Deformation-Mechanism Maps. The plasticity and creep of metals and ceramics. Pergamon Press

Garofalo F (1965) Fundamentals of creep and creep-rupture in metals. The MacMillan Company, New York

Gilman JJ (1959) The plastic resistance of crystals. Australian Journal of Physics 13:327

Herring C (1950) Diffusional viscosity of a polycrystalline solid. J Appl Phys 21:437–445

Hoff NJ (1953) The necking and the rupture of rods subjected to constant tensile loads. J Appl Mech Trans ASME 20:105–108

Kaddour D, Fréchinet S, Gourgues A-F, Brachet JC, Portier L, Pineau A (2004) Experimental determination of creep properties of zirconium alloys together with phase transformation. Scripta Mater 51:515–519

Kaddour D, Gourgues-Lorenzon AF, Brachet JC, Portier L, Pineau A (2011) Microstructural influence on high temperature creep flow of Zr – 1%NbO alloy in near α, $(\alpha + \beta)$, and β temperature ranges in a high vacuum environment. J Nucl Mater 408:116–124

Larson FR, Miller J (1952) A time-temperature relationship for rupture and creep stresses. Trans ASME 74:765–775

Lemaitre J, Chaboche JL (1990) Mechanics of solids materials. Cambridge University Press, Cambridge

Mason WP (1968) In Physical Acoustic. Academic Press, New York

McClintock FA, Argon A (1966) Mechanical behavior of materials. Addison-Wesley Publishing Company, Reading, MA, Melon Park, California, London, Amsterdam, Don Mills, Ontario, Sidney

McLean D (1962) Deformation at high temperatures. Met Rev 7:481–527

Meyers MA, Chawla KK (1984) Mechanical metallurgy: Principles and applications. Prentice-Hall - Englewood Cliffs, NJ

Mukherjee AK, Bird JE, Dorn JE (1969) Experimental correlations for high temperature creep. Trans Am Inst Min 62:155–179

Muto H, Sakai M (1998) Grain boundary sliding and grain interlocking in the creep deformation of two-phase ceramics. J Am Ceram Soc 81:1611–1621

Nabarro FRN (1947) Dislocations in a simple cubic lattice. Proc Phys Soc 59:256–272

Nabarro FRN (1952) Mathematical theory of stationary dislocations. Advances in Physics 1: 269–394

Nabarro FRN (1967) Theory of crystal dislocation. Clarendon, Oxford, Clarendon 22:799–804

Perevezentsev VN, Rybin VV, Chuvil'deev VN (1992) The theory of structural superplasticity Parts I-IV. Acta Metall et Mater 40:887–923

Poirier JP (1985) Creep of crystals: high temperature deformation processes in metals, ceramics and minerals. Cambridge University Press, Cambridge

Raj R, Ashby MF (1971) On grain boundary sliding and diffusional creep. Met Trans 2:1113–1127

Regrain C, Laiarinandrasana L, Toillon S, Saï K (2009) Multimechanism models for semi-crystalline polymer: constitutive relations and finite element implementation. Int J Plasticity 25:1253–1279

Saï K, Laiarinandrasana L, Ben Naceur J, Besson J, Jeridi M, Cailletaud G (2011) Multi-mechanism damage – plasticity model for semi-crystalline polymer: creep damage of notched specimen of PA6. Mat Sci Eng A528:1087–1093

Sellars CM, Teggart WJMcG (1966) La relation entre la résistance et la structure dans la déformation à chaud. Mem Sci Rev Metall 63:731–746

Stein DF, Low JR (1969) Mobility of edge dislocations in silicon-iron crystals. J Appl Phys 31:362–369

Vairis A (2008) Superplastic effects and strain rate dependency in a material joining process. J Eng Sci Tech Rev 1:28–32

Weertman J (1968) Dislocation climb theory of steady-state creep. Trans ASME 61:681–694

Wilkinson DS (1998) Creep mechanisms in multiphase ceramic materials. J Am Ceram Soc 81:275–299

Williams SJ, Bach MR, Wilshire B (2010) Recent developments in the analysis of high temperature creep and creep fracture behaviour. Met Sci Tech 26:1332–1337

Chapter 5
Viscoelasticity

Abstract Experimental demonstration of viscoelastic behaviour includes creep and relaxation tests, recovery tests and dynamic mechanical analysis (DMA). Formulation of linear viscoelastic functions are given and the Boltzman superposition principle explained. The Laplace-Carson transform allows simplifying the one-dimensional constitutive behaviour in non-ageing linear viscoelasticity. This leads to representation with the use of partial differential equations. Spectral representation is a generalisation. Thus, DMA can be performed. It is explained how to check the linearity of the behaviour.

Polymers are an important case of viscoelastic materials. This behaviour is linked to the conformation of chains in amorphous polymers. Temperature effects result: glass transition temperature, time-temperature superposition principle. The relaxation mechanisms and viscoelastic behaviour are discussed. The glass transition and the β transition are related to the structure of polymers. Physical ageing is discussed as well as the case of semi-crystalline polymers.

Another manifestation of viscoelasticity is internal friction in metals. There is a distinction between relaxation peaks and hysteretic behaviour. The point defect relaxations include the Snoek and the Zener effects. The Bordoni, the Hasiguti, the Snoek-Köster, the damping background at elevated temperatures are dislocation induced relaxation phenomena. Grain boundaries are also sources of relaxation internal friction. At higher amplitudes, dislocations produce hysteretic damping. Damping can be very high when due to phase transformations.

The 3-D formulation of constitutive equations is then given. The correspondence theorem allows structural design. Finally, the analysis of the overall behaviour of heterogeneous materials through the estimation of the effective creep moduli and compliances is discussed.

This chapter is devoted to the study of one kind of constitutive mechanical behaviour, which is intermediate between the one of Chap. 2 (viscosity is considered now in addition to elasticity) and the one of Chap. 4 (no plastic strain is

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1_5, © Springer Science+Business Media B.V. 2012

present anymore). The sole *coupling* of elasticity and viscosity is responsible for specific mechanical responses and it needs specific experimental investigations and modelling approaches. Almost all materials obey this behaviour within a definite domain of stress amplitude and temperature but this domain is especially large for *polymers*, which are mainly concerned with what follows. Nevertheless, metals are considered too, in particular for what concerns *internal friction* investigations. Attention is focussed on *linear* (mainly *non-ageing*) viscoelasticity which has been the matter of deeper analysis and of the development of many efficient experimental and theoretical tools. After a phenomenological analysis of 1-D typical responses, we take up the study of microstructural aspects of viscoelasticity and then we introduce methods for structural design and for the analysis and modelling of the behaviour of viscoelastic heterogeneous materials.

5.1 Phenomenological Analysis of 1-D Mechanical Responses

5.1.1 Experiments

Let q and Q be the displacement and force variables, respectively, and $\dot{q}$ and $\dot{Q}$ their rates. As we know (see Sect. 1.4.2), for perfect elasticity, when either q (or Q) is imposed instantaneously and maintained constant, the other variable, Q (or q), takes its equilibrium value instantaneously and does not evolve any longer. For pure viscosity, the same is true for $\dot{q}$ and Q: with Q fixed, $\dot{q}$ is fixed too; when q is constant, $\dot{q}$ is null and Q must be constant (actually, q cannot be applied instantaneously so that Q is not necessarily null, but it remains constant as soon as q is constant). The *viscoelastic coupling* is responsible for more complex responses: generally speaking, the response at any time *depends on the whole load history* from the very beginning. Since the material keeps some memory of all what it has been subjected to in the past, the viscoelastic behaviour is a *hereditary* one (as it was already the case for elasto(visco)plasticity). It can be characterised by a few standard tests which are presented now for unidimensional situations. In this section, we consider only *isothermal evolutions*. Dependence of the viscoelastic behaviour on temperature is discussed in Sect. 5.2, both for metals and polymers.

5.1.1.1 Creep (or Retardation) Test

Q is applied instantaneously at the value Q_0 at time t_0 and q, which was zero up to t_0, is recorded as a function of time t for $t \geq t_0$; it may have the discontinuity q_0 at $t = t_{0_+}$ and then it has a continuous, increasing non-accelerated variation (Fig. 5.1); $\dot{q}$ may tend towards the constant $\dot{q}_\infty$. The *creep test* exhibits the creep faculty of the viscoelastic material, *i.e.*, its ability to flow when subjected to a constant force. The difference with the viscoplastic creep depicted in Sect. 4.2.1.1 lies in the fact that

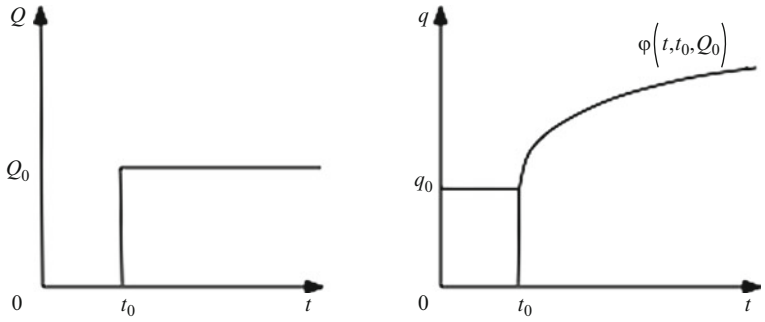


Fig. 5.1 Creep test

there is now no plastic strain in the response q: even if part of this response can be irreversible, this fact occurs for any value of Q_0, as low as it is. For this reason, the term "retardation" is sometimes preferred to "creep" for viscoelasticity, what we do not do in the following.

This test may be summed up as follows:

$$Q(t,t_0) = Q_0 \mathrm{H}(t-t_0) \quad \Rightarrow \quad q(t,t_0) = \varphi(t,t_0,Q_0) \tag{5.1}$$

where $\mathrm{H}(t-t_0)$ is the Heaviside[1] unit step function at time t_0 and $\varphi(t,t_0,Q_0)$ (which could also be written as $\varphi(t,t_0,Q_0)\mathrm{H}(t-t_0)$ since q was zero up to t_0) is the general form of the creep function. At this stage, it can be a linear as well as a nonlinear function of Q_0.

5.1.1.2 Relaxation Test

The role of Q and q are inverted with respect to the creep test: q is applied instantaneously at the value q_0 at time t_0 and Q, which was zero up to t_0, is recorded as a function of time t for $t \geq t_0$; it may have the discontinuity Q_0 at $t = t_{0+}$ and then it has a continuous, decreasing variation (Fig. 5.2); Q may tend towards the constant Q_∞. The *relaxation test* exhibits the relaxation faculty of the viscoelastic material, *i.e.*, the ability of the stress variable to decrease with time at constant strain.

This test may be summed up as follows:

$$q(t,t_0) = q_0 \mathrm{H}(t-t_0) \quad \Rightarrow \quad Q(t,t_0) = \rho(t,t_0,q_0) \tag{5.2}$$

where $\rho(t,t_0,q_0)$ (which could also be written as $\rho(t,t_0,Q_0)\mathrm{H}(t-t_0)$ since Q was zero up to t_0) is the general form of the relaxation function. At this stage, it can also be a linear as well as a nonlinear function of q_0.

[1] Oliver Heaviside (1850–1925) was a British physicist.

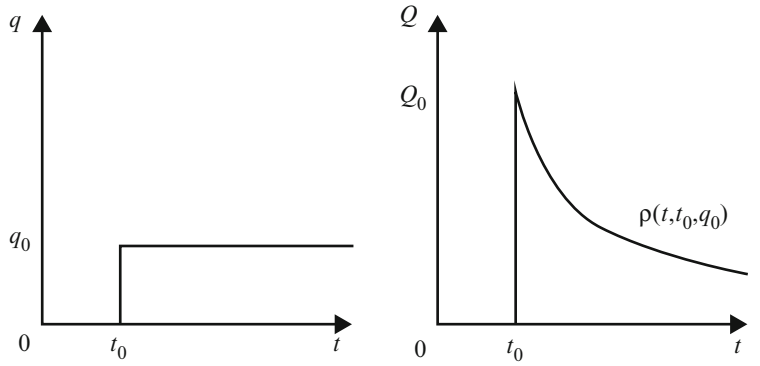


Fig. 5.2 Relaxation test

Note that, despite their similarity, there are some differences between the creep and relaxation tests: first, it is always possible to impose instantaneously a force but not a displacement (*e.g.* for a dashpot, which has no instantaneous elasticity); second, the time constants of these two tests may be significantly different from one another (relaxation is generally much faster than creep).

5.1.1.3 Recovery Tests

These tests are concerned with the response to boxcar (rectangle) functions, instead of step functions above, for either Q (strain-recovery) or q (stress-recovery) (Fig. 5.3). The *strain-recovery test* is much more used (Fig. 5.3a): it begins like a creep test at time t_0 but the force Q_0 is suddenly removed at time $t_1 > t_0$. At $t = t_1$, the strain variable q may have a discontinuity and it decreases continuously afterwards. The strain-recovery may be complete ($\lim_{t\to\infty} q(t) = 0$) or only partial ($\lim_{t\to\infty} q(t) = q_\infty > 0$), so that q_∞ is irreversible (but not plastic since this occurs with Q_0 as low as wished).

Note that the Q-history can be expressed as

$$Q(t, t_0, t_1) = Q_0 \left[\mathrm{H}(t - t_0) - \mathrm{H}(t - t_1)\right] \tag{5.3}$$

but the q-response after t_1 cannot be predicted without additional assumption, such as linearity (see after).

5.1.1.4 Dynamic Mechanical Analysis (DMA)

A number of other tests exist for the experimental investigation of viscoelasticity, including the classical test at constant strain-rate. The most frequent experimental

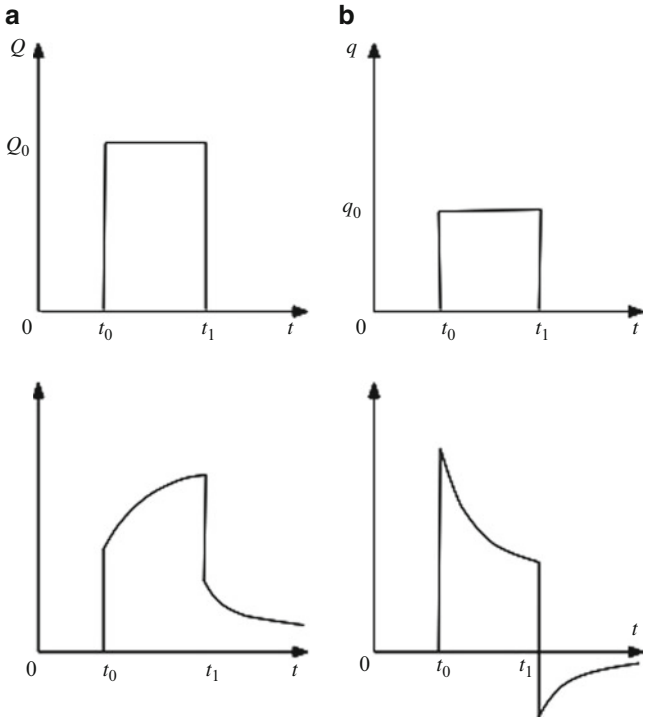


Fig. 5.3 Recovery tests (**a**) strain-recovery; (**b**) stress-recovery (The two lower plots relate q and t at left and Q and t at right)

investigation refers to the "*dynamic mechanical analysis*" (DMA) for polymers and to "*internal friction experiments*" for metals. The principle is the same for both tests: for several temperatures, the sample is subjected to an oscillatory (usually sinusoidal) force or displacement at a given frequency, not too high for the inertia effects to be negligible, and the response is recorded as a function of time. The interpretation of the results is much easier for linear viscoelasticity: as shown below, the response oscillates then at the same frequency as the load but there is a phase lag between them. The variation of the amplitude and of the phase lag of the response with frequency and temperature gives very useful information on both the *damping capacity* of the material and its mechanical behaviour. This is shown below in this section (Sect. 5.1.3.5) and illustrated for various materials in Sect. 5.2.

5.1.2 Linear Viscoelasticity

5.1.2.1 Linear Creep and Relaxation Functions

We now restrict ourselves to *linear* viscoelasticity. Before giving a general definition for it, we can notice simply that, since linearity implies that in a creep test, when the

force Q_0 is multiplied by λ, the displacement response given by (5.1) as $q(t, t_0) = \varphi(t, t_0, Q_0)$ must also be multiplied by λ, the general form of the creep function $\varphi(t, t_0, Q_0)$ is now specified as

$$\varphi(t, t_0, Q_0) = Q_0 \, \mathrm{f}(t, t_0) \tag{5.4}$$

where $\mathrm{f}(t, t_0)$ is the *linear* creep function. Similarly, for a relaxation test, we must have

$$\rho(t, t_0, q_0) = q_0 \mathrm{r}(t, t_0) \tag{5.5}$$

where $\mathrm{r}(t, t_0)$ is the *linear* relaxation function.

5.1.2.2 The Boltzmann[2] Superposition Principle

These comments can be generalised by application of the Boltzmann superposition principle (Boltzmann 1874) which gives the general condition for a physical system to behave linearly: this principle states that the response of this system to the superposition of several loads is the superposition of the responses to each one. For a linear viscoelastic system defined by the parameters (Q, q), this reads for two Q-histories $Q^{(1)}$ and $Q^{(2)}$

$$\left(Q^{(1)} \Rightarrow q^{(1)},\ Q^{(2)} \Rightarrow q^{(2)}\right) \quad \Rightarrow \quad \left(\lambda_1 Q^{(1)} + \lambda_2 Q^{(2)} \Rightarrow \lambda_1 q^{(1)} + \lambda_2 q^{(2)}\right) \tag{5.6}$$

Similarly, for two q-histories $q^{(1)}$ and $q^{(2)}$,

$$\left(q^{(1)} \Rightarrow Q^{(1)},\ q^{(2)} \Rightarrow Q^{(2)}\right) \quad \Rightarrow \quad \left(\lambda_1 q^{(1)} + \lambda_2 q^{(2)} \Rightarrow \lambda_1 Q^{(1)} + \lambda_2 Q^{(2)}\right) \tag{5.7}$$

These definitions obviously imply in particular the properties (5.4) and (5.5). In addition, one can note that the recovery tests can be easily related to the creep and relaxation tests. For example, since the strain-recovery test corresponds to the Q-history given by the superposition (5.3), the response must be

$$\begin{cases} Q_0 \mathrm{H}(t - t_0) \Rightarrow Q_0 \, \mathrm{f}(t, t_0) \\ Q_0 \mathrm{H}(t - t_1) \Rightarrow Q_0 \, \mathrm{f}(t, t_1) \end{cases} \Rightarrow q_0(t, t_0, t_1) = Q_0 \left[\mathrm{f}(t, t_0) - \mathrm{f}(t, t_1)\right] \tag{5.8}$$

Similarly, for a stress-recovery test, the response must be

$$\begin{cases} q_0 \mathrm{H}(t - t_0) \Rightarrow q_0 \mathrm{r}(t, t_0) \\ q_0 \mathrm{H}(t - t_1) \Rightarrow q_0 \mathrm{r}(t, t_1) \end{cases} \Rightarrow Q_0(t, t_0, t_1) = q_0 \left[\mathrm{r}(t, t_0) - \mathrm{r}(t, t_1)\right] \tag{5.9}$$

[2] Ludwig Boltzmann (1844–1906) was an Austrian physicist famous for his founding contributions in the fields of statistical mechanics and thermodynamics.

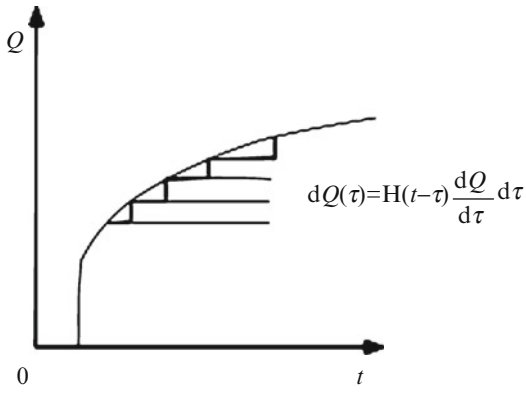


Fig. 5.4 Decomposition of a Q-history into incremental steps

More generally, any Q-history starting from $t = t_0$ may be considered (Fig. 5.4) as the superposition of the incremental steps $\mathrm{d}Q(\tau) = \frac{\mathrm{d}Q}{\mathrm{d}\tau}\mathrm{d}\tau \mathrm{H}(t-\tau)$.

Since, according to (5.4), the response to such a step is $\frac{\mathrm{d}Q}{\mathrm{d}\tau}\mathrm{d}\tau\, \mathrm{f}(t,\tau)$, the whole response must be

$$q(t) = Q(0_+)\mathrm{f}(t,0) + \int_0^t \mathrm{f}(t,\tau)\mathrm{d}Q(\tau) \tag{5.10}$$

where the response to a possible discontinuity of Q at $\tau = 0_+$ has been set apart. Integration by parts leads to another, equivalent, expression, namely

$$q(t) = Q(t)\,\mathrm{f}(t,t) - \int_0^t Q(\tau)\frac{\partial \mathrm{f}(t,\tau)}{\partial \tau}\mathrm{d}\tau \tag{5.11}$$

where the instantaneous and the delayed responses have been clearly separated.

Dual expressions are obtained when a q-history is prescribed. They involve the linear relaxation function $\mathrm{r}(t,\tau)$:

$$\begin{cases} Q(t) = q(0_+)\mathrm{r}(t,0) + \int_0^t \mathrm{r}(t,\tau)\mathrm{d}q(\tau) \\ Q(t) = q(t)\mathrm{r}(t,t) - \int_0^t q(\tau)\frac{\partial \mathrm{r}(t,\tau)}{\partial \tau}\mathrm{d}\tau \end{cases} \tag{5.12}$$

At this stage, the 1-D behaviour of a linearly viscoelastic material is fully determined by the knowledge of the two scalar functions $\mathrm{f}(t,\tau)$ and $\mathrm{r}(t,\tau)$. Actually, these two functions are not independent and any of them can be derived from the other by calculation. As a matter of fact, for instance, if $\mathrm{f}(t,\tau)$ is known, (5.10) or (5.11) can be used for a Q-history, which reduces to the linear relaxation function

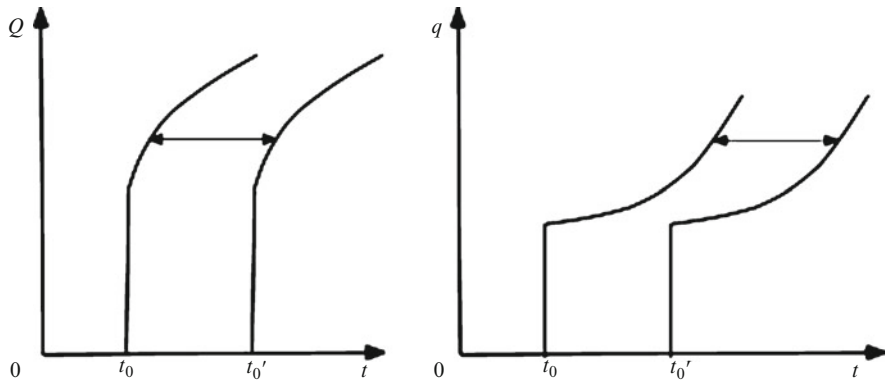


Fig. 5.5 Response to a time-translation of the Q-history for a non-ageing material

$r(t, \tau)$: the q-response must then be the unit step-function $H(t - \tau)$. When Q and q in these equations are replaced by the relaxation function and the unit step-function respectively, an integral equation is obtained which allows us in principle to calculate the relaxation function from the creep function. Obviously, the reverse is true when q and Q in (5.12) are replaced by the creep function and the unit step-function, respectively.

Though the property of linearity greatly simplifies, as we can see, the treatment of viscoelasticity, this treatment still remains rather difficult since complex convolution products appear in the constitutive equations (5.10), (5.11), (5.12). An important simplification derives from the additional assumption of *non-ageing* linear viscoelasticity.

5.1.3 Non-ageing Linear Viscoelasticity

5.1.3.1 Definition and Constitutive Equations

Ageing is a general property of materials which makes its physical properties evolve with time due to microstructural evolutions deriving from various causes: chemical reactions, radiations, long term thermally-activated phenomena... Nevertheless, it may exist a sufficiently long time interval during which such evolutions are negligible. If so (Fig. 5.5), two identical Q-histories starting at two times t_0 and t'_0 will cause identical q-responses, except for the translation $(t'_0 - t_0)$. In particular, for a non-ageing linear viscoelastic material, two creep tests with the same unit amplitude $H(t-t_0)$ and $H(t-t'_0)$ are such that the responses $f(t,t_0)$ and $f(t,t'_0)$ are translated from each other by $(t'_0 - t_0)$. This means that the creep function $f(t,\tau)$ actually only depends on $(t-\tau)$: we write it $f(t-\tau)$. The same must be true for the relaxation function which reads now $r(t-\tau)$.

The constitutive equation (5.11) now reads

$$\begin{aligned} q(t) &= Q(t)\mathrm{f}(0) + \int_0^t Q(\tau)\mathrm{f}'(t-\tau)\mathrm{d}\tau \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t Q(\tau)\mathrm{f}(t-\tau)\mathrm{d}\tau\right] \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t Q(t-\tau)\mathrm{f}(\tau)\mathrm{d}\tau\right] \end{aligned} \tag{5.13}$$

So, $q(t)$ is simply the time-derivative with respect to t of the classical convolution product of Q and f, say $(\mathrm{f} * Q)$. We write this derivative of a convolution product, which is named a Stieljes[3] convolution product,

$$q(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{f} * Q) = \mathrm{f} \odot Q = Q \odot \mathrm{f} \tag{5.14}$$

Similarly, we have for the response to a q-history:

$$Q(t) = \frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{r} * q) = \mathrm{r} \odot q = q \odot \mathrm{r} \tag{5.15}$$

We can combine (5.14) and (5.15) to obtain a relation between f and r. It reads

$$\begin{cases} q(t) = \mathrm{f} \odot (\mathrm{r} \odot q) = (\mathrm{f} \odot \mathrm{r}) \odot q \\ Q(t) = \mathrm{r} \odot (\mathrm{f} \odot Q) = (\mathrm{r} \odot \mathrm{f}) \odot Q \end{cases} \tag{5.16}$$

Since the Heaviside function $\mathrm{H}(t)$ is the only solution to the equation

$$\mathrm{H} \odot F = F, \quad \forall F \tag{5.17}$$

this means that we must have

$$\mathrm{r} \odot \mathrm{f} = \mathrm{f} \odot \mathrm{r} = \mathrm{H}(t) \tag{5.18}$$

so that *one* single function of *one* variable, either $\mathrm{f}(t-\tau)$ or $\mathrm{r}(t-\tau)$, is needed to characterise the whole 1-D mechanical behaviour (at a given temperature).

[3] Thomas Joannes Stieljes (1856–1894) was a Dutch mathematician.

5.1.3.2 Use of the Laplace-Carson Transformation

Calculations of mechanical responses of a linear non-ageing viscoelastic material are made much easier by use of a mathematical transformation which could help dealing with the Stieljes convolution products. The Laplace[4] transform $\mathcal{L}\{\mathrm{f}(t)\}$ is known to transform the convolution product of two functions of time, say $\mathrm{f} * \mathrm{g}$, into the simple algebraic product of the transformed functions:

$$\begin{cases} \mathcal{L}\{\mathrm{f}(t)\} = \int_0^\infty \exp(-pt)\,\mathrm{f}(t)\mathrm{d}t \\ \mathcal{L}\{\mathrm{f} * \mathrm{g}\} = \mathcal{L}\{\mathrm{f}\}\,\mathcal{L}\{\mathrm{g}\} \end{cases} \tag{5.19}$$

where p is a complex variable. In addition, we have the property

$$\mathcal{L}\left\{\frac{\mathrm{df}}{\mathrm{d}t}\right\} = p\left[\mathcal{L}\{\mathrm{f}\} - \mathrm{f}(0_+)\right] \tag{5.20}$$

Consequently, a Laplace-transformed Stieljes convolution product reads

$$\mathcal{L}\{f \odot \mathrm{g}\} = \mathcal{L}\left\{\frac{\mathrm{d}}{\mathrm{d}t}(f * \mathrm{g})\right\} = p\mathcal{L}\{\mathrm{f}\}\,\mathcal{L}\{\mathrm{g}\} \tag{5.21}$$

This result could be of great use in order to deal with (5.14) or (5.15), but it would be still more convenient to use the so-called "Laplace-Carson (or Carson[5]) transform" $\mathrm{f}^*(p)$ of any function $\mathrm{f}(t)$ defined by

$$\mathrm{f}^*(p) = p\int_0^\infty \exp(-pt)\,\mathrm{f}(t)\mathrm{d}t \tag{5.22}$$

It is such that we get simply

$$\{\mathrm{f} \odot \mathrm{g}\}^* = p\mathcal{L}\{\mathrm{f} \odot \mathrm{g}\} = \mathrm{f}^*(p)\mathrm{g}^*(p) \tag{5.23}$$

So, the Carson-transformed constitutive equations (5.14) and (5.15) read

$$\begin{cases} q^*(p) = \mathrm{f}^*(p)Q^*(p) \\ Q^*(p) = \mathrm{r}^*(p)q^*(p) \end{cases} \tag{5.24}$$

[4] Pierre-Simon de Laplace (1749–1827) was a French mathematician and astronomer.

[5] John Renshaw Carson (1886–1940) was an American mathematician and electrical engineer.

which, for any fixed value of the parameter p, are nothing but the constitutive equations of *linear elasticity*. This comment will take its whole importance when we deal with 3-D behaviour and structural design in Sect. 5.3.

In addition to (5.23), the most useful properties of the Carson transform are the following

$$\begin{gathered}
[\mathrm{CH}(t)]^* = \mathrm{CH}^*(p) = \mathrm{C} \quad \forall \mathrm{C\ constant} \\
\left(\frac{\mathrm{df}}{\mathrm{d}t}\right)^* = p\,[\mathrm{f}^*(p) - \mathrm{f}(0_+)] \quad (\text{or } p\mathrm{f}^*(p) \text{ if f is a distribution}) \\
\left[\int_0^t \mathrm{f}(\tau)\mathrm{d}\tau\right]^* = \tfrac{1}{p}\mathrm{f}^*(p)
\end{gathered} \tag{5.25}$$

From (5.18) and (5.25), $\mathrm{f}^*(p)$ and $\mathrm{r}^*(p)$ are simply connected by

$$\mathrm{f}^*(p)\mathrm{r}^*(p) = 1 \tag{5.26}$$

From the knowledge of $\mathrm{f}(t)$ (or $\mathrm{r}(t)$) and of the prescribed Q(or q)-history, one can then calculate $\mathrm{f}^*(p)$ (or $\mathrm{r}^*(p)$) and $Q^*(p)$ (or $q^*(p)$) and deduce from them $q^*(p)$ (or $Q^*(p)$) by (5.24). It still remains to derive $q(t)$ (or $Q(t)$) from $q^*(p)$ (or $Q^*(p)$). This can be done in principle through the Mellin[6]-Bromwich[7] inversion formula

$$\mathrm{f}(t) = \int_\Delta \frac{\mathrm{f}^*(p)}{p} \exp(pt)\,\mathrm{d}p \tag{5.27}$$

where the integration is done in the complex plane along the vertical line Δ, from the bottom to the top, leaving at left all singularities of $\mathrm{f}^*(p)/p$. In practice, computing this complex integral can be done by complementing Δ by an adequate path in order to get a closed loop and by using the Cauchy[8] residue theorem. Fortunately, simpler algebraic methods can often be used from the knowledge of several Carson-transformed basic functions, as listed in Table 5.1. Note in addition these two important mathematical properties which are especially useful when studying limit responses:

$$\begin{cases}
\mathrm{f}(t)_{t\to\infty} = \mathrm{f}^*(p)_{p\to 0} \\
\mathrm{f}(t)_{t\to 0} = \mathrm{f}^*(p)_{|p|\to\infty}
\end{cases} \tag{5.28}$$

[6] Hjalmar Mellin (1854–1933) was a Finnish mathematician.

[7] Thomas John l'Anson Bromwich (1875–1929) was an English mathematician.

[8] Augustin-Louis Cauchy (1789–1857) was a French mathematician.

Table 5.1 Some basic Carson-transformed functions

Original function f(t)	Carson-transformed function f*(p)
H(t)	1
$\delta(t)$	p
$\exp(-\lambda t)$	$p/(p+\lambda)$
$(1-\exp(-\lambda t))/\lambda$	$1/(p+\lambda)$
t	$1/p$
f(t−h)	$\exp(-hp)\,f^*(p)$
f(bt)exp(−at)	$p/(p+a)f^*[(p+a)/b]$

5.1.3.3 Special Case: Differential Constitutive Equations

The general expression for the constitutive equations of viscoelastic bodies is known to be a *functional relation* between stresses and strains. For the present 1-D description, the general relation (1.23) of Chap. 1 reduces to

$$Q(t) = f\left[q(\tau)_{-\infty}^{t}\right] \text{ or } q(t) = g\left[Q(\tau)_{-\infty}^{t}\right] \tag{5.29}$$

For linear viscoelasticity, the functionals f and g must be linear. As shown above by use of the Boltzmann superposition equation (Eqs. 5.10, 5.11, 5.12), the linearity of f and g implies that they can be represented as *time integrals*. When, in addition, non-ageing is assumed, these time integrals reduce to *Stieljes convolution products* (Eqs. 5.14 and 5.15). As a special case, these convolution products may degenerate into *differential equations of finite order with constant coefficients*, say

$$\sum_i a_i \frac{\mathrm{d}^i}{\mathrm{d}t^i} Q = \sum_j b_j \frac{\mathrm{d}^j}{\mathrm{d}t^j} q \tag{5.30}$$

This differential equation is converted by the Carson transform into

$$Q^*(p) \sum_i a_i\, p^i = q^*(p) \sum_j b_j\, p^j \tag{5.31}$$

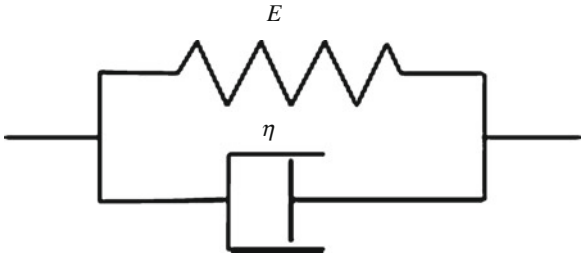
Consequently, the Carson-transformed creep and relaxation functions

$$\mathrm{f}^*(p) = \frac{\sum_i a_i\, p^i}{\sum_j b_j\, p^j} \qquad \mathrm{r}^*(p) = \frac{\sum_j b_j\, p^j}{\sum_i a_i\, p^i} \tag{5.32}$$

are rational fractions with respect to p which can be decomposed into simple elements. According to Table 5.1, the original of every simple element is an exponential function of the form $\exp(-\lambda t)$ or $(1-\exp(-\lambda t))/\lambda$, so that the creep and relaxation functions read as series of such exponential functions (they are named

Fig. 5.6 The Maxwell model

Fig. 5.7 The Kelvin model



"Prony[9] series"). This is easily checked on first-order differential equations which correspond to the classical Maxwell[10] and Kelvin[11] models.

The Maxwell model: it is imaged (Fig. 5.6) by a series assemblage of a linear spring (stiffness E) and a linear dashpot (viscosity η), whose deformation rates must be added to give the total deformation rate $\dot{q}$.

The differential equation reads

$$\dot{q} = \frac{\dot{Q}}{E} + \frac{Q}{\eta} \quad \Rightarrow \quad pq^* = \left(\frac{p}{E} + \frac{1}{\eta}\right) Q^* \tag{5.33}$$

Consequently, the relaxation and creep functions are

$$\begin{cases} \mathrm{r}^*(p) = \dfrac{E\eta p}{E + \eta p} \quad \Rightarrow \quad \mathrm{r}(t) = E\exp(-\frac{E}{\eta}t) \\ \mathrm{f}^*(p) = \dfrac{1}{E} + \dfrac{1}{\eta p} \quad \Rightarrow \quad \mathrm{f}(t) = \dfrac{1}{E} + \dfrac{t}{\eta} \end{cases} \tag{5.34}$$

We can check on these expressions that the Maxwell model exhibits instantaneous elasticity, indefinite creep, complete relaxation, no strain-recovery (except for the instantaneous elastic strain) and complete stress-recovery.

The Kelvin model: it is imaged (Fig. 5.7) by a parallel assemblage of a linear spring (stiffness E) and a linear dashpot (viscosity η), whose forces must be added to give the total force Q.

---

[9] Gaspard Clair François Marie Riche de Prony (1755–1839) was a French mathematician and engineer.

[10] James Clerk Maxwell (1831–1879) was a Scottish physicist and mathematician.

[11] William Thomson, 1st Baron Kelvin (1824–1907), was a British mathematical physicist and engineer.

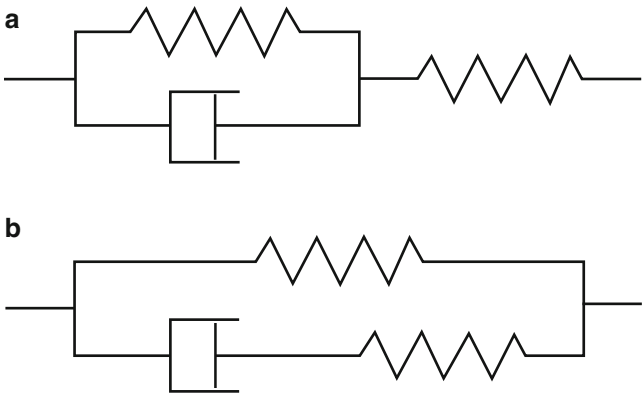


Fig. 5.8 The equivalent Kelvin-Voigt (**a**) and Zener (**b**) models

The differential equation reads

$$Q = Eq + \eta\dot{q} \quad \Rightarrow \quad Q^* = (E + \eta p)\, q^* \tag{5.35}$$

Consequently, the relaxation and creep functions are

$$\begin{cases} \mathrm{r}^*(p) = E + \eta p & \Rightarrow \quad \mathrm{r}(t) = E + \eta\delta(t) \\ \mathrm{f}^*(p) = \dfrac{1}{E + \eta p} & \Rightarrow \quad \mathrm{f}(t) = \dfrac{1}{E}\left(1 - \exp(-\frac{E}{\eta}t)\right) \end{cases} \tag{5.36}$$

We can check on these expressions that the Kelvin model exhibits no instantaneous elasticity, limited creep and complete strain-recovery. Since there is no instantaneous elasticity, the relaxation and stress-recovery tests cannot be performed (see the Dirac $\delta(t)$ singularity in the relaxation function).

Some other elementary models: This last difficulty with regard to the relaxation test can be overcome by adding a spring in series to the Kelvin model: the corresponding Kelvin-Voigt[12] model, which is equivalent to the Zener[13] model (Fig. 5.8), can be considered as a *standard linear solid model* since it does not flow indefinitely at constant stress. On the other hand, the Maxwell model, which is a *standard linear fluid model*, may be enriched by putting a Kelvin model in series with it (Fig. 5.9): this is the so-called Burgers[14] model.

[12] Woldemar Voigt (1850–1919) was a German physicist.

[13] Clarence Melvin Zener (1905–1993) was an American physicist.

[14] Johannes (Jan) Martinus Burgers (1895–1981) was a Dutch physicist (see also Chap. 3).

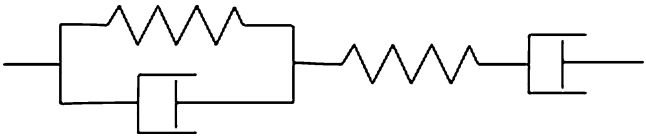


Fig. 5.9 The Burgers model

5.1.3.4 Spectral Representation

Up to now, we have considered systems which are defined by only one pair of external parameters (q, Q). This treatment can be easily extended to systems defined by a finite number n of such pairs, say (q_i, Q_i), such that the incremental work of the external forces for a variation $\mathrm{d}q_i$ of the geometrical parameters is $\sum_{i=1}^{n} Q_i \mathrm{d}q_i$. Like for one single pair of parameters (q, Q), the superposition principle and the assumption of a non-ageing behaviour lead easily, instead of (5.13), (5.14) and (5.15), to the constitutive equations

$$\begin{cases} q_i(t) = \dfrac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t \mathrm{f}_{ij}(t-\tau) Q_j(\tau)\mathrm{d}\tau\right] = \mathrm{f}_{ij} \odot Q_j \\ Q_i(t) = \dfrac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t \mathrm{r}_{ij}(t-\tau) q_j(\tau)\mathrm{d}\tau\right] = \mathrm{r}_{ij} \odot q_j \end{cases} \tag{5.37}$$

Application of the Carson transform to these equations leads to

$$\begin{cases} q_i^*(p) = \mathrm{f}_{ij}^*(p) Q_j^*(p) \\ Q_i^*(p) = \mathrm{r}_{ij}^*(p) q_j^*(p) \end{cases} \tag{5.38}$$

Thermodynamic arguments lead to show that the creep (or retardation) and relaxation symbolic matrices $[\mathrm{f}_{ij}^*(p)]$ and $[\mathrm{r}_{ij}^*(p)]$, which are inverse from one another, are symmetric. In the special case where, among the parameters q_i or Q_i, only one is an external ("observable") parameter whereas the others are internal ("hidden") variables, the same thermodynamic analysis shows that the Carson-transformed creep and relaxation functions read necessarily

$$\begin{cases} \mathrm{f}^*(p) = J_0 + \frac{1}{\eta_\infty p} + \sum_{s=1}^{k} \dfrac{J_s \lambda_s}{p + \lambda_s} \\ \mathrm{r}^*(p) = E_\infty + \eta_0 p + \sum_{s=1}^{k'} \dfrac{E_s p}{p + \mu_s} \end{cases} \tag{5.39}$$

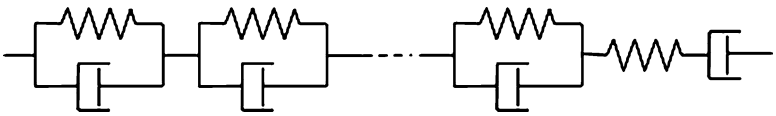


Fig. 5.10 Representation by a series assemblage

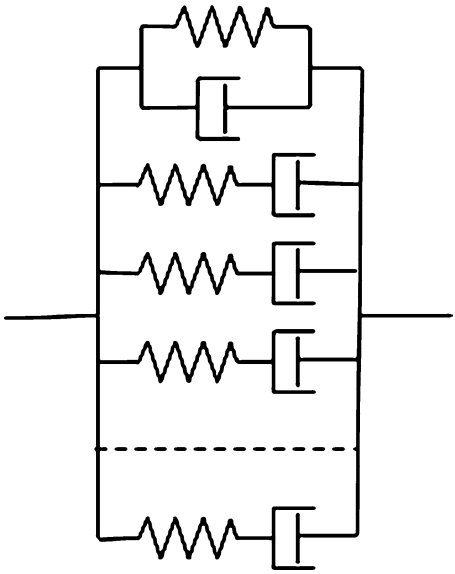


Fig. 5.11 Equivalent representation by a parallel assemblage

with $J_s > 0,\ \eta_\infty > 0,\ J_0 \geq 0,\ \lambda_s \geq 0,\ E_s > 0,\ E_\infty \geq 0,\ \eta_0 \geq 0,\ \mu_s \geq 0,\ k < n,\ k' < n.$ Consequently, the creep and relaxation functions must read

$$\begin{cases} \mathrm{f}(t) = J_0 + \frac{t}{\eta_\infty} + \sum\limits_{s=1}^{k} J_s \left(1 - \exp(-\lambda_s t)\right) \\ \mathrm{r}(t) = E_\infty + \eta_0 \delta(t) + \sum\limits_{s=1}^{k'} E_s \exp(-\mu_s t) \end{cases} \tag{5.40}$$

This means that any linear non-ageing viscoelastic system defined by a finite number of internal parameters can be represented by *a series assemblage of Kelvin models and one (at most) Maxwell model* (Fig. 5.10) or, equivalently, by *a parallel assemblage of Maxwell models and one (at most) Kelvin model* (Fig. 5.11).

In the series assemblage, the stiffness coefficients are $E_s = 1/J_s, E_0 = 1/J_0$ and the viscosity constants are $\eta_s = 1/(\lambda_s J_s)$ and η_∞. Every Kelvin model may be characterised by the compliance J_s and the "retardation time" $\theta_s = 1/\lambda_s$ and the whole set of the retardation times associated with the corresponding

compliances constitutes the *retardation spectrum*. This spectrum has a finite number of discrete lines. Similarly, in the parallel assemblage, every Maxwell model may be characterised by the moduli E_s and the "relaxation time" $\tau_s = 1/\mu_s$. The corresponding *relaxation spectrum* (E_s, τ_s) consists also of a finite number of discrete lines.

There is no difficulty to pass from that to an infinite number of discrete lines and then to a continuous spectrum. For the retardation spectral representation, let $\mathrm{j}(\theta)$ be the sum of the elementary compliances for retardation times lying between θ and $(\theta + \mathrm{d}\theta)$. The global creep function will read

$$\mathrm{f}(t) = J_0 + \frac{t}{\eta_\infty} + \int_0^\infty \mathrm{j}(\theta)(1 - \exp(-t/\theta))\mathrm{d}\theta \tag{5.41}$$

whereas, with $\mathrm{g}(\tau)$ the sum of the elementary moduli for relaxation times lying between τ and $(\tau + \mathrm{d}\tau)$, the global relaxation function is

$$\mathrm{r}(t) = E_\infty + \eta_0\delta(t) + \int_0^\infty \mathrm{g}(\tau)\exp(-t/\tau)\mathrm{d}\tau \tag{5.42}$$

Note that the retardation and relaxation spectra $\mathrm{j}(\theta)$ and $\mathrm{g}(\tau)$ cannot be obtained directly from experiment. Nevertheless, formulae have been proposed to connect them in an approximate manner with the time derivatives of the creep and relaxation functions, respectively.

5.1.3.5 "Dynamic" Investigation

The so-called "dynamic" investigation, which actually is a quasi-static one, refers to forced oscillation tests where q or Q is imposed as a harmonic function of time. After a while, the transitory regime has faded enough for the response to be harmonic too. The comparison of this steady-state response with the applied load yields rich information on the viscoelastic behaviour of the investigated material. Let us consider for example an imposed sinusoidal q-history which we write

$$q(t) = q_0 \cos \omega t = q_0 \mathrm{Re}\,(\exp(\mathrm{i}\omega t)) = \mathrm{Re}\,(\hat{q}(t)) \tag{5.43}$$

where $\mathrm{i}^2 = -1$ and $\mathrm{Re}\,(\hat{q}(t))$ is the real part of the complex function $\hat{q}(t)$ given by

$$\hat{q}(t) = q_0 \exp(\mathrm{i}\omega t) = q_0(\cos \omega t + \mathrm{i} \sin \omega t) \tag{5.44}$$

The Q-response in the steady-state regime can be obtained by application of the superposition principle with the time origin at $\tau = -\infty$:

$$\begin{aligned} Q(t) &= q(-\infty)\mathrm{r}(\infty) + \int_{-\infty}^{t} \dot{q}(\tau)\mathrm{r}(t-\tau)\mathrm{d}\tau \\ &= q(-\infty)\mathrm{r}(\infty) + \int_{-\infty}^{t} q_0 \mathrm{Re}(\mathrm{i}\omega \exp(\mathrm{i}\omega\tau))\mathrm{r}(t-\tau)\mathrm{d}\tau \end{aligned} \tag{5.45}$$

The initial value $q(-\infty)$ can be considered as zero; so we have, with $u = t - \tau$:

$$\begin{aligned} Q(t) &= -\int_{\infty}^{0} q_0 \mathrm{Re}(\mathrm{i}\omega \exp(\mathrm{i}\omega(t-u)))\mathrm{r}(u)\mathrm{d}u \\ &= q_0 \mathrm{Re}\left\{ \exp(\mathrm{i}\omega t)\mathrm{i}\omega \int_0^{\infty} \exp(-\mathrm{i}\omega u)\mathrm{r}(u)\mathrm{d}u \right\} \end{aligned} \tag{5.46}$$

Remember now that the Carson-transformed relaxation function reads

$$\mathrm{r}^*(p) = p \int_0^{\infty} \exp(-pt)\mathrm{r}(t)\mathrm{d}t \tag{5.47}$$

So, we can recognise within the term between braces of (5.46) the value of $\mathrm{r}^*(p)$ for $p = \mathrm{i}\omega$, *i.e.*, $\mathrm{r}^*(\mathrm{i}\omega)$. Finally, we get

$$Q(t) = q_0 \mathrm{Re}\left\{ \exp(\mathrm{i}\omega t)\mathrm{r}^*(\mathrm{i}\omega) \right\} = \mathrm{Re}\left\{ \hat{q}(t)\mathrm{r}^*(\mathrm{i}\omega) \right\} \tag{5.48}$$

or

$$\hat{Q}(t) = \hat{q}(t)\mathrm{r}^*(\mathrm{i}\omega), \quad \text{with } Q(t) = \mathrm{Re}(\hat{Q}(t)) \tag{5.49}$$

This result shows that the steady-state Q-response is sinusoidal too, with the same frequency as $q(t)$ but with some phase lag $\varphi(\omega)$ and an amplitude $Q_0(\omega)$ which depend on the frequency. If the complex function $\mathrm{r}^*(\mathrm{i}\omega)$ is written in the form

$$\mathrm{r}^*(i\,\omega) = R(\omega)\exp(\mathrm{i}\delta(\omega)), \quad R(\omega) = |\mathrm{r}^*(\mathrm{i}\omega)| \tag{5.50}$$

it is easy to see that the phase lag $\varphi(\omega)$ reduces to $\delta(\omega)$, the argument of $\mathrm{r}^*(\mathrm{i}\omega)$, and that the amplitude $Q_0(\omega)$ is equal to $q_0 R(\omega)$, where $R(\omega)$ is the modulus of $\mathrm{r}^*(\mathrm{i}\omega)$. So, $\mathrm{r}^*(\mathrm{i}\omega)$ gives all information on the steady-state Q-response: it is named the "*complex modulus*" and often written as

$$\mathrm{r}^*(\mathrm{i}\omega) = G'(\omega) + \mathrm{i}G''(\omega) \tag{5.51}$$

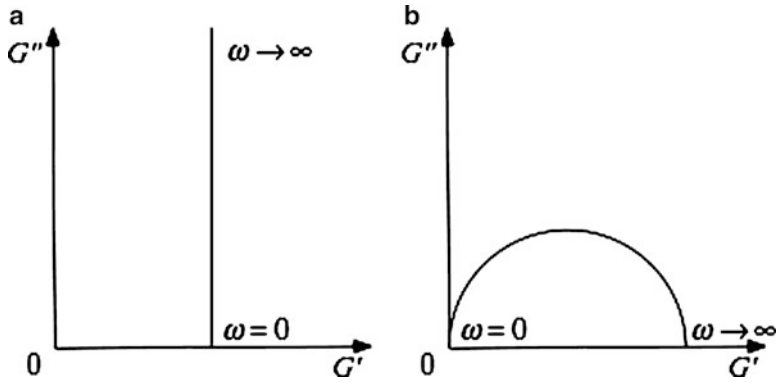


Fig. 5.12 Cole-Cole diagram of (**a**) the Kelvin model; (**b**) the Maxwell model

$G'(\omega)$ is called the *storage modulus* and $G''(\omega)$ the *loss modulus*. The phase angle $\delta(\omega)$ is the *loss angle* and its tangent is the *loss tangent* or the *damping factor*:

$$\tan\delta = G''/G' \tag{5.52}$$

All these terms originate in the *hysteretic* analysis of one cycle in the steady-state regime: in a (Q, q) plot, this cycle is *elliptic*. One part of the mechanical energy is stored (and then restituted) and the other part, ΔW_d, is dissipated (or lost) as heat. The *internal friction coefficient* ρ is defined as the ratio of the dissipated energy to twice the average elastic energy per cycle W_e

$$\rho = \frac{\Delta W_\mathrm{d}}{2W_\mathrm{e}} = \frac{\pi q_0 Q_0 \sin\delta}{(1/2) q_0 Q_0 \cos\delta} = 2\pi\tan\delta \tag{5.53}$$

The complex modulus is conveniently represented by two plots: the Cole[15] and Cole diagram (G', G'') and the $(\tan\delta, \omega)$ plot. It is easy to see that the Cole and Cole diagram for a Kelvin model is a vertical half straight line whereas it is a half-circle for a Maxwell model (Fig. 5.12) since we have

$$\text{Kelvin model:}\quad \mathrm{r}^*(\mathrm{i}\omega) = E + \mathrm{i}\eta\omega \quad\Rightarrow\quad G' = E,\ G'' = \eta\omega$$

$$\text{Maxwell model:}\quad \mathrm{r}^*(\mathrm{i}\omega) = \frac{\mathrm{i}E\eta\omega}{E + \mathrm{i}\eta\omega} \quad\Rightarrow\quad \begin{cases} G' = \dfrac{E\eta^2\omega^2}{E^2 + \eta^2\omega^2} \\ G'' = \dfrac{E^2\eta\omega}{E^2 + \eta^2\omega^2} \end{cases} \tag{5.54}$$

[15]Kenneth Stewart Cole (1900–1984) was an American biophysicist. Robert H. Cole was his younger brother.

More general diagrams and their dependence on temperature are shown in Sect. 5.2.1.

The connexion between the complex modulus and the relaxation function can be used to complement a given set of experimental data by computation in the real time space without recourse to the Carson transform. As a matter of fact, we have

$$\mathrm{r}^*(\mathrm{i}\omega) = \mathrm{i}\omega \int_0^\infty \exp(-\mathrm{i}\omega t)\mathrm{r}(t)\mathrm{d}t = \mathrm{i}\omega \int_0^\infty \mathrm{r}(t)\,(\cos\omega t - \mathrm{i}\sin\omega t)\,\mathrm{d}t \tag{5.55}$$

and then, after an analysis of the limit response at $t \to \infty$

$$\begin{cases} G'(\omega) = \mathrm{r}(\infty) + \omega \int_0^\infty [\mathrm{r}(t) - \mathrm{r}(\infty)]\,\sin\omega t\,\mathrm{d}t \\ G''(\omega) = \omega \int_0^\infty [\mathrm{r}(t) - \mathrm{r}(\infty)]\,\cos\omega t\,\mathrm{d}t \end{cases} \tag{5.56}$$

Conversely, when the storage and loss moduli $G'(\omega)$ and $G''(\omega)$ have been determined experimentally, the relaxation function $\mathrm{r}(t)$ can be computed through the Fourier[16] inverse transform

$$\mathrm{r}(t) = \frac{2}{\pi}\int_0^\infty \frac{G'(\omega)}{\omega}\sin\omega t\,\mathrm{d}\omega = \mathrm{r}(\infty) + \frac{2}{\pi}\int_0^\infty \frac{G''(\omega)}{\omega}\cos\omega t\,\mathrm{d}\omega \tag{5.57}$$

Finally, note that all the foregoing developments related to the complex modulus could have been applied with few obvious modifications to the *complex compliance* $\mathrm{f}^*(\mathrm{i}\omega) = 1/\mathrm{r}^*(\mathrm{i}\omega)$.

Applications of this paragraph to polymers will be found in Sect. 5.2.1.5.

5.1.3.6 Criteria of Linearity

As shown above, the properties of linearity and non-ageing are of great use for the study of viscoelasticity. It is then interesting to generate practical methods for the experimental determination of the *domain of linearity* of a given material. We briefly comment some of these methods.

Proportionality criterion: various creep tests are performed on the same material for different loads Q. The creep responses are $q = q(t, Q)$. Linearity implies that $q(t, Q) = Q\mathrm{f}(t)$. For any time t, the relation between q and Q for different values of Q must be linear. On the corresponding plot (q, Q), the domain of linearity is limited by the occurrence of bends on the experimental curves (Fig. 5.13). Of course, the same approach can be applied to relaxation tests at various imposed displacement levels.

[16] Jean Baptiste Joseph Fourier (1768–1830) was a French mathematician.

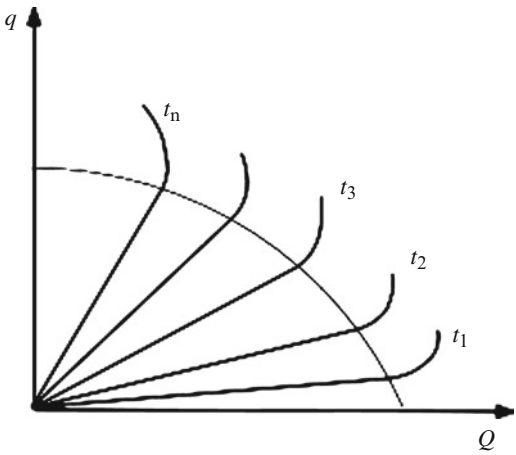


Fig. 5.13 Determination of the domain of linearity from creep tests

Prediction of the recovery response: according to (5.3), the Q-history for a strain-recovery test starting at $t = 0$ with unloading at $t = \mathrm{T}$ is given by:

$$Q(t, \mathrm{T}) = Q_0 \left[\mathrm{H}(t) - \mathrm{H}(t - \mathrm{T})\right] \tag{5.58}$$

Linearity leads to (5.8) and non-ageing implies the response to be

$$q(t) = Q_0 \left[\mathrm{f}(t)\mathrm{H}(t) - \mathrm{f}(t - \mathrm{T})\mathrm{H}(t - \mathrm{T})\right] \tag{5.59}$$

Consequently, if, on the plot (q, Q), from $t = \mathrm{T}$ up to $t = 2\mathrm{T}$, we add $q(t)$ and $q(t-\mathrm{T})$, we must get points which lie on the continued creep curve between T and 2T (Fig. 5.14). Thus, linearity can be easily checked in this way. The same treatment can work for a stress-recovery test.

Harmonic tests : as we know, a (Q, q) plot during such a test must be and stay the same ellipse at any cycle for linear viscoelastic non-ageing materials. This property could be used in principle as a linearity criterion. Nevertheless, it is practically rather difficult to be sure that a closed curve is an ellipse and that it does not evolve from one cycle to another. That is why static tests are preferred for checking linearity.

5.2 Microstructural Aspects and Physical Mechanisms

5.2.1 Viscoelasticity of Polymers and Related Phenomena

Before studying viscoelasticity of polymers as such, we need to look at their elastic properties as they derive from their structures, described in Annex 1. We are then led to some repetitions for easier comprehension.

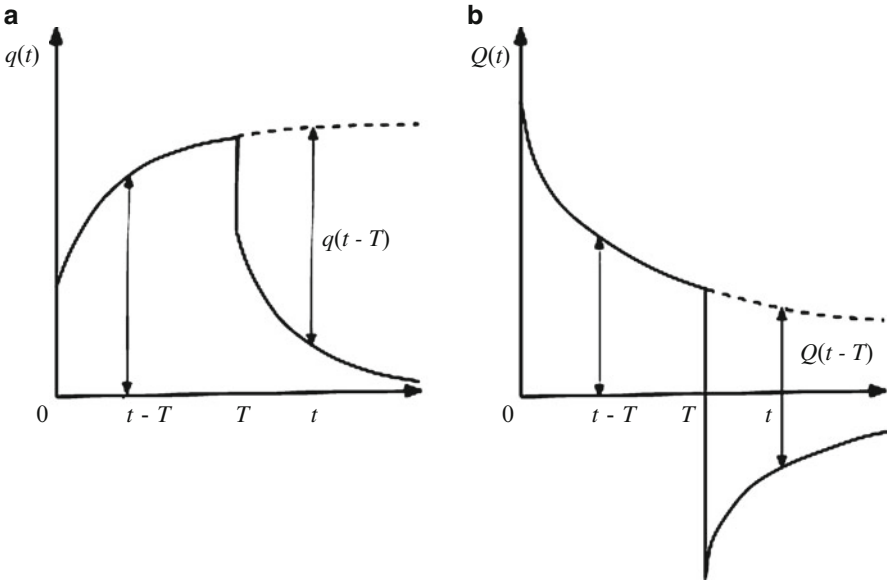


Fig. 5.14 Linearity check by a recovery test: (**a**) strain-recovery; (**b**) stress-recovery

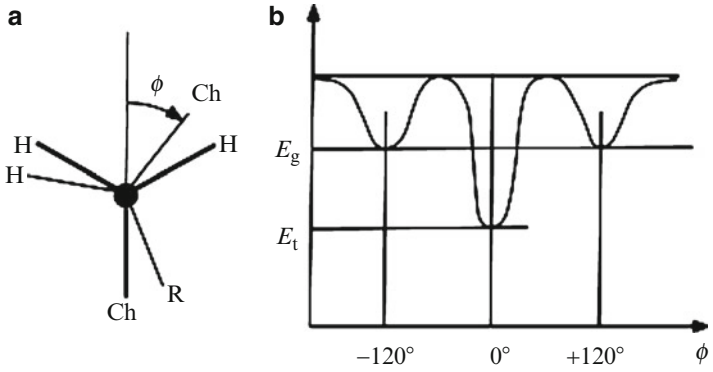


Fig. 5.15 Conformation of a carbon-carbon bond in a vinyl polymer. (**a**) Newman's representation (**b**) shape of the variation of the potential energy with the angle ϕ

5.2.1.1 Chain Conformation in Amorphous Polymers

Chain conformation in amorphous polymers is described in Annex 1. It is shown that the potential energy variation as a function of the angle of rotation of carbon atoms with respect to one another has minima for some orientations. To make a complete rotation the molecules must jump over the highest potential barrier *e.g.* E_T in Fig. 5.15.

The probability for a jump is given by:

$$P_D = \exp\left(-\frac{E_\mathrm{T}}{\mathrm{k}T}\right) \tag{5.60}$$

P_D is called the *dynamic stiffness* of the molecule.

Trans and gauche configurations can coexist at a given temperature. Their population ratio is directly linked to the *static stiffness* factor of the molecule defined as:

$$P_S = \exp\left(-\frac{E_\mathrm{T} - E_\mathrm{G}}{\mathrm{k}T}\right) \tag{5.61}$$

If the chain is long enough relative to the persistence length (see Annex 1), the average of a large number of chain configurations corresponding to a trans/gauche ratio is a random coil, which can be characterised by its mean end-to-end distance $<r>$, which is a decreasing function of the chain tortuosity. This latter is characterised by the chain characteristic ratio C_∞ for a chain having N bonds of length l (Flory 1953):

$$C_\infty = \lim_{N\to\infty}\left(\frac{< r^2 >}{Nl^2}\right) \tag{5.62}$$

For totally free rotations, $C_\infty = 1$, where C_∞ expresses the constraints (geometric linked to valence angles and energetic linked to energetic barriers) opposed to free rotations. For industrial polymers, generally C_∞ ranges from about 2 (aromatic polymers of the polycarbonate type) to about 10 (polymers with bulky side groups of the polystyrene type). There are important relationships between certain thermomechanical properties and C_∞ (see later). It will be just remarked here that, if for a long chain in the undeformed state, the end-to-end distance is:

$$< r >_0 = \left(l{C_\infty}^{1/2}\right) N^{1/2} \tag{5.63}$$

for the fully extended chain it is:

$$< r >_f = gN \tag{5.64}$$

where g is a factor linked to the geometry of the monomer unit. The maximum extension ratio of a chain is thus:

$$\Lambda_\mathrm{max} = \frac{< r >_\mathrm{f}}{< r >_0} \propto N^{1/2} \tag{5.65}$$

In rubbers, if the elastically active chains have a molar mass M_e and a number of bonds N_e (proportional to M_e), the maximum elongation is proportional to ${N_\mathrm{e}}^{1/2}$

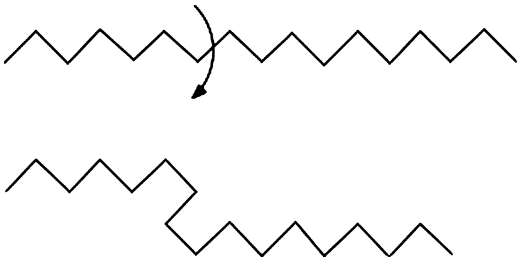


Fig. 5.16 Schematisation of conformational changes in an initially all trans segment

(Flory1953; Treloar1975; Sperling1992) whereas the shear modulus G (also named μ elsewhere in this book) is proportional to ${M_e}^{-1}$, *i.e.,* to ${N_e}^{-1}$. We expect thus that:

$$G \propto {\Lambda_{max}}^{-2} \tag{5.66}$$

The maximum elongation of rubber is a decreasing function of its stiffness. In a linear polymer the maximum extensibility is mainly linked to the chain length between entanglements.

A chain can undergo conformational changes caused by thermal agitation or mechanical loading. A trans → gauche local change can be schematised by Fig. 5.16.

5.2.1.2 Temperature Effect on the Conformational State. Transitions

Let us consider the case, schematised in Fig. 5.15. where the gauche conformer has the highest potential energy, *i.e.,* is the less stable. Since the population ratio gauche/trans is linked to the factor $\exp\left(-\frac{E_G - E_T}{kT}\right)$, it is expected to increase with temperature. The polymer entropy S, which is directly linked to the conformational state, also increases with temperature. Let us now consider the change of the molar fraction g of gauche conformations during cooling at a constant rate θ started at a temperature T_1 in the liquid state (Fig. 5.17). During cooling, when the temperature varies between T and $T - dT$ the mole fraction g of gauche conformatiofn is expected to vary from g to $(g - dg)$. Conformational changes are however not instantaneous; they are characterised by a rate which is an increasing function of temperature. At high temperature, rotations are fast and the polymer adopts almost instantaneously its equilibrium conformation. Below a certain temperature, however, rotations become too slow relatively to the time scale fixed by the cooling rate. Then, the chains leave their thermodynamic equilibrium and tend rapidly towards a "frozen" conformational state. This transition is called *glass transition* and the state out of equilibrium is called *glassy state.* This process calls for the following remarks:

1. The glass transition displays many characteristics of a secondary thermodynamical transition. It is characterised by jumps of the volume expansibility $\alpha = \alpha_l - \alpha_g$, heat capacity $\Delta C_p = C_{pl} - C_{pg}$ and compressibility $\Delta k = k_l - k_g$.

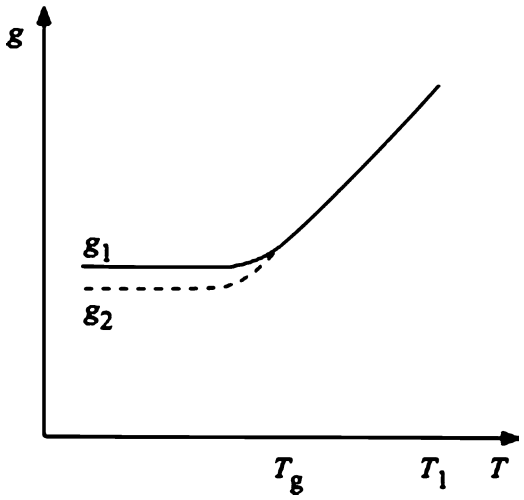


Fig. 5.17 Schematisation of the change of the molar fraction of gauche conformers during a cooling experiment at rate θ_1 (*full line*) and $\theta_2 < \theta_1$ (*dashed line*)

For a true thermodynamical transition, we expect:

$$T_g\left(\Delta\alpha^2\right) = \Delta C_p \Delta k \tag{5.67}$$

Polymers can deviate from this law because glass transition occurs between an equilibrium (liquid/rubbery) and a non-equilibrium (glass) state, whereas in a true thermodynamical transition, the material is expected to be in equilibrium on both sides of the transition.

2. Owing to the kinetic character of the glass transition, its coordinates depend on the cooling rate θ. The lower θ, the lower T_g and the lower the gauche molar fraction in glassy state (Fig. 5.17). In the range of experimentally available cooling rates, however, T_g cannot vary by more than 10–20°C.
3. At T_g the molecular mobility, as characterised for instance by viscosity or NMR relaxation time, varies by several orders of magnitude. There is a great diversity of physical methods giving access to quantities more or less sharply linked to molecular mobility, for instance ESR (Electron Spin Resonance), NMR (Nuclear Magnetic Resonance), photophysical methods, etc.

These methods allow putting in evidence four important temperature domains separated by three more or less fuzzy transitions (Fig. 5.18).

There are eventually several types of local motions, each one characterised by the temperature at which it begins to be significant. These *sub-glass transition temperatures* are called T_β, T_γ, T_δ, etc., in order of decreasing temperature, starting at T_g (T_α). Cooperative motions are also named α motions. Local motions are named β, γ, δ, etc. In the following, only the first sub-glass transition T_β will be considered for the sake of simplicity.

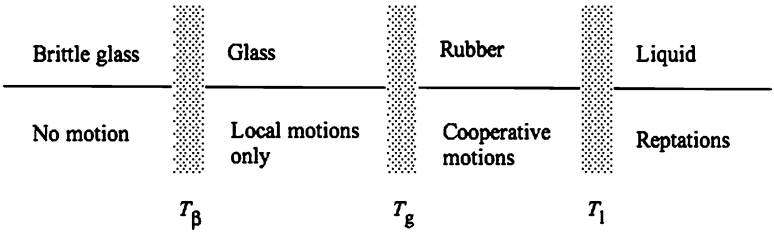


Fig. 5.18 Temperature domains of the main distinct mechanical behaviours of polymers and corresponding transitions. N.B. The transition temperatures depend more or less on the timescale in which they are observed

5.2.1.3 Time-Temperature Superposition

Let us consider for instance a relaxation experiment in which the strain $\varepsilon = \varepsilon_0$ is fixed; the stress $\sigma(t)$ is recorded and variations of the modulus $E(t) = \sigma(t)/\varepsilon_0$ are studied. Experiments of this type are difficult to perform in very short time scales, for instance < 0.01 h, owing to the key importance of the loading history. They are also difficult to perform in very long time scales, for instance $>10^4$ h, for obvious economical reasons. This is the reason why the application of a time-temperature superposition principle appeared soon as the key research objective in the field of polymer mechanics. Such a principle would be especially interesting if it was expressed by the following equality (Ferry 1970):

$$P\,(t, T) = P\left(\frac{t}{a_T}, T_R\right) \tag{5.68}$$

where T_R is an arbitrary reference temperature and a_T a shift factor depending only on temperature. As a matter of fact, using such a principle, it is possible to build a *master curve* by a simple translation along a logarithmic scale, as shown in Fig. 5.19.

Indeed, if a mathematical function $a_T = f(T)$ is found, it must keep an invariant form with the choice of the reference temperature. Williams et al. (1955) found for the first time such a relationship, which is now called the WLF law:

$$\log a_T = \frac{-C_1\,(T - T_R)}{C_2 + (T - T_R)} \tag{5.69}$$

The following remarks and observations can be made:

1. This law is not equivalent to the well known Arrhenius law. In other words the temperature effect on relaxation kinetics does not result only from the fact that the system jumps over a barrier of potential.
2. $a_T \to \infty$ when $T \to T_\infty = T_R - C_2$. It can be written: $\log a_T = -C_1 + C_1 C_2/(T - T_\infty)$, *i.e.,* $a_T = B_V \exp[C_V / (T - T_\infty)]$. This law was previously known as the Vogel's law.

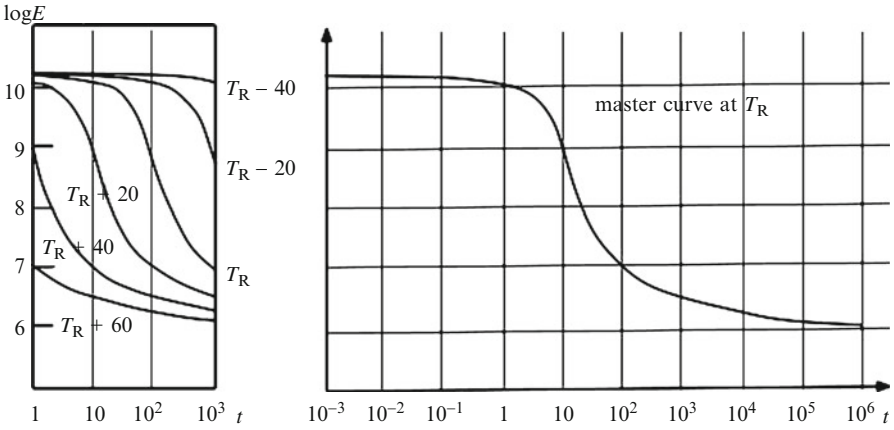


Fig. 5.19 Schematisation of the elaboration of a master curve at T_R from experiments made over three time decades, at temperatures ranging from $T_R - 40°C$ to $T_R + 60°C$

3. It was remarked that if the reference temperature is chosen as equal to T_g, the corresponding parameters: $C_{1g} \approx 17$ and $C_{2g} \approx 50$ K are quasi universal.
4. It is easy to show that the WLF law can be derived from the free volume considerations, writing that a relaxation time τ, for instance, is linked to the free volume fraction f by:

$$\tau(T) = \tau_0 \exp \frac{B}{f}$$
$$\tau\left(T_g\right) = \tau_0 \exp \frac{B}{f_g} \tag{5.70}$$

where B is a constant, so that:

$$\log \frac{\tau(T)}{\tau\left(T_g\right)} = -\frac{C_{1g}\left(T - T_g\right)}{C_{2g} + \left(T - T_g\right)} = \log a_T \tag{5.71}$$

where $C_{1g} = B/f_g$ and $C_{2g} = f_g/\alpha$.

The constant B is usually taken equal to unity but this hypothesis lacks a physical justification. Since $C_{2g} \approx 50$ K and $\alpha \approx 5 \cdot 10^{-4}$ K^{-1}, it comes $f_g \approx 0.025$.

In spite of its seductive simplicity, this interpretation does not constitute a rigorous proof in favour of the free volume theory. First the product $C_{1g}C_{2g}\alpha$, which is expected to be equal to unity, varies significantly from one polymer to another. Second, it is possible to reach the WLF equation starting from entropy considerations as established by Gibbs and Di Marzio (1958) who observed that, plotting the chain entropy against temperature, a graph similar to Fig. 5.17. is obtained, in which extrapolation of the straight line corresponding to equilibrium gives a zero entropy at $T_\infty \approx T_g - 50$ K (Fig. 5.20).

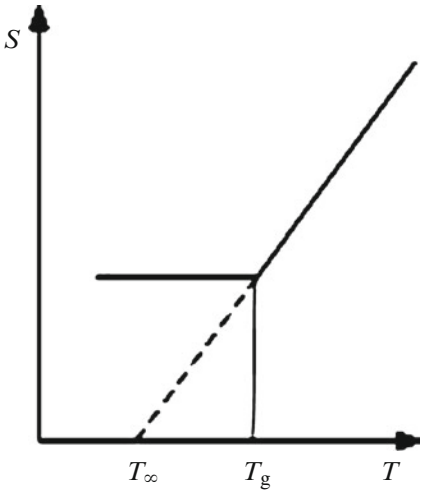


Fig. 5.20 Shape of the chain entropy – temperature plot according to Gibbs and Di Marzio (1958)

5.2.1.4 Molecular Relaxations

Let us consider a polymer in solid state undergoing a stimulus S and a method to record its response R. The stimulus modifies the polymer structure (conformational state, free volume, intermolecular distances). When the stimulus is released, the polymer tends to relax, *i.e.,* to recover equilibrium thanks to molecular motions. The polymer response, initially R_0, tends to the equilibrium value R_0 at a rate which is an increasing function to the distance to equilibrium $(R - R_0)$. In the simplest case:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = -k\,(R - R_0) \tag{5.72}$$

so that:

$$R - R_0 = (R_0 - R_\infty)\exp(-kt) = (R_0 - R_\infty)\exp(-t/\tau) \tag{5.73}$$

τ is called *relaxation time*.

In reality, there is a diversity of situations. For instance segmental motions can be faster near the chain ends than in their middle. Motions are especially hindered in the vicinity of entanglements, etc. Relaxation kinetics are thus characterised by a relaxation spectrum rather than by a unique relaxation time.

It is then usual to represent relaxation kinetics by the Kolrausch-Williams-Watt (KWW) equation:

$$R - R_0 = (R_0 - R_\infty)\exp\left[-(t/\tau)^\beta\right] \tag{5.74}$$

with $0 < \beta < 1$.

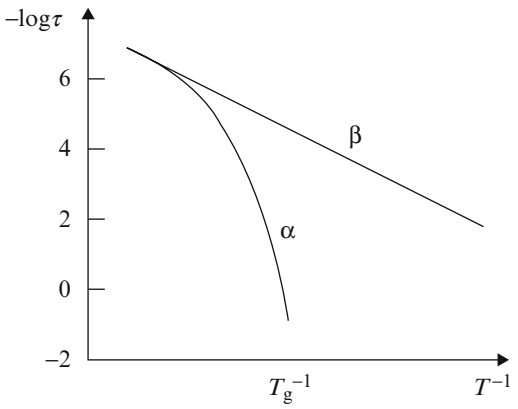


Fig. 5.21 Common shape of relaxation maps for amorphous polymers

Table 5.2 Relaxation characteristics for α and β motions

Transition	β	Temperature dependence of τ	Apparent activation energy (kJ/mol)
α (glass)	0.3–1.0	WLF	290–1,000
β (sub-glass)	0.01–0.1	Arrhenius	0–150

Here τ is an average relaxation time and β is a measure of the width of the relaxation spectrum. Arrhenius plots of the relaxation time are called *relaxation maps* (Fig. 5.21). The glass transition T_g and the sub-glass transitions can be distinguished by many characteristics (Table 5.2).

5.2.1.5 Linear Non-ageing Viscoelasticity

Phenomenological Aspects

In the domain of small strains, polymers are assumed to obey the Boltzmann superposition principle according to which, when there are changes of stimulus, the effects of the changes are additive when the corresponding responses are considered at equivalent times (see Sect. 5.1.2.2).

$$R(t) = \sum_{i=0}^{N} Y\,(t - t_i)\,S_i \tag{5.75}$$

If S_i is a stress σ_i applied at time t_i, $R(t)$ is a strain $\varepsilon(t)$ and $Y(t)$ is the time dependent creep compliance $J(t)$. If S_i is a strain ε_i applied at time t_i, $R(t)$ is a stress $\sigma(t)$ and $Y(t)$ is the time dependent relaxation modulus $E(t)$. Note that $J(t)$ and $E(t)$ are proportional to $f(t-t_0)$ and $r(t-t_0)$ defined above in Sect. 5.1.3.1.

Thus applying a programmed stress $\sigma(t)$, the resulting strain $\varepsilon(t)$ is:

$$\varepsilon(t) = \int_{-\infty}^{t} J\,(t-\tau)\,\dot{\sigma}\,(\tau)\,\mathrm{d}\tau \tag{5.76}$$

which is the equivalent of Eq. 5.14. $\dot{\sigma}(t)$ is the loading rate.

In the same way for a programmed strain $\varepsilon(t)$:

$$\sigma(t) = \int_{-\infty}^{t} E\,(t-\tau)\,\dot{\varepsilon}\,(\tau)\,\mathrm{d}\tau \tag{5.77}$$

which is the equivalent of Eq. 5.15.

Dynamic Testing

It is the most common way to establish relaxation maps. A sinusoidal load is applied and the sample response is recorded (see Sect. 5.1.3.5). It is also sinusoidal but with a delay δ linked to the polymer viscosity. For instance, applying $\varepsilon = \varepsilon_0 \sin\omega t$, one obtains $\sigma = \sigma_0 \sin(\omega\ t + \delta)$. Using complex variables $\varepsilon^* = \varepsilon_0 \exp(\mathrm{i}\omega\ t) \rightarrow \sigma^* = \sigma_0 \exp[\mathrm{i}(\omega\ t + \delta)]$. The complex modulus is thus $E^* = E' + \mathrm{i}E''$ (equivalent of Eq. 5.51), where $E' = (\sigma_0/\varepsilon_0)\cos\delta$ (*storage modulus*), and $E'' = (\sigma_0/\varepsilon_0)\sin\delta$ (*loss or dissipation modulus*). $\tan\delta = E''/E'$ is the *damping factor* (Eq. 5.52).

Dynamic testing allows to study the variations of E' and E'' in a more or less large domain of temperatures (typically 150–600 K) and frequencies (typically 10^{-1}–10^3 Hz). The curves display several transitions, each one corresponding to a polymer relaxation of characteristic time τ such that $\omega\tau = 1$.

The shape of E' and E'' variations is shown in Figs. 5.22 and 5.23.

Viscoelastic Spectra in the Frequency Domain

It is usual to transpose the results of Figs. 5.22 and 5.23 into Cole-Cole diagrams $E'' = \mathrm{f}(E')$ (see Sect. 5.1.3.5). There is a wide variety of mathematical models expressing the complex modulus E^* as a function of angular frequency ω. Some models and the corresponding Cole-Cole plots are shown in Fig. 5.24.

Experimental Cole-Cole plots have generally a shape close to the Harriliak-Nagami (1967) one. It seems that, to simulate experimental curves, a model must have at least two exponents. They are in the case of the model proposed by Perez (1992):

$$E^* = E_0 + \frac{E_\infty - E_0}{1 + (\mathrm{i}\omega t)^{\kappa'} + Q\,(\mathrm{i}\omega t)^{\kappa''}} \tag{5.78}$$

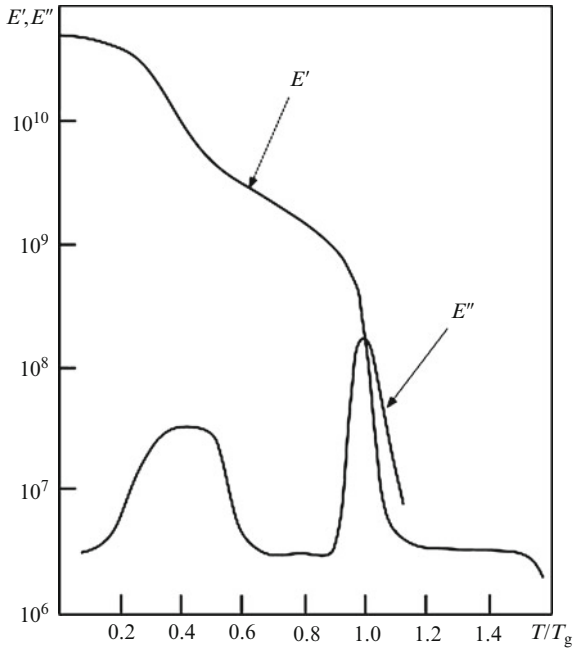


Fig. 5.22 Shape of the temperature variation of the real and imaginary modulus at a fixed frequency. NB: the location and amplitude of the β transition can vary from one polymer to another

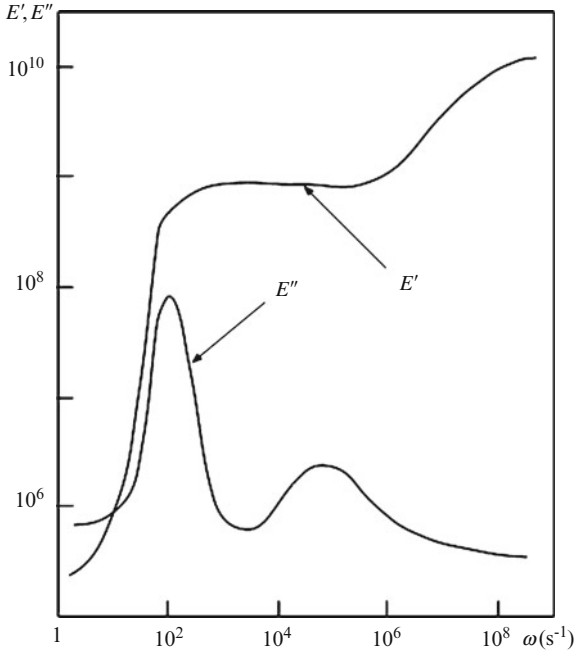


Fig. 5.23 Shape of the frequency variations of E' and E'' at a fixed temperature

Maxwell $E^* = \dfrac{E_\infty}{1 + \left(\mathrm{i}\omega\tau\right)^{-1}}$

Zener $E^* = E_0 + \dfrac{E_\infty - E_0}{1 + \left(\mathrm{i}\omega\tau\right)^{-1}}$

Generalised Maxwell $E^* = \dfrac{E_\infty}{1 + \left(\mathrm{i}\omega\tau\right)^{-\gamma}}$

Harrilak-Negami $E^* = E_\infty - \dfrac{E_\infty - E_0}{\left[1 + \left(\mathrm{i}\omega\tau\right)^{-\alpha}\right]^{-\gamma}}$

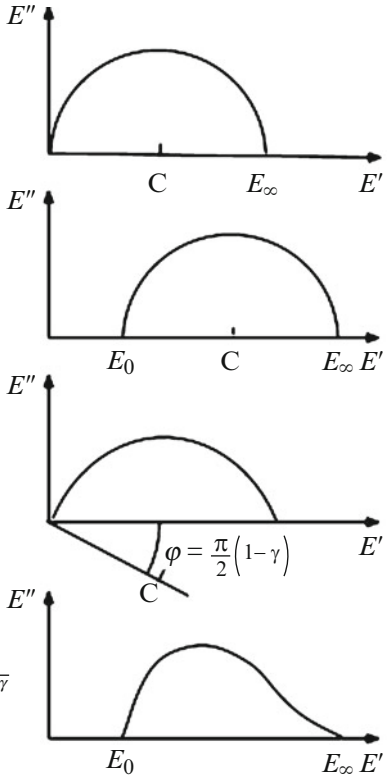


Fig. 5.24 Some models for the frequency dependence of complex moduli and shape of corresponding Cole-Cole plots. NB C is the centre of the circle

According to this author all parameters could be interpreted in terms of structure-properties relationships and the counterpart of the model in the time domain would be the KWW equation.

An example for the modelling of the temperature dependence of modulus is the approach of Gilbert, Ashby and Beaumont (1986) (Fig. 5.25).

The unrelaxed modulus E_u obeys a dilatation law:

$$E_\mathrm{u} = E_0\left(1 - \alpha\frac{T}{T_\mathrm{g}}\right) \tag{5.79}$$

The modulus E at a given temperature T is:

$$E = E_\mathrm{u} - \sum \Delta E_i \tag{5.80}$$

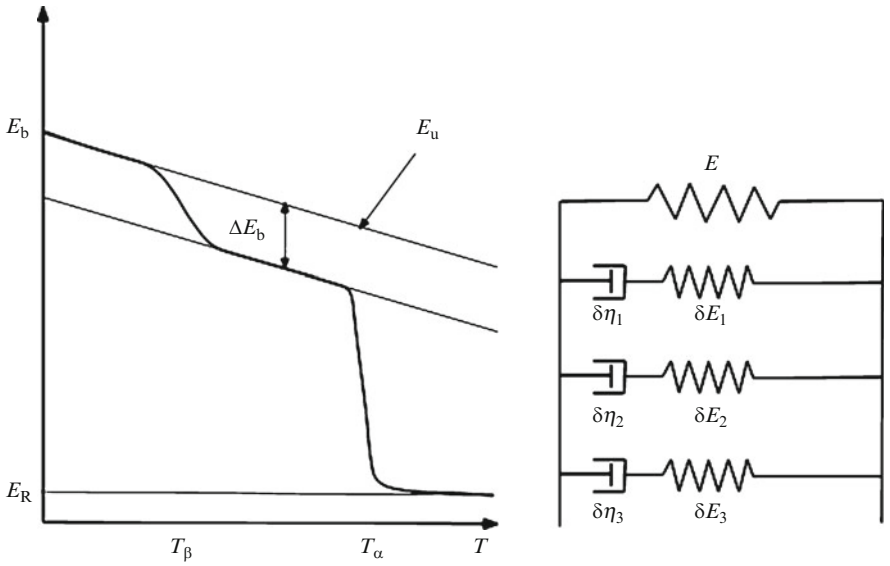


Fig. 5.25 Shape of the temperature variation of modulus (*left*) and chosen rheological model (*right*)

where ΔE_i is the modulus drop corresponding to the i^{th} transition below the temperature T and α is a dimensionless parameter of the order of 0.3.

The rheological model consists of a spring in parallel with an infinity of Maxwell elements differing by the viscosity η of the dashpot. The viscosities obey the Arrhenius law in the glassy domain and the WLF law in the glass transition and rubbery domain:

$$\eta = \eta_0 \exp\left(\frac{Q}{kT}\right) \quad \text{and} \quad \eta = \eta_0 \exp\left[\frac{C_{1g}\left(T - T_g\right)}{C_{2g} + \left(T - T_g\right)}\right] \tag{5.81}$$

It is assumed that viscosities differ only by their activation energy for the Arrhenius equation and by C_1 for the WLF equation. Q and C_1 are assumed to have a Gaussian distribution. For instance the proportion of units with activation energies between Q and $Q + \delta Q$ is f(Q)δQ where:

$$f(Q) = \frac{1}{\sqrt{2\pi}\Delta Q} \exp\left[-\frac{1}{2}\left(\frac{Q - Q_\alpha}{\Delta Q}\right)^2\right] \tag{5.82}$$

The corresponding modulus drop is (for the i^{th} transition):

$$\delta E_i = \Delta E_i f(Q)\delta Q \tag{5.83}$$

Table 5.3 Characteristics of the β transition of PMMA and PS according to Gilbert et al. (1986)

Polymer	PMMA	PS
T_g (K)	378	373
E_0 (GPa)	8.6	6.2
α	0.28	0.30
T_β/T_g	0.75	0.80
Q_β (kJ/mol)	121.3	125
$\Delta E_\beta/E_0$	0.47	0.03
$\Delta Q_\beta/Q_\beta$	0.2	0.4
η_0 (Pa s)	$1.5{\cdot}10^{-15}$	$8.3{\cdot}10^{-17}$

The standard deviation ΔQ is adjusted to fit experimental data. Some parameter values were given by Gilbert et al. (1986) for two polymers poly (methylmethacrylate) and polystyrene (Table 5.3).

For the glass transition the authors take $C_{1g} = 17.4$ and $C_{2g} = 52$ K for both polymers. The fractional spread of C_1 is noticeably lower than for the β transition: $\Delta C_1/C_1 = 0.05$ (PMMA) and 0.08 (PS).

5.2.1.6 Structure-Property Relationships

Glass Transition

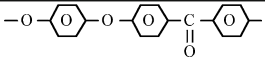
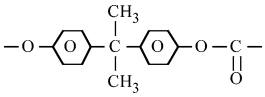
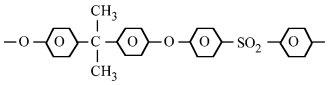
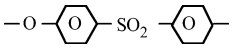
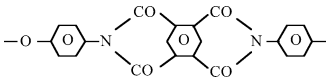
In linear polymers, T_g is an increasing function of molar mass M. This is due to the free volume excess of chain ends. According to the free volume theory (Fox and Flory1954):

$$T_g = T_{g\infty} - \frac{K_{FF}}{M_n} \tag{5.84}$$

where $T_{g\infty}$ is characteristic of the chemical structure and K_{FF} is an increasing function of $T_{g\infty}$. It can be shown empirically that $K_{FF} \approx 10^{-3} {T_{g\infty}}^2$. $T_{g\infty}$ is sharply linked to the chain dynamic stiffness. All the factors hindering segmental rotations tend to increase $T_{g\infty}$. In the case of flexible chain (aliphatic) polymers, the main factor is cohesion. Poly(vinylalcohol) ($T_g \approx 360$ K) can be compared to polypropylene ($T_g \approx 270$ K), a chain of close geometry but considerably less cohesive. In the same way, poly(vinylchloride) ($T_g \approx 353$ K) can be compared to polypentene ($T_g \approx 233$ K). The geometry of side groups plays a complex role. For instance, polystyrene ($T_g \approx 378$ K) appears stiffer than poly(vinylchloride) despite the fact that the latter is more cohesive. In other structural series, for instance polyalkenes or polyalkylmethacrylates, T_g tends to decrease with the bulkiness of side groups.

At high temperatures, typically 400 K, the network of polar interactions collapses and another factor becomes predominant: the inertia of rotatable units in the polymer skeleton. The influence of the "rigid rods" inertia appears clearly in the series of Table 5.4 where the ether groups are the unique "knee joints".

Table 5.4 Glass transition temperature of some aromatic backbone polymers and their molar mass per ether link

Acronym	Structure	T_g (K)	M/ether link (g/mol)
PEEK		414	144
PC		425	127
PSU		463	221
PES		483	232
PPMA		623	382

Here, isopropylene ($-C(CH_3)-$) and sulfone ($-SO_2-$) bridges block the rotations of phenylene groups and do not work as "knee joints".

In cross-linked polymers, T_g is an increasing function of the cross-link density. There are many theories about this influence. The most pertinent one might be Di Marzio's (1964) theory according to which:

$$T_g = \frac{T_{gl}}{1 - K_{DM} F \nu} \tag{5.85}$$

where T_{gl} is the glass transition temperature of a hypothetical linear copolymer containing all the difunctional units of the network; K_{DM} is a universal constant of the order of unity; F is the flex parameter expressing the inertia of the chain units: $F =$ (molar mass of chains/number of rotatable bonds); ν is the cross-link density expressed in moles of elastically active chains per mass unit.

Bellenger et al. (1987), from a study of a large series of epoxy networks (with trifunctional cross-links), showed that $K_{DM} \approx 2$; F varies from about 14 g/mol (polymethylene sequences) to about 60 g/mol (highly aromatic chains). T_g varies in the same way as F and obeys the same structure-properties relationships as for linear polymers. Note that:

$$\frac{dT_g}{d\nu} = K_{DM} F \frac{{T_g}^2}{T_{gl}} \tag{5.86}$$

The sensitivity of T_g to the variations of the cross-links density is an increasing function of T_{gl}. It is very low for networks made of flexible chains; especially rubbers and high for networks made of stiff chains (aromatic thermosets).

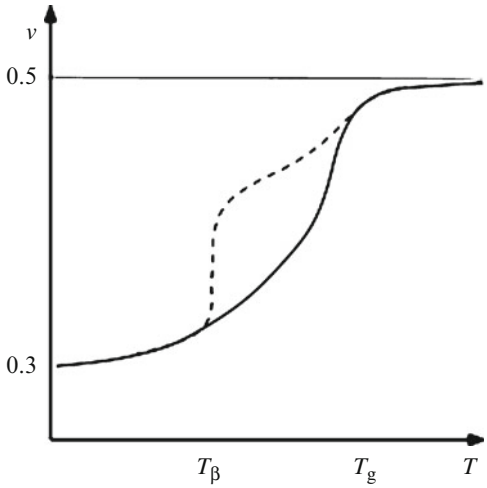


Fig. 5.26 Schematic shape of the temperature dependence of the Poisson ratio for a polymer having a β transition of low activity (*full line*) or of high activity (*dashed line*)

Beta Transition. Engineering Elastic Properties of Glassy State

The relationships between the characteristics of the β transition and the polymer structure are not as clear as for the glass transition. Wu (1992) found empirically that there is an important influence of chain characteristic ratio C_∞(Sect. 5.2.1.1):

$$\frac{T_\beta}{T_g} = 0.135 \pm 0.027 + (0.082 \pm 0.004)\, C_\infty \tag{5.87}$$

In other words, aromatic polymers of the polycarbonate type have a β transition well separated from T_g (more than 200°C in the case of PC), whereas statically stiff polymers as polystyrene have a β transition close to T_g.

The separation between T_g and T_β plays also a great role in other properties, for instance structural relaxation and fracture properties. β transitions are more or less active, *i.e.,* the corresponding modulus drop is more or less strong depending on various factors among which the difference of extension between relaxed and unrelaxed conformers. Polycarbonate, polysulfones, aromatic amine, cross-linked epoxys have a very active β transition that explains why they have a relatively low modulus compared to polymers having the same cohesivity. For instance polycarbonate, polysulfones and some aromatic amine cured epoxies have tensile modulus values at ambient temperature, $10^{-3\pm2}$ s^{-1} tensile rate, of the order of 2.6 ± 0.2 GPa, whereas for polystyrene, despite its lower cohesivity, it is 3.1 ± 0.1 GPa. Let us recall that the bulk modulus K decreases slowly with temperature but is unaffected by secondary (sub-glass) transitions. As a result the Poisson ratio ν varies in the opposite way as the shear and tensile moduli (Fig. 5.26).

For instance at ambient temperature, for a tensile rate of the order of $10^{-3\pm2}$ s^{-1}, the Poisson ratio is of the order of 0.41 ± 0.01 for polymers of the polycarbonate or epoxy (amine cross-linked) type, and of the order of 0.37 ± 0.01 for polystyrene or styrene cross-linked polyester type.

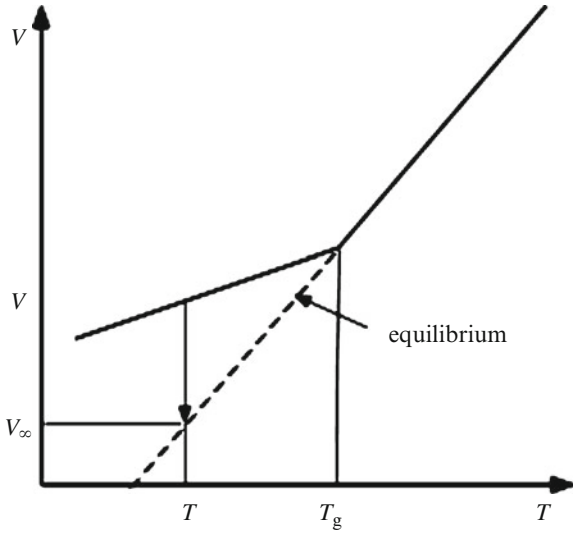


Fig. 5.27 Schematisation of volume variation resulting from structural relaxation at constant temperature below T_g. The arrow indicates the extent of physical ageing. The change is of the order of the 3rd or the 4th decimal of density

5.2.1.7 Structural Relaxation. Physical Ageing

Since polymers are out of equilibrium in the glassy state and since they have a certain residual (β) mobility in a certain temperature interval below T_g, they can evolve towards equilibrium. This structural relaxation can be described as well as a progressive disappearance of the excess of unstable conformations or as a progressive collapse of the free volume excess. In both cases the phenomenon consumes the sources of mobility so that it is self-retarded. In a first approach it could be assumed that the rate of evolution, for instance the rate of volume V decrease, could be proportional to the distance to equilibrium as schematised in Fig. 5.27.

According to the hypothesis of first-order process:

$$\frac{dV}{dt} = -\frac{V - V_\infty}{\tau} \tag{5.88}$$

Kovacs (1963), who discovered the phenomenon in 1963, showed that it (as viscoelasticity) cannot be modelled with a single relaxation time but is characterised by a relaxation spectrum. Struik (1978) studied the structural relaxation using essentially creep tests. Structural relaxation, called here physical ageing, leads to a decrease of mobility and thus a decrease of creep rate. Creep curves for samples differing by their ageing time t_V at a given temperature have the shape shown in Fig. 5.28.

Struik found that in a log (compliance) *vs* log (time of creep) graph the curves corresponding to various ageing times could be obtained one from an another by a

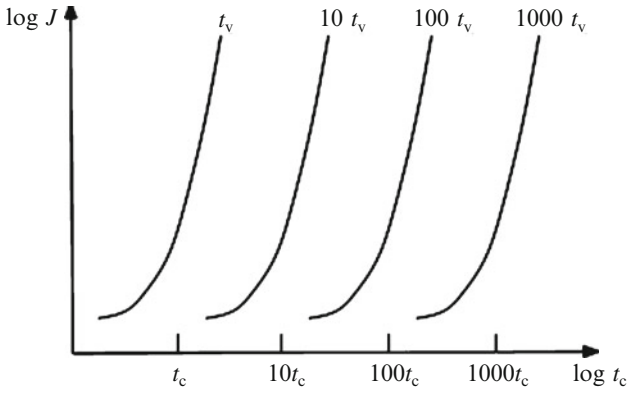


Fig. 5.28 Shape of creep curves (in a double logarithmic scale)

translation along the time axis and that, in a more or less wide temperature interval, the *shift factor* a_V was such that:

$$a_V = \frac{\Delta \log t \text{ (creep)}}{\Delta \log t \text{ (ageing)}} \approx 1 \tag{5.89}$$

It appeared later that this is a common property of a wide variety of phenomena of physical ageing, from emulsions to sand heaps. As expected, the shift factor a_v remains close to unity in a wide temperature interval below T_g for polymers of the polycarbonate type having their β transition well separated from T_g. In contrast, the curve $a_v = f(T)$ displays a relatively sharp maximum just below T_g for polymers of the polystyrene type having their T_β close to T_g.

5.2.1.8 The Case of Semi-crystalline Polymers

The temperature dependence of tensile or shear moduli of semi-crystalline polymers is schematised in Fig. 5.29.

In glassy state, the crystalline phase is stiffer than the amorphous phase because it is more densely packed ($\rho_c/\rho_a \approx 1.17$ in average). However the moduli of both phases are of the same order of magnitude. In rubbery state, however, the crystalline phase is about one thousand times stiffer than the amorphous phase. In a first approximation, the modulus of a semi-crystalline polymer above T_g is quasi proportional to the degree of crystallinity x_c. A more precise approach needs to take into account the crystalline morphology, especially lamellae dimensions.

Practitioners define the *heat deflection temperature* (*HDT*) as the temperature above which a sample submitted to a fixed stress undergoes a deformation higher than an arbitrary value. It can be understood from Fig. 5.29 why, in an amorphous

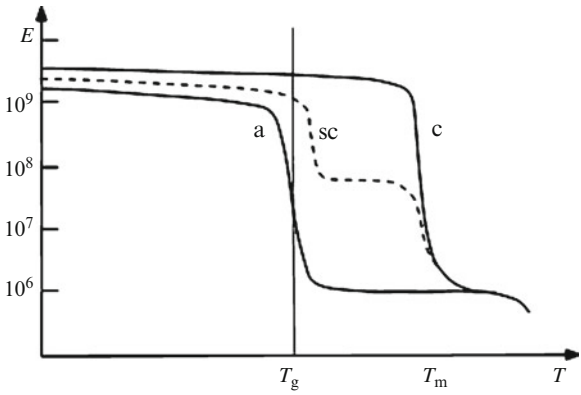


Fig. 5.29 Schematic shape of the modulus – temperature curves for the amorphous phase (a) and the crystalline phase (c) of a semi-crystalline (sc) polymer

polymer, $HDT \approx T_g - \Delta T$, where $\Delta T \approx 5$–10 K, whereas for a semi-crystalline polymer HDT is higher than T_g and is often given by: $HDT \approx T_m - \Delta T$, where $\Delta T \approx 40$–100 K depending on the crystallinity ratio.

5.2.2 Internal Friction of Metals

5.2.2.1 Terminology and Methodology

We referred to internal friction in Chap. 3 as far as it provides information concerning various dislocations mechanisms. Internal friction manifests itself by energy dissipation in cyclic deformation.[17] It means that the loading path in a load *vs* deformation plot is a closed loop, the area of which measures the energy dissipation. Small local stress sensitive phenomena produce measurable internal friction, which makes *mechanical spectroscopy* an interesting tool to study these phenomena or to detect their presence and their importance.

Internal friction is a manifestation of viscoelasticity. The notions introduced in Sect. 5.1 are applicable. However, internal friction specialists have their own terminology. It is appropriate to give it here, at the expense of repeating what has been already introduced in Sect. 5.1 or developed in Sect. 5.2.1 about dynamic testing of polymers.

An important distinction must be made between *anelastic relaxation* and *hysteretic behaviour*. Linear viscoelasticity, for which the *loss factor* is amplitude independent, is associated with anelastic relaxation. On the contrary, the loss factor in hysteretic behaviour is amplitude dependent. It is usually connected with a frequency independent contribution to the hysteresis. In particular it remains present in purely static loading.

[17] The book of Blanter et al. (2007) was of great help in writing this section.

In the following, we will deal most of all with anelastic relaxation phenomena, like for instance point defects relaxation. We will afterwards briefly mention some phenomena producing hysteretic behaviour, like phase transformations or dislocations escape.

Figure 5.3 represents a typical load-displacement curve of anelastic behaviour, with complete recovery, if, after complete unloading, the deformation goes back to zero. A completely similar figure can be drawn representing the stress-strain behaviour. Upon application of the stress, the strain takes instantaneously a value called the unrelaxed strain ε_U and then the creep strain $\varepsilon(t)$ (related to the linear creep function $f(t,t_0)$ defined in Sect. 5.1) tends to the relaxed strain ε_R. The behaviour is characterised by the *relaxation strength* $\Delta = (\varepsilon_R - \varepsilon_U)/\varepsilon_R$. The relaxation strength can as well be expressed as a function of the modulus $E(t)$.

The *standard anelastic solid* or *standard linear solid* (see Sect. 5.1.3.3) is characterised by three parameters only: ε_U (or E_U), ε_R (or E_R) and a single *relaxation time* τ ($\tau = \tau_\sigma = \tau_\varepsilon$), so that the creep strain is proportional to $\exp(-t/\tau)$. (τ is designated by θ in Sect. 5.1.3.3). In cyclic loading, the response to an applied sinusoidal stress $\sigma^* = \sigma_0 \exp(i\omega t)$ is a strain $\varepsilon^* = \varepsilon_0 \exp(i\omega t - \phi)$. The complex modulus E^* is then:

$$E^* = \frac{\sigma^*}{\varepsilon^*} = E(\omega)\exp[i\phi(\omega)] = E'(\omega) + iE''(\omega) \tag{5.90}$$

Thus are distinguished the storage modulus E' and the loss modulus E'', whereas ϕ is the loss angle. And now:

$$\tan\phi = \frac{E''}{E'} = \frac{\Delta W}{2\pi W} = Q^{-1} \tag{5.91}$$

All this is but a repetition of what has been written in Sect. 5.1.3.4 (Eqs. 5.51, 5.52, 5.53), giving however the usual notations used in the discipline of internal friction. Note in particular that the energy dissipation is denoted Q^{-1}.

The dynamic response functions of the standard anelastic solid are given by the Debye[18] equations:

$$E'(\omega) = E_R\left(1 + \Delta\frac{\omega^2\tau^2}{1+\omega^2\tau^2}\right) = E_U\left(1 + \frac{\Delta}{1+\Delta}\frac{1}{1+\omega^2\tau^2}\right) \tag{5.92}$$

and

$$Q^{-1}(\omega) = \frac{\Delta}{\sqrt{1+\Delta}}\frac{\omega\tau}{1+\omega^2\tau^2} \tag{5.93}$$

As a general rule Δ can be neglected in front of 1 and the preceding equations are simplified.

[18] Peter Joseph Wilhem Debye (1884–1966) was a Dutch physicist and chemist who won the Nobel Prize in chemistry in 1936.

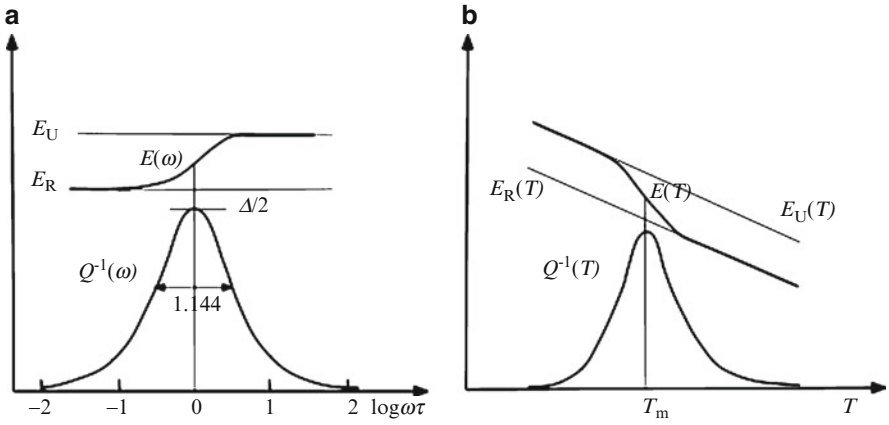


Fig. 5.30 Dynamic modulus E and internal friction Q^{-1} of the standard anelastic solid: (**a**) as a function of frequency on a log$\omega\tau$ scale; (**b**) as a function of temperature at constant frequency

Figure 5.30 shows the *Debye peak* and these various parameters (see also Figs. 5.22 and 5.23). Note that the loss of internal friction is only detectable around $\omega\tau = 1$. At high frequencies the defects responsible for relaxation do not move. At low frequencies they have all the time needed to achieve equilibrium.

Most of the relaxation phenomena are thermally activated, so that the relaxation time can be assumed to depend on temperature in the following way:

$$\tau = \tau_0 \exp\left(\frac{H}{kT}\right) \tag{5.94}$$

τ represents then the reciprocal jump frequency ν^{-1} over energy barriers of height H. The energy dissipation Q^{-1} can then be expressed as a function of temperature at fixed frequency $f = \omega/2\pi$. If T_m is the temperature at the Debye peak for $\omega\tau = 1$, the internal friction is:

$$Q^{-1} = \frac{\Delta}{2}\operatorname{sech}\frac{H}{k}\left(\frac{1}{T} - \frac{1}{T_m}\right) \tag{5.95}$$

in which expression sech $x = (\cosh x)^{-1}$. The Debye peak is symmetric with a half width equal to 2.635 k/H.

The activation parameters: apparent activation energy H and limit relaxation time (reciprocal attempt frequency) $\tau_0 = \nu_0{}^{-1}$ are obtained from the shift of the peak temperature when changing the frequency:

$$\log\left(\frac{f_2}{f_1}\right) = \frac{H}{k}\left(\frac{1}{T_{m1}} - \frac{1}{T_{m2}}\right) \tag{5.96}$$

Table 5.5 Some internal friction parameters for some metals

Composition	f (Hz)	T_m (K)	Q_m^{-1} (10^{-4})	H (kJ:mol)	τ_0(s)
Cu4-14Mn10Al	1.1	273–113	350–2,000	–	–
Cu13.75Al4.95Ni	1	260–290	3,300	–	–
Al	0.03–0.6	450	60–120	101	$2.5 \cdot 10^{-12}$
Al	10^7	185–190	50–85	14.5–16	–
Pb	1	30	30	–	–
		110	250		
		250	200		
Ti3-4N	1	773	–	240.6	–
Ti51Ni	1	178 cooling	∼300		
		181 heating			
FeC	1	314	3.5	83.7	$1.9 \cdot 10^{-15}$
FeN	1	296–298	–	73.4–76.9	$(5–18)10^{-15}$
FeH	1	105–210	<10	29	–

These examples are there to provide orders of magnitudes for typical systems. There are many different values given in the literature

All this can be repeated when there is not a single relaxation time but a relaxation spectrum. The internal friction peaks as a function of temperature are then broader, and what is determined is an average value of the parameters.

It must be noted that experimentation as a function of temperature is much easier than as a function of frequency. Internal friction determinations (see Table 5.5) are usually performed at fixed frequency and varying temperature.

5.2.2.2 Measurement of Internal Friction

There exists a wide variety of techniques for measuring the internal friction. We cannot give here a complete description of them. We limit ourselves to evoke the most widely used ones.

The *low frequency forced torsion pendulum* works well below the resonance frequency in forced oscillations. The determination of the internal friction consists in measuring the time lag ϕ. The main advantage of this technique is that it offers the possibility to perform isothermal experiments over a continuous frequency range from 10^{-4} to 10^2 Hz.

However, *resonant experiments* are much more common. They use torsion pendulums, vibrating reeds, composite oscillators and resonant ultrasound spectroscopy. It is possible to measure directly the energy loss by analysing the relative magnitude of the input and output signals. However, it is more usual to measure either the *width of the resonant peak* or the *free decay*. This last method is the most widely spread.

The *free decay* method uses the free damped vibrations after turning off the excitation. The *logarithmic decrement* δ is determined, which is equal to $\log (A_n/A_{n+1})$, A representing the successive amplitudes of oscillations. Then

$$Q^{-1} = \delta / \pi \tag{5.97}$$

It is usual practice to put the equipment under vacuum in order to suppress the friction in air. Various devices allow varying the temperature.

5.2.2.3 Point Defect Relaxation

Point defect relaxation is usually caused by the redistribution by diffusion of point defects under the influence of an applied stress. This requires an elastic interaction between the applied stress and the lattice distortion due to the defect. Shear components of the applied stress interacting with anisotropic distortions around the defects induce short-range diffusion, over distances of the order of the interatomic distances. The corresponding relaxation times are short.

The Snoek Relaxation

In Sect. 3.3.4.1 we mentioned that interstitial atoms C, N in α-Fe, in octahedral sites, being larger than the available interstice between atoms of iron increased the cube edge of the lattice (Fig. 3.75) and we briefly described the redistribution of the interstitial atoms when an applied stress elongated the lattice in one direction. Under cyclic stress, this produces a relaxation called the Snoek relaxation. This phenomenon is not limited to C and N in α-Fe. It is observed as well for O, C and N solid solutions in BCC metals V, Nb, Ta, Cr, Mo and W.

When a tensile stress is applied, the dimensions of part of the octahedral interstices are increased while the Poisson effect reduces the dimensions of the other. A difference in the concentration of interstitials in the elongated and contracted interstices results. When the stress is reversed there is a redistribution of the interstitials in the second ones now of greater length than the first. This redistribution requires diffusion of the interstitials over atomic scale distances and the phenomenon is thermally activated with an activation energy, which is that of the diffusion. The relaxation time is related to the diffusion coefficient D:

$$\tau = \frac{{a_0}^2}{36D} \tag{5.98}$$

The Snoek peak temperature T_m for $2\pi f\tau = 1$ is then given by:

$$T_m = \frac{H}{R \log\left(\pi {a_0}^2 f / 18 D_0\right)} \tag{5.99}$$

Figure 5.31 shows an example of a Snoek peak for niobium.

The Snoek peak temperatures at a frequency of 1 Hz are 314 and 300 K for Feα – C and for Feα − N respectively. The corresponding activation energies H are

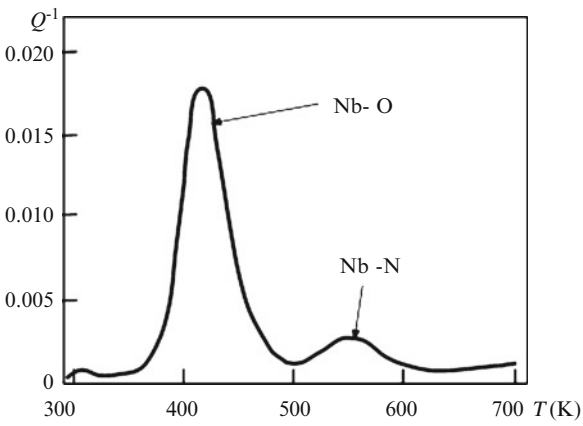


Fig. 5.31 Snoek peak for Nb at 0.8 Hz (Adapted from Grandini et al. 2005)

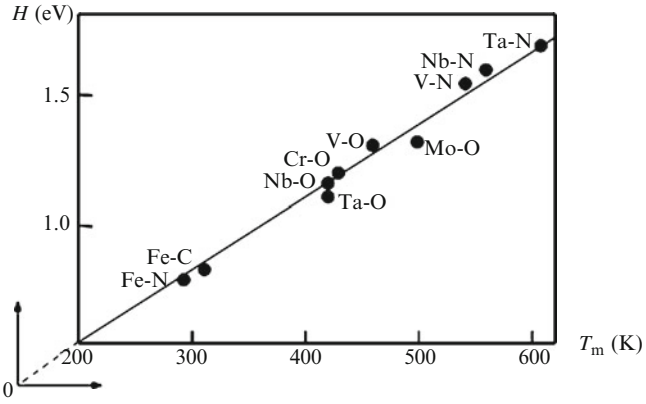


Fig. 5.32 Activation energy versus Snoek peak temperature for various BCC solid solutions (Adapted from Weller 1985)

83.7 and 78.8 kJ/mol while the relaxation times τ_0 are $1.89{\cdot}10^{-15}$ and $2.38{\cdot}10^{-15}$ s. Parameters for other solid solutions of BCC metals can be found in Blanter et al. (2007).

It was found that the D_0 values recorded were almost the same for various solid solutions. Hence, the Snoek peak temperature is proportional to the activation energy of diffusion (Fig. 5.32). This corresponds to $\tau_0 = 2.08{\cdot}10^{-15}$ s and T_m (K) $= 3.765\,H$ (kJ/mol).

The height of the Snoek peaks is proportional to the solute concentration (Fig. 5.33). The slope is higher for carbon atoms (0.215 per 1 at% solute) than for nitrogen ones (0.2 per 1 at% solute) as they are bigger and distort more the BCC lattice. Measurement of Snoek peaks is a very sensitive method, which is able to detect small amounts of solutes.

More about Snoek peaks can be found in Blanter et al. (2007) concerning for instance the influence of alloying elements in FCC and HCP metals.

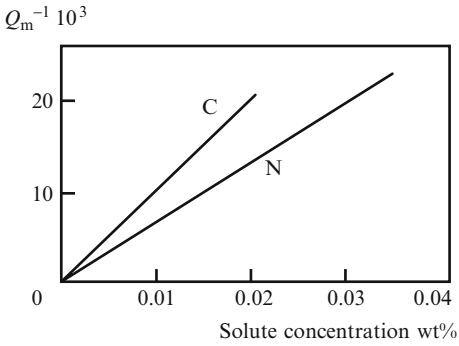


Fig. 5.33 Heights of the Snoek peaks in Fe-C and Fe-N at 1 Hz (Adapted from Lenz and Dahl 1974)

The Zener Relaxation

Substitutional alloying elements do not distort the solvent lattice. Isolated they do not produce any relaxation. However, if they are grouped by pairs, they introduce an anisotropy, which is responsible for internal friction relaxation peaks. These are called Zener relaxation peaks. They are found in the temperature range around 700 K where the solute atoms are able to move by diffusion.

The relaxation strength Δ_Z is proportional to the square of the solvent concentration. The activation energies are found to be slightly less than the activation energies for diffusion. This is due to the grouping by pairs of the solvent atoms responsible for the relaxation.

The heights of the peaks are much lower than the ones of the Snoek peaks. This is because the distortion of the lattice by substitution atom pairs is weaker than by interstitials.

5.2.2.4 Dislocation Relaxations

Applying a small enough stress moves the dislocations in a reversible fashion with nevertheless a time lag for the strain response giving rise to internal friction. When the yield strength is reached, plastic deformation produces an amplitude dependent internal friction. The relaxation phenomena intervening in the first type of internal friction at very low strain amplitude give rise to a number of peaks, which are denoted B and P in FCC and HCP metals and α, β and γ in BCC metals in order of increasing temperature. The internal friction is also related to the interactions between foreign atoms and dislocations. We will now describe the main phenomena responsible for dislocation relaxation.

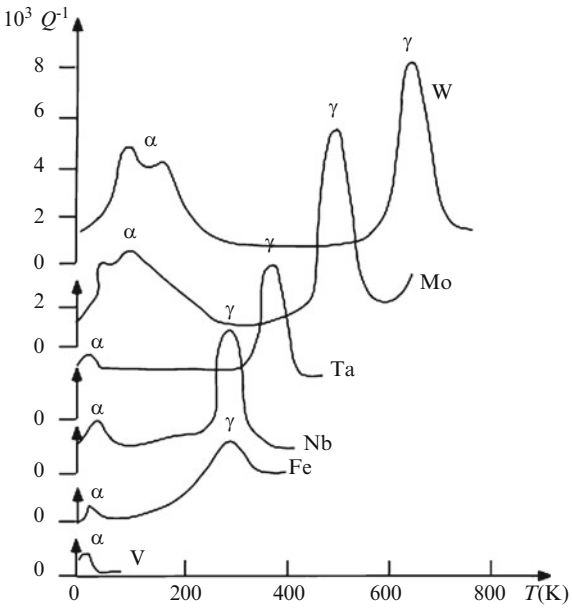


Fig. 5.34 Bordoni peaks in pure BCC metals (Adapted from Schultz 1991)

The Bordoni Relaxation

The displacement of dislocations from a Peierls trough to another occurs by nucleation and displacement of kinks (Sect. 3.4.2.1). Low temperature relaxation peaks are associated with these kinks. The B_1 and B_2 peaks, in order of increasing temperature, in FCC metals for which the kink energy is low, are dependent on the strain amplitude. The second one appears only if this is sufficient and it increases with the strain amplitude. In BCC metals, the α peaks are observed at low temperatures (around 100 K), while the γ ones occur in the temperature range from 300 to 700 K (Fig. 5.34). In general the presence of impurities decreases the height of the Bordoni peaks owing to the anchoring of dislocations, notwithstanding some stacking fault energy effects.

The presence of two B peaks in FCC metals and of α and γ peaks in BCC ones is explained by the difference of the nucleation energies of kink pairs on pure screw dislocations (B_1 and γ) and on mixed ones (B_2 and α_2). The α_1 peak in BCC metals is attributed to the displacement of kinks.

The relaxation strength of the Bordoni relaxation is proportional to the dislocation density and the relaxation time to the drag coefficient of the dislocation velocity.

Hasiguti Relaxation

These peaks called P in FCC metals are found at higher temperatures than the Bordoni peaks; called β in BCC metals, they are found at intermediate temperatures

between the α and γ Bordoni peaks. In FCC metals there are three Hasiguti peaks P_1, P_2 and P_3 in order of increasing temperature. The Hasiguti peaks are attributed to the interaction of dislocations with self-point defects created by plastic deformation: vacancies, interstitials and divacancies.

Snoek-Köster Relaxation

This kind of relaxation is due to the interaction between interstitial solute atoms and dislocations. The corresponding peaks (SK peaks) are found at higher temperatures than the ones due to other relaxation mechanisms (except when the foreign atoms are H, because then their diffusion is very fast). The height of the SK peaks increases with the concentration of solute atoms and with prior plastic deformation. It also increases during annealing. The peak temperatures and the activation energies are also concentration and deformation dependent.

As the solute interstitial atoms as well as the dislocations and their kinks are all moving in internal friction experiments, the detailed explanation of the SK relaxation is complicated and there are several possible mechanisms. It could be due to the bowing of the dislocations segments combined with the dragging of the interstitial atoms. At the same time there could be diffusion of these atoms along the dislocation lines. Another mechanism could be the interaction of mobile interstitials with the formation of double kinks, related then to the Bordoni peaks.

Another related relaxation is the *dislocation enhanced Snoek effect*. This enhancement of the height of the Snoek peaks is brought about by strain-hardening. It shows again an interaction between interstitials and dislocations.

Damping Background at Elevated Temperatures

Finally, it should be mentioned that can be observed a background, which increases exponentially with temperature above 0.6–0.7 the absolute melting temperature. The mechanism could be interactions between dislocations and grain boundaries or dislocation walls. In fact it is grain size dependent and varies also with the degree of polygonisation.

5.2.2.5 Grain Boundary Relaxation

Kê (1947) found a relaxation peak in polycrystalline aluminium at about 500 K, which is absent in single crystals. It was attributed to grain boundary mobility. It was found as well in various alloys.

The relaxation time of grain boundary relaxation should be proportional to the grain size. The relaxation strength should not depend on the grain size.

Heat treatments, which produce precipitates in the grain boundaries, suppress the grain boundary relaxation. An example is precipitation of chromium carbides in Ni-Cr alloys.

We would expect this kind of relaxation to be very important in nanocrystalline materials. It has been found that these materials exhibit an internal friction peak with an activation energy of about 1 eV, which does not exist in materials with grains of ordinary size. Although this relaxation could be of the Kê type, it could be attributed also to recrystallisation.

5.2.2.6 Dislocation Related Amplitude-Dependent Internal Friction

At very low stress amplitude the motion of dislocations gives rise, as we have previously discussed, to pure amplitude-independent relaxation phenomena. Now, if the amplitude increases above a yielding threshold the dislocations can escape from their anchorage points and move over relatively large distances. This produces amplitude-dependent internal friction of the hysteretic type. It is observed that above this critical stress, internal friction rises exponentially when the stress amplitude increases.

In Sect. 3.4.3 we studied the interaction of dislocations with foreign atoms: we saw that they first bow between the pinning points and we calculated the stress needed for the dislocations to escape. When the stress is reversed the dislocations move backwards and are pinned again by the same atoms.

The corresponding internal friction was modelled by Granato and Lücke (1956). They get the following expressions for the internal friction Q^{-1} and for the internal friction strength Δ (at 0 K):

$$\begin{aligned} Q^{-1} &= \frac{\Delta}{\pi}\frac{\sigma_c}{\sigma_0}\exp\left(-\frac{\sigma_c}{\sigma_0}\right) \\ \Delta &= \frac{\rho_m b^2 \Lambda^2}{12 t \varepsilon_0 / \sigma_0} \end{aligned} \tag{5.100}$$

where σ_c is the stress needed to break away the dislocations from their pinning points, that is K/bL if K is the force exerted by the obstacle on the dislocation, b the Burgers vector and L the mean distance between pinning points; σ_0 is the stress amplitude; ρ_m is the density of mobile dislocations; Λ is the length of dislocations; t the line tension ($\mu b^2/2$); ε_0 the unrelaxed strain.

The model yields indeed an internal friction, which is an exponential function of the stress amplitude. Plotting $\log \varepsilon_0 Q^{-1}$ as a function of $1/\varepsilon_0$ should give a straight line, the slope of which yields the critical stress and the intercept the dislocation density. Thus internal friction experiments, performed in the kHz range, can provide interesting information about the pinning obstacles and their evolution during heat treatment for instance.

For temperatures different from 0 K, the activation energy for the overcoming of obstacles must be taken into account and this leads to the same expressions as (5.100) with a slight additional frequency dependence of Q^{-1}. The internal friction is then a mixture of relaxation and hysteresis.

5.2.2.7 Hysteretic Internal Friction due to Phase Transformations. High Damping Materials

Hysteretic internal friction occurs owing to phase transformation when this is accompanied by volume change of shear deformation. A common example is the martensitic transformation. It was studied in Sect. 3.4.4.3. Transformation plasticity (Sect. 3.4.4.4) is also to be associated with this kind of internal friction. Cyclic straining can produce back and forth transformations with a lag of the strain response.

The *transient damping* ${Q_{\mathrm{TR}}}^{-1}$, as opposed to *non-transient* and to *intrinsic damping*, is by far the most important contribution. The non-transient damping is due to interfaces and dislocations displacements and the intrinsic damping to the combination of the damping of the two coexisting phases. The transient damping is attributed to the transformation rate accompanied by anelastic strain. It depends on the rate of cooling and heating ($\dot{T}$ effect). It increases with this rate. It is inversely proportional to the frequency. At low amplitudes, the mechanism is that of Granato-Lücke escape of dislocations, but the phase transformation internal friction itself, increasing with the amplitude, appears above amplitudes larger than $5{\cdot}10^6$ to $2{\cdot}10^5$.

The shape memory alloys, a prototype of which is *nitinol* NiTi at near equi-atomic composition, are examples of materials producing phase transformation damping. For these alloys, during cooling, in the temperature range between 150°C and −40°C depending on the composition, the high temperature BCC (B2) is transformed in a monoclinic (B19') martensite. During cooling then, the internal friction Q^{-1} goes through a maximum. It is again the case upon heating but at a higher temperature.

Other examples are zinc, aluminium or tin based copper alloys, which display a martensitic transformation, the high temperature disordered phase transforming to an ordered one upon cooling. The damping for these alloys, in particular Cu-Al-Ni alloys, is very high allowing their use as *Hidamets* or *high damping materials (HDM)* in various practical applications. Cu-Mn alloys (32–42 wt% Mn and 2–4 wt% Al) known as "Sonoston" are used to fabricate ship propellers owing to their high damping capacity. Classifying an alloy as HDM cannot be done without reference to a particular application, because the damping characteristics depend on many different factors (as amplitude, temperature, frequency).

In view of the large number of applications in which the damping capacity of materials is an important parameter, we limit ourselves to these typical examples of HDM only, referring the reader to Blanter et al. (2007) and to Lakes (2009).

5.3 Viscoelastic Structures and Heterogeneous Materials

In this section, we restrict ourselves to *linear* (mainly *non-ageing*) viscoelastic behaviour, *isothermal* evolutions and *infinitesimal transformations*. We first extend the consequences of the superposition principle, which have been described in Sect. 5.1 for a global 1-D description, to general local 3-D situations. Then, we take advantage of the assumption of non-ageing and of the use of the Carson transform to reduce viscoelastic to *elastic* structural design through the *correspondence theorem*. Finally, we apply this theorem to heterogeneous materials and we introduce the basic homogenisation techniques for viscoelasticity.

5.3.1 Local Viscoelastic Constitutive Equations

5.3.1.1 General Case

We have already studied the case of a global linear non-ageing viscoelastic system defined by a finite number of pairs of parameters (Q_i, q_i) and obtained the constitutive equations (5.37). The same treatment can be applied to the local description of a volume element defined by the stress and strain tensors ($\underline{\underline{\sigma}}$, $\underline{\underline{\varepsilon}}$). When the local behaviour is supposed to be linear viscoelastic, the superposition principle, applied in the same way as in Sect. 5.1.2.2, leads to the generalisation of (5.10), (5.11) and (5.12), namely

$$\begin{cases} \underline{\underline{\varepsilon}}(t) = \mathbf{f}(t,0) : \underline{\underline{\sigma}}(0_+) + \displaystyle\int_0^t \mathbf{f}(t,\tau) : \mathrm{d}\underline{\underline{\sigma}}(\tau) \\ \underline{\underline{\varepsilon}}(t) = \mathbf{f}(t,t) : \underline{\underline{\sigma}}(t) - \displaystyle\int_0^t \frac{\partial \mathbf{f}(t,\tau)}{\partial \tau} : \underline{\underline{\sigma}}(\tau)\mathrm{d}\tau \end{cases} \tag{5.101}$$

and

$$\begin{cases} \underline{\underline{\sigma}}(t) = \mathbf{r}(t,0) : \underline{\underline{\varepsilon}}(0_+) + \displaystyle\int_0^t \mathbf{r}(t,\tau) : \mathrm{d}\underline{\underline{\varepsilon}}(\tau) \\ \underline{\underline{\sigma}}(t) = \mathbf{r}(t,t) : \underline{\underline{\varepsilon}}(t) - \displaystyle\int_0^t \frac{\partial \mathbf{r}(t,\tau)}{\partial \tau} : \underline{\underline{\varepsilon}}(\tau)\mathrm{d}\tau \end{cases} \tag{5.102}$$

Here, $\mathbf{f}(t,\tau)$ and $\mathbf{r}(t,\tau)$ are *fourth-order* tensorial functions of the two independent variables t and τ, the *creep and relaxation tensors*, respectively. This meaning is proved by considering the responses to the stress and strain imposed histories

$\underline{\underline{\sigma}}_0 \mathrm{H}(t-\tau)$ and $\underline{\underline{\varepsilon}}_0 \mathrm{H}(t-\tau)$, respectively. This can be done, for instance, for homogeneous uniaxial stress or strain histories so as to recover the 1-D description of Sect. 5.1.

When *non-ageing* linear viscoelasticity is assumed, as in all what follows, the creep and relaxation tensors only depend on the variable $(t-\tau)$ and we find the generalisation of (5.37), namely

$$\begin{cases} \underline{\underline{\varepsilon}}(t) = \dfrac{\mathrm{d}}{\mathrm{d}t}\left[\displaystyle\int_0^t \mathbf{f}(t-\tau) : \underline{\underline{\sigma}}(\tau)\mathrm{d}\tau\right] = \mathbf{f}\odot : \underline{\underline{\sigma}} \\ \underline{\underline{\sigma}}(t) = \dfrac{\mathrm{d}}{\mathrm{d}t}\left[\displaystyle\int_0^t \mathbf{r}(t-\tau) : \underline{\underline{\varepsilon}}(\tau)\mathrm{d}\tau\right] = \mathbf{r}\odot : \underline{\underline{\varepsilon}} \end{cases} \tag{5.103}$$

where the symbol $\odot$: denotes at the same time a Stieljes convolution product and the double product of a fourth-order tensor and a second-order tensor, *e.g.*:

$$\left(\mathbf{f}\odot : \underline{\underline{\sigma}}\right)_{ij} = \frac{\mathrm{d}}{\mathrm{d}t}\left[\int_0^t f_{ijkl}(t-\tau)\sigma_{kl}(\tau)\mathrm{d}\tau\right] \tag{5.104}$$

Now, the Carson transform works in the same way as in (5.38), *i.e.*,

$$\begin{cases} \underline{\underline{\varepsilon}}^*(p) = \mathbf{f}^*(p) : \underline{\underline{\sigma}}^*(p) \\ \underline{\underline{\sigma}}^*(p) = \mathbf{r}^*(p) : \underline{\underline{\varepsilon}}^*(p) \end{cases} \tag{5.105}$$

Due to the symmetries of $\underline{\underline{\varepsilon}}^*(p)$ and $\underline{\underline{\sigma}}^*(p)$, the fourth-order tensors $\mathbf{f}^*(p)$ and $\mathbf{r}^*(p)$ have only 36 independent components instead of 81 since we can state that

$$\begin{cases} f^*_{ijkl}(p) = f^*_{jikl}(p) = f^*_{ijlk}(p) \\ r^*_{ijkl}(p) = r^*_{jikl}(p) = r^*_{ijlk}(p) \end{cases} \tag{5.106}$$

In addition, a thermodynamical analysis within the framework of the Biot[19] theory (Biot 1954, 1956), based on the use of the Onsager[20] principle, leads to establish their diagonal symmetry

$$\begin{cases} f^*_{ijkl}(p) = f^*_{klij}(p) \\ r^*_{ijkl}(p) = r^*_{klij}(p) \end{cases} \tag{5.107}$$

Finally, $\mathbf{f}^*(p)$ and $\mathbf{r}^*(p)$ are *inverse* from each other.

So, in the space of the Carson-transformed functions, the "Carson space", say, for any fixed value of p, there is no difference between the constitutive equations

[19]Maurice Anthony Biot (1905–1985) was a Belgian-American physicist.

[20]Lars Onsager (1903–1976) was a Norwegian-born American physical chemist and theoretical physicist, winner of the 1968 Nobel Prize in Chemistry.

(5.105) and those of linear elasticity. This remark is more developed below in the next paragraph. So far, we could specify $\mathbf{f}^*(p)$ and $\mathbf{r}^*(p)$ for different types of anisotropy as we did for linear elasticity in Chap. 2 (Sect. 2.3.2): the results would be the same by substituting $\mathbf{c}$ with $\mathbf{r}^*(p)$ and $\mathbf{s}$ with $\mathbf{f}^*(p)$. Let us only mention the case of *isotropy*.

5.3.1.2 Isotropic Non-ageing Viscoelasticity

Instead of (2.78) and (2.81), we write now

$$\begin{cases} \underline{\underline{\sigma}}^*(p) = \lambda^*(p)\mathrm{Tr}\left[\underline{\underline{\varepsilon}}^*(p)\right]\underline{\underline{\delta}} + 2\mu^*(p)\underline{\underline{\varepsilon}}^*(p) \\ \underline{\underline{\varepsilon}}^*(p) = \dfrac{1+\nu^*(p)}{E^*(p)}\underline{\underline{\sigma}}^*(p) - \dfrac{\nu^*(p)}{E^*(p)}\mathrm{Tr}\left[\underline{\underline{\sigma}}^*(p)\right]\underline{\underline{\delta}} \end{cases} \tag{5.108}$$

where $\lambda^*(p)$ and $\mu^*(p)$ are the Carson-transformed functions of $\lambda(t)$ and $\mu(t)$ which are the equivalent of the Lamé coefficients λ and μ; $E^*(p)$ and $\nu^*(p)$ are the Carson-transformed functions of $E(t)$ and $\nu(t)$ which are the equivalent of the Young modulus and Poisson ratio, respectively. The elastic bulk modulus k has also its viscoelastic counterpart $k(t)$ which is linked with $\lambda(t)$ and $\mu(t)$ like k with λ and μ. More generally, since isotropy implies that the elastic moduli $\mathbf{c}$ can be put in the form $\mathbf{c} = 2\mu\mathbf{K} + 3k\mathbf{J}$ with $\mathrm{J}_{ijkl} = (1/3)\delta_{ij}\delta_{kl}$ and $\mathbf{K} = \mathbf{I} - \mathbf{J}$ (see footnote 24 of Chap. 2), we have for isotropic non-ageing viscoelasticity

$$\begin{aligned} &\mathbf{r}^*(p) = 2\mu^*(p)\mathbf{K} + 3k^*(p)\mathbf{J} \\ &\quad\Rightarrow \underline{\underline{\sigma}}^*(p) = 2\mu^*(p)\underline{\underline{e}}^*(p) + k^*(p)\mathrm{Tr}\left[\underline{\underline{\varepsilon}}^*(p)\right]\underline{\underline{\delta}} \end{aligned} \tag{5.109}$$

where $\underline{\underline{e}}$ is the strain deviator.

It must not be forgotten that the condensed form of (5.108) relates to a complex expression in the time-space, namely

$$\begin{cases} \underline{\underline{\sigma}}(t) = \lambda(t) \odot \mathrm{Tr}\left[\underline{\underline{\varepsilon}}(t)\right]\underline{\underline{\delta}} + 2\mu(t) \odot \underline{\underline{\varepsilon}}(t) \\ E(t) \odot \underline{\underline{\varepsilon}}(t) = [1+\nu(t)] \odot \underline{\underline{\sigma}}(t) - \nu(t) \odot \mathrm{Tr}\left[\underline{\underline{\sigma}}(t)\right]\underline{\underline{\delta}} \end{cases} \tag{5.110}$$

These expressions are a little bit simplified when the Poisson ratio is assumed to be constant, as it is frequently considered for polymers, especially with $\nu = 1/2$ (isochoric deformation).

Like for the 1-D global description, additional simplifying assumptions can be made, *e.g.* the degeneracy of Stieljes convolution products into ordinary differential

equations with constant coefficients. First-order differential equations can be illustrated by the isotropic 3-D generalisation of the Kelvin and Maxwell models.

The generalised Kelvin model: instead of $Q = Eq + \eta\dot{q}$, we have simply

$$\underline{\underline{\sigma}} = \left[\lambda_1 \mathrm{Tr}(\underline{\underline{\varepsilon}}) + \lambda_2 \mathrm{Tr}(\dot{\underline{\underline{\varepsilon}}})\right]\underline{\underline{\delta}} + 2(\mu_1 \underline{\underline{\varepsilon}} + \mu_2 \dot{\underline{\underline{\varepsilon}}}) \tag{5.111}$$

This equation corresponds to a special case of (5.108) with

$$\begin{cases} \lambda^*(p) = \lambda_1 + \lambda_2 p & \Rightarrow \quad \lambda(t) = \lambda_1 + \lambda_2 \delta(t) \\ \mu^*(p) = \mu_1 + \mu_2 p & \Rightarrow \quad \mu(t) = \mu_1 + \mu_2 \delta(t) \end{cases} \tag{5.112}$$

The generalised Maxwell model: the 3-D transcription of $\dot{q} = \dot{Q}/E + Q/\eta$ reads

$$\dot{\underline{\underline{\varepsilon}}} = (\alpha_1 \dot{\underline{\underline{\sigma}}} + \alpha_2 \underline{\underline{\sigma}}) - \left[\beta_1 \mathrm{Tr}(\dot{\sigma}) + \beta_2 \mathrm{Tr}(\underline{\underline{\sigma}})\right]\underline{\underline{\delta}} \tag{5.113}$$

This is a special case of (5.108) too, with

$$\begin{cases} \dfrac{1+\nu^*(p)}{E^*(p)} = \alpha_1 + \dfrac{\alpha_2}{p} \\ \dfrac{\nu^*(p)}{E^*(p)} = \beta_1 + \dfrac{\beta_2}{p} \end{cases} \quad \Rightarrow \quad \begin{cases} \dfrac{1}{E^*(p)} = \dfrac{(\alpha_1 - \beta_1)p + \alpha_2 - \beta_2}{p} \\ \nu^*(p) = \dfrac{\beta_1 p + \beta_2}{(\alpha_1 - \beta_1)p + \alpha_2 - \beta_2} \end{cases} \tag{5.114}$$

whence the derivation of $E(t)$ and $\nu(t)$. Of course, other 1-D rheological models can be generalised in the same way.

5.3.1.3 Comments

For the sake of brevity, we do not develop here other 3-D generalisations of methods already presented for the 1-D description of Sect. 5.1 (recovery tests, DMA and complex moduli . . .). Let us only mention that the spectral representation approach also works for a 3-D local description. For example, the integral representation of any scalar relaxation function (5.42) would read now:

$$\mathbf{r}(t) = \mathbf{c}_\infty + \boldsymbol{\eta}_0 \delta(t) + \int_0^\infty \mathbf{g}(\tau)\exp(-t/\tau)\mathrm{d}\tau \tag{5.115}$$

with use of the fourth-order tensors $\mathbf{r}(t)$, $\mathbf{c}_\infty$, $\boldsymbol{\eta}_0$ and $\mathbf{g}(\tau)$.

5.3.2 *The Correspondence Theorem: Principle of Structural Design*

5.3.2.1 The Correspondence Theorem

We have already noted that the Carson-transformed constitutive equations for linear non-ageing viscoelasticity (5.105, 5.106, 5.107) look like the equations of linear elasticity, depending on the parameter p. It may then appear interesting to study the conditions for which the whole set of field equations for a boundary-value problem of viscoelasticity could reduce, after application of the Carson-transform, to those of some boundary problem of elasticity. The answer to this question is given by the "correspondence theorem" (Lee 1955; Mandel 1955, 1966; Salençon 2009).

Like for elasticity (see Sect. 2.3.4), the whole set of field equations for a quasi-static problem of linear non-ageing viscoelastic boundary-value problem at infinitesimal strain is the following:

(a) equilibrium: $\operatorname{div} \underline{\underline{\sigma}}(\underline{x}, t) + \underline{f}(\underline{x}, t) = 0$

(b) compatibility: $\begin{cases} \mathbf{Inc}\left[\underline{\underline{\varepsilon}}(\underline{x}, t)\right] = 0 \quad \text{or} \\ \underline{\underline{\varepsilon}}(\underline{x}, t) = \frac{1}{2}\left[\operatorname{grad} \underline{u}(\underline{x}, t) + \operatorname{grad}^{\mathrm{T}} \underline{u}(\underline{x}, t)\right] \end{cases}$

(c) constitutive equations $\begin{cases} \underline{\underline{\varepsilon}}(\underline{x}, t) = \left(\mathbf{f}\odot : \underline{\underline{\sigma}}\right)(\underline{x}, t) \\ \underline{\underline{\sigma}}(\underline{x}, t) = \left(\mathbf{r}\odot : \underline{\underline{\varepsilon}}\right)(\underline{x}, t) \end{cases}$

(d) boundary conditions: we have to specify that the complementary parts $S_{\mathrm{u}}(t)$ and $S_{\mathrm{T}}(t)$ of the whole surface $S = \partial V$ actually *do not depend* on time t, at least during the considered time interval, that is

$$\begin{cases} S_{\mathrm{u}} \cup S_{\mathrm{T}} = \partial V, \ S_{\mathrm{u}} \cap S_{\mathrm{T}} = \emptyset \\ \underline{u}(\underline{x}, t) = \underline{u}^{\mathrm{g}}(\underline{x}, t) \text{ on } S_{\mathrm{u}} \ \forall t \\ \underline{T}(\underline{x}, t) = \underline{T}^{\mathrm{g}}(\underline{x}, t) \text{ on } S_{\mathrm{T}} \ \forall t \end{cases} \tag{5.116}$$

where $\underline{u}^{\mathrm{g}}$ and $\underline{T}^{\mathrm{g}}$ are imposed on S_{u} and S_{T}, respectively.

We can now apply the Carson transform to all these equations, *at any point*, and we obtain the following field equations, depending on the space variable and the parameter p:

$$\operatorname{div} \underline{\underline{\sigma}}^{*}(\underline{x}, p) + \underline{f}^{*}(\underline{x}, p) = 0$$

$$\underline{\underline{\varepsilon}}^{*}(\underline{x}, p) = \frac{1}{2}\left[\operatorname{grad} \underline{u}^{*}(\underline{x}, p) + \operatorname{grad}^{\mathrm{T}} \underline{u}^{*}(\underline{x}, p)\right]$$

$$\begin{cases} \underline{\underline{\varepsilon}}^*(\underline{x}, p) = \mathbf{f}^*(\underline{x}, p) : \underline{\underline{\sigma}}^*(\underline{x}, p) \\ \underline{\underline{\sigma}}^*(\underline{x}, p) = \mathbf{r}^*(\underline{x}, p) : \underline{\underline{\varepsilon}}^*(\underline{x}, p) \end{cases} \text{ with } \mathbf{f}^*(\underline{x}, p) : \mathbf{r}^*(\underline{x}, p) = \mathbf{I}$$

$$\begin{cases} S_u \cup S_T = \partial V, \ S_u \cap S_T = \emptyset \\ \underline{u}^*(\underline{x}, p) = [\underline{u}^g]^*(\underline{x}, p) \text{ on } S_u \\ \underline{T}^*(\underline{x}, p) = [\underline{T}^g]^*(\underline{x}, p) \text{ on } S_T \end{cases} \tag{5.117}$$

These equations are nothing but those of a linear elastic boundary-value problem with the "elastic" moduli and compliances $\mathbf{f}^*(\underline{x}, p)$ and $\mathbf{r}^*(\underline{x}, p)$ (which means that this problem may concern an heterogeneous body) and with regular boundary conditions. The only difference with a classical problem of elasticity lies in the fact that all the considered variables depend on the complex parameter p, but no special difficulty arises from that: this only means that the solution, derived by use of classical (including numerical) methods of elasticity, also depends on this parameter. Nevertheless, the solution to the viscoelastic problem is expected to depend on the real time variable t, which still needs an *inverse Carson-transform* operation.

5.3.2.2 The Inverse Carson-Transform Operation

As already discussed in Sect. 5.1.3.2, this inversion can be performed in principle through the Mellin-Bromwich formula (5.27) or, for simple cases, by using tables for direct Carson transform of current analytical functions and expressing the function to be inversed in terms of them. Nevertheless, this is not always easy nor even possible, especially when recourse has been made to numerical treatments for the resolution of the symbolic elastic problem (parametrised with p). In such cases, some techniques have been developed to get an approximation of the viscoelastic solution. We briefly mention two of them.

The collocation technique: an approximation $f_{\text{col}}(t)$ to the original of some given function (let it be analytical or numerical) $f^*(p)$ is searched for, after substraction of instantaneous and asymptotic contributions, in the form of a Prony series[21] such as

$$f_{\text{col}}(t) = \sum_{i=1}^{N} f_i \exp(-t/\tau_i) \tag{5.118}$$

with, at this stage, arbitrary parameters f_i and τ_i. So, $f^*(p)$ is approximated by $f^*_{\text{col}}(p)$ given by

$$f^*_{\text{col}}(p) = \sum_{i=1}^{N} f_i \frac{p}{p + (1/\tau_i)} \tag{5.119}$$

[21]Note that when the considered function $f(t)$ is a relaxation function, according to (5.42), this approach reduces to the approximation of its continuous relaxation spectrum by a discrete one.

It can be proved that, when the times τ_i have been chosen, as proposed by the simplest collocation technique (Schapery 1962), the best approximation to $f^*(p)$ deriving from the condition

$$\min_{f_i} E = \min_{f_i} \int_0^\infty [f(t) - f_{\text{col}}(t)]^2 \mathrm{d}t \quad \Leftrightarrow \quad \partial E / \partial f_i = 0 \tag{5.120}$$

is obtained when we have

$$f^*_{\text{col}}(1/\tau_i) = f^*(1/\tau_i) \quad \forall i = 1, ..., \mathrm{N} \tag{5.121}$$

These conditions define a system of N linear equations with N unknowns whose solution yields $f_{\text{col}}(t)$ by (5.118). In practice, N is of the order of 10 and the times τ_i are often chosen equidistant on a logarithmic time scale.

The quasi elastic approximation: this method, also named the "direct method", derives from the approximation (Schapery 1965)

$$\log_{10} 10^w \mathrm{e}^{-10^w} \cong \delta(w - w_0) \tag{5.122}$$

with an adequate value of w_0. From that we get

$$\varphi(t) \cong \varphi^*(p)|_{p=\frac{10^{w_0}}{t}} \tag{5.123}$$

With the usual choice $w_0 = 0$, this means that the original $\varphi(t)$ is approximated by $\varphi^*(p)$ itself at $p = 1/t$. This rather crude but very convenient approximation is found to be rather satisfactory for smooth enough functions. Improvements result from specific choices for w_0, depending on the shape of $\varphi(t)$ (Brenner et al. 2002).

5.3.3 Homogenisation

5.3.3.1 Introductory Comments

From what precedes we can guess that the basic approach to linear non-ageing viscoelastic homogenisation consists in converting this problem into a linear elastic one through the use of the Carson transform. This is only partly true since one essential part of the elastic homogenisation theory has no viscoelastic counterpart, namely the variational approach and the bounding techniques. The reason for that lies in the fact that, at variance with elasticity, viscoelasticity cannot be defined energetically by one single potential: two potentials are needed for that, due to the *viscoelastic coupling* and, at the time being, no *general* variational approach is available for the homogenisation of the overall behaviour of heterogeneous materials whose local behaviour cannot be defined by one single potential. Some

partial results can be derived in terms of bounds for specific loadings (*e.g.* harmonic tests) but not for *any* loading, as it was the case for elasticity. Thus, we only give in what follows basic elements for the definition of the overall (relaxation) moduli and (creep) compliances and for the derivation of *estimates* for these quantities.

5.3.3.2 Effective Relaxation Moduli and Creep Compliances

The arguments of linearity which allowed us for the definition of the strain and stress concentration tensors **A** and **B** in elasticity still work for linear viscoelasticity, except for the fact that they are no more simple fourth-order algebraic tensors but operators. Instead of (1.27), namely

$$\begin{cases} \underline{\underline{\varepsilon}}(\underline{x}) = \mathbf{A}(\underline{x}) : \underline{\underline{E}} \\ \underline{\underline{\sigma}}(\underline{x}) = \mathbf{B}(\underline{x}) : \underline{\underline{\Sigma}} \end{cases} \tag{1.27}$$

we get for linear non-ageing viscoelasticity

$$\begin{cases} \underline{\underline{\varepsilon}}(\underline{x},t) = \mathbf{A}(\underline{x},\tau)\odot : \underline{\underline{E}}(\tau)\Big|_{\tau=-\infty}^{t} \\ \underline{\underline{\sigma}}(\underline{x},t) = \mathbf{B}(\underline{x},\tau)\odot : \underline{\underline{\Sigma}}(\tau)\Big|_{\tau=-\infty}^{t} \end{cases} \tag{5.124}$$

where the whole macroscopic strain or stress history has to be taken into account. These concentration equations have then to be combined with the local constitutive equations (5.103) in order to yield, with simplified notation

$$\begin{cases} \underline{\underline{\sigma}} = \mathbf{r}\odot : \left(\mathbf{A}\odot : \underline{\underline{E}}\right) \quad \Rightarrow \quad \underline{\underline{\Sigma}} =< \underline{\underline{\sigma}} >=< \mathbf{r}\odot : \mathbf{A} > \odot : \underline{\underline{E}} \\ \underline{\underline{\varepsilon}} = \mathbf{f}\odot : \left(\mathbf{B}\odot : \underline{\underline{\Sigma}}\right) \quad \Rightarrow \quad \underline{\underline{E}} =< \underline{\underline{\varepsilon}} >=< \mathbf{f}\odot : \mathbf{B} > \odot : \underline{\underline{\Sigma}} \end{cases} \tag{5.125}$$

and then

$$\begin{cases} \mathbf{R}^{\text{eff}} =< \mathbf{r}\odot : \mathbf{A} > \\ \mathbf{F}^{\text{eff}} =< \mathbf{f}\odot : \mathbf{B} > \end{cases} \tag{5.126}$$

As expected, the Carson-transformed effective relaxation moduli and creep compliances then read

$$\begin{cases} \left(\mathbf{R}^{\text{eff}}\right)^{*}(p) =< \mathbf{r}^{*}(p) : \mathbf{A}^{*}(p) > \\ \left(\mathbf{F}^{\text{eff}}\right)^{*}(p) =< \mathbf{f}^{*}(p) : \mathbf{B}^{*}(p) > \end{cases} \tag{5.127}$$

Since $\mathbf{r}^*(p)$ and $\mathbf{f}^*(p)$ are uniform within every phase (r), these formulae can also be written in the form

$$\begin{cases} \left(\mathbf{R}^{\text{eff}}\right)^* = \sum_{r=1}^{n} f_r \mathbf{r}_r^* : \mathbf{A}_r^*, \ \mathbf{A}_r^* = <\mathbf{A}^*>_{(r)} \text{ with } \sum_{r=1}^{n} f_r \mathbf{A}_r^* = \mathbf{I} \\ \left(\mathbf{F}^{\text{eff}}\right)^* = \sum_{r=1}^{n} f_r \mathbf{f}_r^* : \mathbf{B}_r^*, \quad \mathbf{B}_r^* = <\mathbf{B}^*>_{(r)} \text{ with } \sum_{r=1}^{n} f_r \mathbf{B}_r^* = \mathbf{I} \end{cases} \tag{5.128}$$

Like in elasticity, these general results could be complemented by dealing with initial strains and stresses, that is with eigenstrains and eigenstresses: this can be especially useful to study the case of heterogeneous thermo-viscoelasticity.

Nevertheless, like in elasticity, these theoretical results are of small usefulness for random heterogeneous materials because little is known about the spatial distribution of the phases in the RVE so that, in general, the per phase average concentration tensors $\mathbf{A}_r^*$ and $\mathbf{B}_r^*$ cannot be computed unambiguously. Since variational approaches aiming at deriving bounds for the effective quantities are not available, the only way to go further is to search for estimates for these quantities.

5.3.3.3 Estimates for the Effective Relaxation Moduli and Creep Compliances

The general approach to derive such estimates is quite similar to the one defined in Sect. 2.6 for elasticity, except for the use of Carson-transformed quantities instead of constant ones. Moreover, thanks to the *correspondence theorem*, the Green techniques which have been used to solve the inclusion problem and other parent situations in elasticity have their viscoelastic counterpart when dealing with Carson-transformed quantities. Since the solution to these elementary problems has been used to build definite estimates, we can go directly to these elastic estimates and replace the elastic constants on which they depend by the Carson-transformed equivalent quantities (Laws and McLaughlin 1978): thus we get in the Carson space as many viscoelastic estimates as we got in elasticity (laws of mixtures, Mori-Tanaka model, self-consistent scheme, Hashin estimates, etc.). Nevertheless, one additional operation is needed, namely the inverse Carson transformation of the final elastic results. This question has been discussed above in Sect. 5.3.2.2.

One simple example can be discussed in conclusion (see Volume III for details): we consider an isotropic incompressible two-phase material with Maxwellian isotropic incompressible phases and we are looking for a self-consistent estimate for the effective shear relaxation modulus $\mu^{\text{eff}}(t)$. The *elastic* solution is known to be the positive root of the second-order equation

$$\begin{gathered} X^2 + 2\left(\frac{5f-3}{6} + \frac{2-5f}{6}\beta\right) X - 2\beta = 0 \\ \text{with } X = \mu^{\text{eff}}/\mu_1, \beta = \mu_2/\mu_1, f = f_2 \end{gathered} \tag{5.129}$$

that is

$$X = -\frac{5f-3}{6} - \frac{2-5f}{6}\beta + \sqrt{\left(\frac{5f-3}{6} + \frac{2-5f}{6}\beta\right)^2 + 2\beta} \tag{5.130}$$

The strain and stress deviators in each phase (i) obey the constitutive equations

$$\underline{\underline{\dot{e}}} = a_i \underline{\underline{s}} + b_i \underline{\underline{\dot{s}}},\ i = 1,2 \quad \Rightarrow \quad 2\mu_i^*(p) = \frac{p}{a_i + b_i\, p} \tag{5.131}$$

Consequently, the Carson-transformed solution $X^*(p)$ reads

$$X^*(p) = -\frac{5f-3}{6} - \frac{2-5f}{6}\beta^*(p) + \sqrt{\left(\frac{5f-3}{6} + \frac{2-5f}{6}\beta^*(p)\right)^2 + 2\beta^*(p)}$$
$$\text{with } X^*(p) = \mu^{\text{eff}*}(p)/\mu_1{}^*(p),\ \beta^*(p) = \mu_2{}^*(p)/\mu_1{}^*(p), f = f_2 \tag{5.132}$$

This expression happens to be invertible in closed form (Rougier et al. 1993); it is still more interesting to compute analytically the associated relaxation spectrum, $g^{\text{eff}}(t)$, defined by

$$\mu^{\text{eff}}(t) = \int_0^\infty g^{\text{eff}}(\tau)\exp(-t/\tau)\mathrm{d}\tau \tag{5.133}$$

Whereas the relaxation spectrum of every (Maxwellian) phase reduces to one single line, the one of the two-phase material is found to be constituted of a *continuous* one on a given time interval and, according to the volume fraction f, of one or two single lines outside this interval. This continuous contribution is a good example of the effect of the viscoelastic coupling and of the complexity of the delayed interactions between the phases, which depend on the whole loading history. This result is a particular case of a general result (Suquet 1987), which states that a multiphase material with Maxwellian phases obeys the "effective" differential equation

$$\underline{\underline{\dot{E}}} = \mathbf{S}^{\text{eff}} : \underline{\underline{\dot{\Sigma}}} + \mathbf{V}^{\text{eff}} : \underline{\underline{\Sigma}} + \int_0^t \mathbf{Y}(t-\tau) : \underline{\underline{\dot{\Sigma}}}(\tau)\mathrm{d}\tau \tag{5.134}$$

which differs from a Maxwellian behaviour by the convolution integral term with the kernel $\mathbf{Y}(t-\tau)$: this term expresses the *"long range memory effect"* which was also responsible for the continuous part of the relaxation spectrum in the foregoing example of a two-phase material. It is also responsible for the need of two potentials to define a viscoelastic behaviour and for the associated difficulty to conceive variational approaches and bounding techniques for the overall behaviour.

Acknowledgments The authors are particularly grateful to Professor Jacques Verdu (École nationale supérieure des Arts et Métiers ParisTech, Paris) for his help in writing the entire subsection devoted to polymers, as well as part of Annex 1.

References

Bellenger V, Verdu J, Morel E (1987) Effect of structure on glass transition temperature of amine crosslinked epoxies. J Polym Sci B 25:1219–1234

Biot M (1954) Theory of stress-strain relations in anisotropic viscoelasticity and relaxation phenomena. J Appl Phys 25:1385–1391

Biot M (1956) Thermoelasticity and irreversible thermodynamics. J Appl Phys 27:240–253

Blanter MS, Golovin IS, Neuhäuser H, Sinning H-R (2007) Internal friction in metallic materials. Springer, New York

Boltzmann L (1874) Zur Theorie der elastischen Nachwirkung. Sitzber Kaiserl Akad Wiss Wien, Math-Naturw Kl 70:275–306

Brenner R, Masson R, Castelnau O, Zaoui A (2002) A quasi-elastic affine formulation for the homogenised behaviour of nonlinear viscoelastic polycrystals and composites. Eur J Mech A/Solids 21:943–960

Di Marzio EA (1964) On the second order transition of a rubber. J Res NBS 68A:611–617

Ferry JD (1970) Viscoelastic properties of polymers. Wiley, New York

Flory P (1953) Principles of polymer chemistry. Cornell U. Press, Ithaca

Fox TG, Flory PJ (1954) The glass temperature and related properties of polystyrene. Influence of molecular weight. J Polym Sci Polym Lett 14:315–319

Gibbs JH, Di Marzio EA (1958) Nature of the glass transition and the glassy state. J Chem Phys 28:373–383

Gilbert DG, Ashby MF, Beaumont PWR (1986) Modulus-maps for amorphous polymers. J Mater Sci 21:3194

Granato AV, Lücke K (1956) Application of dislocation theory to internal friction phenomena at high frequencies. J Appl Phys 27:789–805

Grandini CR, Almeida LH, Niemeyer TC, Nogueria RA, Florêncio O (2005) In: Proceedings of the 1st international conference on diffusion in solids and liquids, University of Aveiro, pp 279–282

Harriliak S Jr, Nagami S (1967) A complex plane representation of dielectric and mechanical relaxation processes in some polymers. Polymer 8:161

Kê TS (1947) Experimental evidence of viscous behavior of grain boundaries in metals. Phys Rev 71:533–546

Kovacs AJ (1963) Transition vitreuse dans les polymères amorphes. Etude phénomènologique. Adv Polym Sci 3:394–541

Lakes RS (2009) Viscoelastic materials. Cambridge University Press, New York

Laws N, McLaughlin R (1978) Self-consistent estimates for the viscoelastic creep compliances of composite materials. Proc R Soc Lond Ser A 359:251–273

Lee EH (1955) Stress analysis in viscoelastic bodies. Quart Appl Math 13:183–190

Lenz E, Dahl W (1974) Determination of dissolved carbon and nitrogen using a torsion pendulum lamellar specimen. Arch Eisenhütten 45:541–544

Mandel J (1955) Sur les corps viscoélastiques à comportement linéaire. C R Acad Sci 241: 1910–1912

Mandel J (1966) Mécanique des milieux continus, vol 2. Gauthier Villars, Paris

Perez J (1992) Physique et mécanique des polymères amorphes. TEC and DOC Lavoisier, Paris

Rougier Y, Stolz C, Zaoui A (1993) Représentation spectrale en viscoélasticité linéaire des matériaux hétérogènes. C R Acad Sci II b 316:1517–1522

Salençon J (2009) Viscoélasticité pour le calcul des structures. Editions de l'Ecole Polytechnique, Palaiseau

Schapery RA (1962) Approximate methods of transform inversion for viscoelastic stress analysis. In: Proceedings of the 4th US National Congress on Applied Mechanics, vol 2, Berkeley, pp 1075–1085

Schapery RA (1965) A method of viscoelastic stress analysis using elastic solutions. J Franklin Inst 279:268–289

Schultz H (1991) Defect parameters in b c c metals: group-specific trends. Mater Sci Eng A 141:149–167

Sperling LH (1992) Introduction to physical polymer science. Wiley, New York

Struik LCE (1978) Physical ageing in amorphous polymers and other materials. Elsevier, Amsterdam

Suquet P (1987) Elements of homogenization for inelastic solid mechanics. In: Sanchez-Palencia E, Zaoui A (eds) Homogenization techniques for composite media Lecture notes in physics, vol 272. Springer, New York, pp 193–278

Treloar LRG (1975) The physics of rubber elasticity. Clarendon Press, London

Weller M (1985) Anelastic relaxation of interstitial foreign atoms and their complexes with intrinsic defects in BCC metals. J Phys 46(C10):7–14

Williams ML, Landel RF, Ferry JD (1955) The temperature dependence of relaxation mechanisms in amorphous polymers and other glass-forming liquids. J Am Chem Soc 77:3701

Wu S (1992) Secondary relaxation, brittle-ductile transition temperature and chain structure. J Appl Polym Sci 46:619–624

Appendix A
Annex 1: Atomic and Molecular Structures

Annex 1 gives background notions dealing with atomic and molecular structures in an abbreviated way, for the convenience of the user of the book. On the other hand, it is also easy to gather useful information on the web.

A1.1 Types of Bonds

The main types of chemical bonds are listed in Table A1.1.

Table A1.1 Types of bonds

Type of bonds	Mechanism	Order of magnitude (kJ/mole)
Covalent	Shared electrons	10^2
Metallic	Free electrons cloud	10^2
Ionic	Electrostatic attraction	10^2
Van der Walls	Molecular attraction	10^{-1}
Hydrogen bond	Dipoles attraction	1

Adding a repulsive term to the attractive one gives the usual expression for the energy:

$$U = \frac{B}{r^m} - \frac{A}{r^n} \tag{A1.1}$$

(A, B positive; B is *Born's constant*[1])
where r is the distance between the atoms

m is of the order of 10
$n = 1$ for ionic bonds, $= 6$ for van der Waals bonds

For an ionic crystal the attractive force is qq'/r^2, where q, q' are the charges on the ions.

[1] Max Born (1882–1970), Nobel Prize winner, was a German physicist.

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

For NaCl, $A = \mu e^2$, where e is the charge on the electron and $\mu = 1.7475$ is *Madelung's constant*[2].

A1.2 Crystalline Solids – Elements of Crystallography

A1.2.1 Symmetry Groups

Figure A1.1 shows the elements of symmetry and the corresponding *Hermann-Mauguin symbols*[3], an integer for axes of symmetry and m for a mirror plane. The notation 2/m corresponds to common axis and normal to the mirror plane.

The following operations are identical

$$\bar{2} \equiv 2 \times \bar{1}$$

$$\bar{3} \equiv 3 \times \bar{1}$$

$$\bar{4} \equiv 4 \times \bar{1}$$

$$\bar{6} \equiv 6 \times \bar{1} \equiv 3 \times \bar{2} \equiv 3\mathrm{m}$$

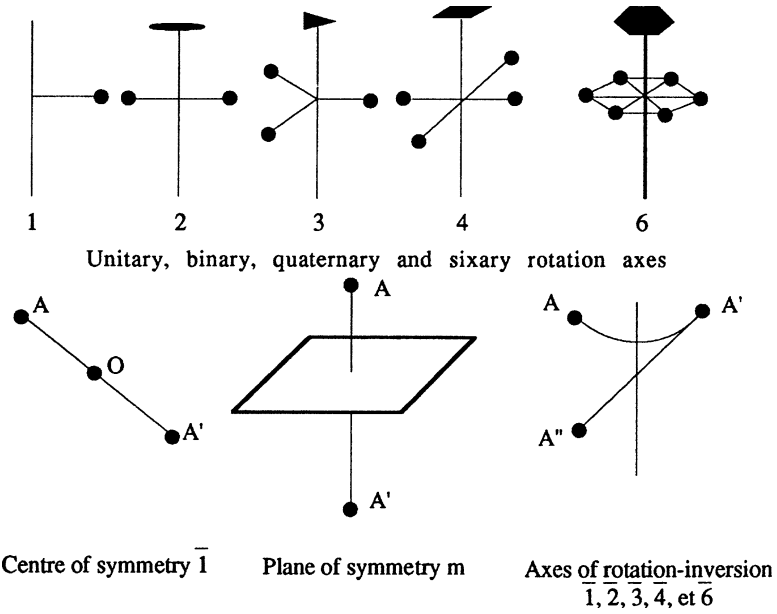


Fig. A1.1 Point groups of symmetry and the corresponding Hermann-Mauguin symbols

[2]Erwin Madelung (1881–1972) was a German physicist.

[3]Charles Victor Mauguin (1878–1958) was a French mineralogist; Carl Hermann (1898–1961) was a German mineralogist.

A1.2.2 Crystallographic Systems

The crystallographic systems are listed in Table A1.2.

Table A1.2 Crystallographic systems (Barrett and Massalski 1988)

System	Characteristics	Symmetry element	Hermann-Mauguin symbol (32 point groups)	Examples
Triclinic	Three unequal axes, no pair at right-angles $a \neq b \neq c$, $\alpha \neq \beta \neq \gamma \neq 90^o$	None	1	K_2CrO_7
Monoclinic	Three unequal axes, one pair not at right angles $a \neq b \neq c$ $\alpha = \gamma = 90° \neq \beta$	One binary axis of rotation or one mirror plane	2, $\bar{2}$ (= m) 2/m	Sβ, $CaSO_4 \cdot 2H_2O$ (gypsum)
Orthorhombic	Three unequal axes, all at right angles $a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$	3 orthogonal binary axes of rotation or 2 perpendicular mirror planes	222, 2 mm 2/m2/m2/m	Sα, Uα, Ga Fe_3C (cementite)
Tetragonal	Three axes at right angles, two equal $a = b \neq c$, $\alpha = \beta = \gamma = 90°$	One quaternary axis of rotation or of rotation-inversion	4, $\bar{4}$, 422, 4 mm, $\bar{4}$2m 4/m, 4/m2/m2/m	Snβ (white) TiO_2
Cubic	Three equal axes, all at right angles $a = b = c$, $\alpha = \beta = \gamma = 90°$	4 ternary axes of rotation	23, 432, $\bar{4}3$ /m 2 /m $\bar{3}$, 4/m$\bar{3}$2/m	Cu, Ag, Au, Fe NaCl
Hexagonal	Three equal coplanar axes at 120º, a fourth orthogonal to the plane $a1 = a2 = a3 \neq c$, $\alpha = \beta = 90°$ $\gamma = 120°$	One 6-ary axis of rotation or of rotation-inversion	6, $\bar{6}$, 6 mm, $\bar{6}$m2, 6/m, 6/m2/m2/m	Zn, Mg, Ti NiAs
Rhombohedric	Three equal axes, angles equal but not right angles $a = b = c$, $\alpha = \beta = \gamma \neq 90°$	One ternary axis of rotation or of rotation-inversion	$\bar{3}$, 32, 3 m $\bar{3}$,$\bar{3}$2 /m	As, Sb, Bi Calcite

Structural types.

The following nomenclature is used in the *Strukturbericht*:

A simple elements
B AB compounds
C AB_2 compounds
D A_mB_m composites
L alloys
O organic compounds
S silicates

A1 materials are FCC; A2 are BCC, A3 are CPH, A4 are diamond cubics,

The most common structures in metallic materials are the *face-centred cubic* (FCC), the *body-centred cubic* (BCC) and the *close-packed hexagonal* one (CPH). Face-centred cubic (FCC) and close-packed hexagonal (CPH) are compact structures that can be created by stacking hard spheres, as in Figs. A1.2 and A1.3.

In the FCC structure the packing is PQRPQR, in the CPH it is PQPQP.

In these structures the *insertion site* is at the centre of the tetrahedron formed by the stacked spheres at position (1/4, 1/4, 1/4)a and of radius $(1/4)(\sqrt{3}-\sqrt{2})a = 0.079a = 0.112r_0$, a being the lattice parameter and r_0 the inter-atomic distance.

In BCC there are two sites:

tetrahedral (1/2, 1/8, 1/8)a, radius $(1/8)(3\sqrt{2}-2\sqrt{3})\text{a} = 0.097\text{a} = 0.112r_0$
octahedral (1/2, 0, 0)a, radius $(1/4)(2-\sqrt{3})a = 0.067a = 0.077r_0$, in term of the inter-atomic distance r_0.

A1.2.3 Ordered Structures

Long-range order. The degree of order S is defined by $S = (p - r)/(1 - r)$, where p is the probability that a site that should be occupied by an atom A is in fact occupied by an atom A, and r is the fraction of sites occupied by atoms A when the order is perfect. S varies between 0 for complete disorder and 1 for perfect order.

Short-range order. The degree of order σ is defined as the difference between the probability of finding a different atom adjacent to a given atom and that of finding an atom of the same kind (Table A1.3).

A1.2.4 Miller Indices[4]

Direction: [uvw] denotes the direction of the vector with co-ordinates u, v, w in terms of the parameters of the lattice

[4]William Hallowes Miller (1801–1880) was a British mineralogist and crystallographer.

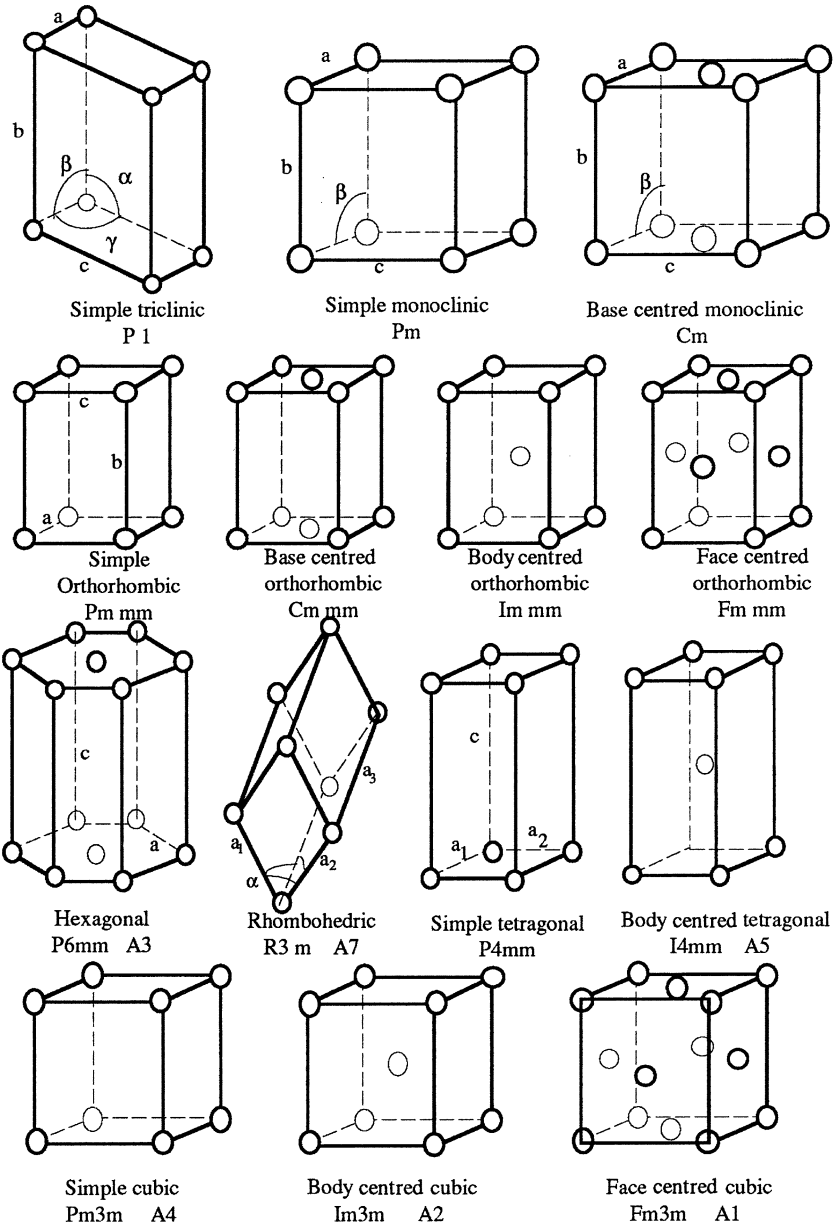


Fig. A1.2 The 14 fundamental Bravais[5] crystal lattices: types of structure and Hermann-Mauguin indices (Hermann 1949, Steuer 1993)

[5] Auguste Bravais (1811–1863) was a French physicist.

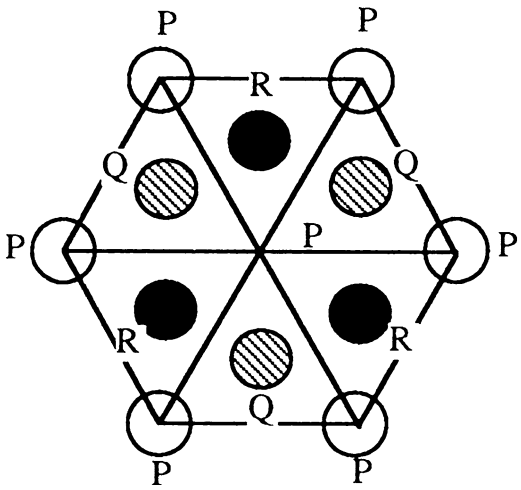
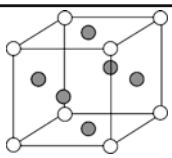
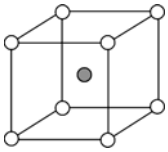
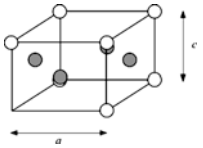
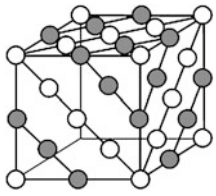


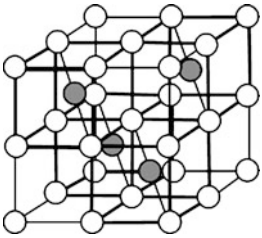
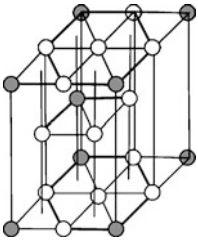
Fig. A1.3 Close packing: the dense planes denoted P, Q, R, are projected one on the other to show the successive positions of the atoms

Table A1.3 Ordered structures

Type	Sketch	Examples
$L1_2$ or Cu_3AuI		Cu_3Au, $AlCo_3$, $AlZr_3$ $FeNi_3$, Ni_3Al
B2 or brass β		CuZnβ, AlNiβ, NiZnβ
$L1_0$ or CuAuI		AlTi, AuCuI, CuTiδ, FePt, NiPt
$L1_1$ CuPt		or CuPt

(continued)

Table A1.3 (continued)

Type	Sketch	Examples
DO_3 or $L2_1$		AlCuβ, $AlFe_3$, Cu_3Sbβ, Fe_3Siα
L2 structures are ferromagnetic DO_{19} or Mg_3Cd Analogous to L1 but consisting of 4 hexagonal sub-lattices		Cd_3Mg, Mg_3Cd, Ni_3Snβ, Ni_3Nb

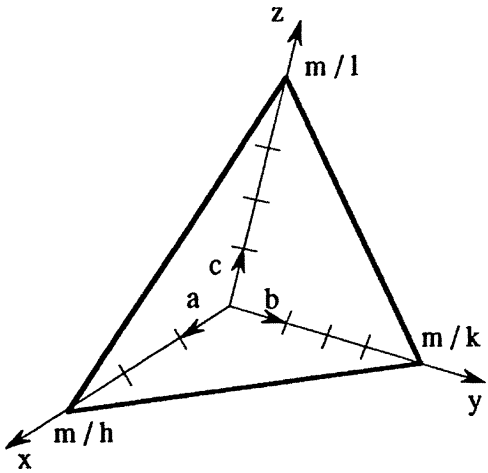


Fig. A1.4 Miller indices for a plane

Plane: (*hkl*) denotes the plane whose intercepts on the lattice axes are *m*/*h*, *m*/*k*, *m*/*l*, where m is chosen so that *h*, *k*, and *l* are the smallest possible integers (Fig. A1.4).

{*hkl*} means the set of all planes with indices $|h|$, $|k|$, $|l|$.
<*uvw*> means the set of all directions with indices $|u|$, $|v|$, $|w|$.

For CPH, it is customary to use a 4-index system, taking 3 axes at 120° in the basal plane: a direction is denoted by [*uvtw*], with $u + v + t = 0$

a plane is denoted by (*hkil*), with $h + k + i = 0$

In this system it is equally practical to use *orthohexagonal axes*, such that $a = a_1$, $b = a_1 + 2a_2$. With these, a direction $[pqr]$ is such that $[uvtw] = [p+q, 2q, -p-3q, r]$ and a plane (efg) is such that $(hkil) = (e, (f-e)/2, -(f+e)/2, g)$.

In cubic systems a direction $[hkl]$ is perpendicular to the plane (hkl).

A1.2.5 Reciprocal Lattice

The reciprocal lattice $\underline{a}^*, \underline{b}^*, \underline{c}^*$ of lattice $\underline{a}, \underline{b}, \underline{c}$ is defined by the relation:

$$\left[\underline{a}^*\underline{b}^*\underline{c}^*\right]^{\mathrm{T}} = \left[\underline{a}\underline{b}\underline{c}\right]^{-1} \tag{A1.2}$$

which can be written as:

$$\underline{a}^* = \frac{\underline{b} \wedge \underline{c}}{\underline{a} \cdot \left(\underline{b} \wedge \underline{c}\right)}$$

$$\underline{b}^* = \frac{\underline{c} \wedge \underline{a}}{\underline{b} \cdot \left(\underline{c} \wedge \underline{a}\right)}$$

$$\underline{c}^* = \frac{\underline{a} \wedge \underline{b}}{\underline{c} \cdot \left(\underline{a} \wedge \underline{b}\right)} \tag{A1.2bis}$$

Thus:

$$\underline{a}^* \cdot \underline{b} = \underline{a}^* \cdot \underline{c} = \underline{b}^* \cdot \underline{c} = \underline{b}^* \cdot \underline{a} = \underline{c}^* \cdot \underline{a} = \underline{c}^* \cdot \underline{b} = 0$$

$$\underline{a}^* \cdot \underline{a} = \underline{b}^* \cdot \underline{b} = \underline{c}^* \cdot \underline{c} = 1 \tag{A1.3}$$

Each point *h, k, l,* in the reciprocal lattice corresponds to a set of (hkl) planes in the real space lattice. The vector *h, k, l,* in the reciprocal lattice is perpendicular to the planes (hkl) in the real space lattice.

The distance between two (hkl) planes is:

$$d_{hkl} = \frac{1}{\left|\underline{d}^*_{hkl}\right|} = \frac{1}{\left|h\underline{a}^* + k\underline{b}^* + l\underline{c}^*\right|} \tag{A1.4}$$

In a cubic crystal: $\underline{a}^* = \underline{a}\,/a^2$, $\underline{b}^* = \underline{b}\,/a^2$, $\underline{c}^* = \underline{c}\,/a^2$.

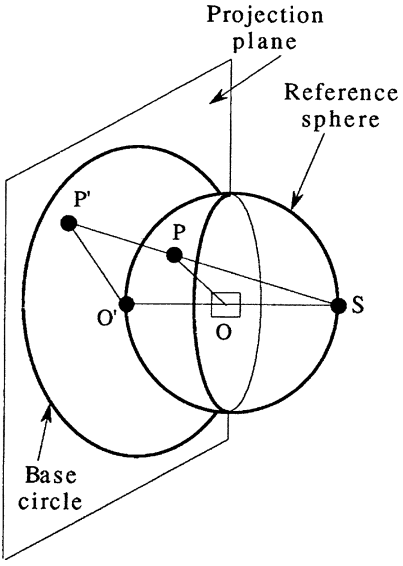


Fig. A1.5 Stereographic projection

A1.2.6 Stereographic Projection

P is the *pole* of the plane, that is, the intersection of the normal to the plane with the reference sphere. The main properties of this projection (see Fig. A1.5) are as follows:

1. The projection on to the sphere of a circle of centre P is a circle whose centre is different from the projection P′ of P
2. Great circles on the sphere project into circles intersecting the base circle (the projection of the equator) in two diametrically opposite points
3. The angle between two poles can be measured if they are on the same meridian
4. The angle between two poles is not changed by rotation about the axis of projection.
5. If R is the radius of the base circle, the distance between its centre and the projection of a pole making an angle φ with the axis of projection is $R \tan(\varphi/2)$.

Figure A1.6 is a Wulff[6] net, the projection of the meridians; Fig. A1.7 is the polar projection. If drawn on transparent paper they enable the angles between any pair of poles to be found by bringing them on to the same meridian.

Figure A1.8 is the stereographic projection of a cubic crystal; Fig. A1.9 that of a CPH crystal with c/a = 1.86 (c/a = 1.89 for Zn and Cu).

[6] Yuri Viktorovitch Wulff was a Russian mineralogist.

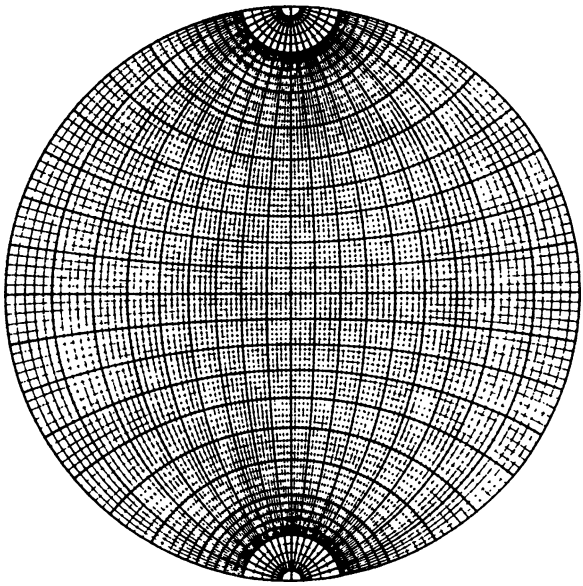


Fig. A1.6 Wulff net ($2^\circ \times 2^\circ$)

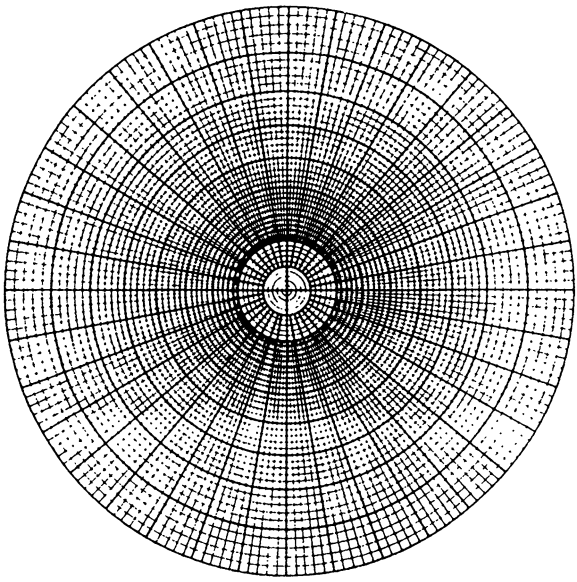


Fig. A1.7 Polar stereographic net

Direction $[uvw]$ is the *zone axis* of the family of planes (hkl) if $uh + vk + wl = 0$. It is the zone axis for two lattice planes $h_1k_1l_1$ and $h_2k_2l_2$ if it satisfies:

$$\frac{u}{\begin{vmatrix} k_1 & l_1 \\ k_2 & l_2 \end{vmatrix}} = \frac{v}{\begin{vmatrix} l_1 & h_1 \\ l_2 & h_2 \end{vmatrix}} = \frac{w}{\begin{vmatrix} h_1 & k_1 \\ h_2 & k_2 \end{vmatrix}} \tag{A1.5}$$

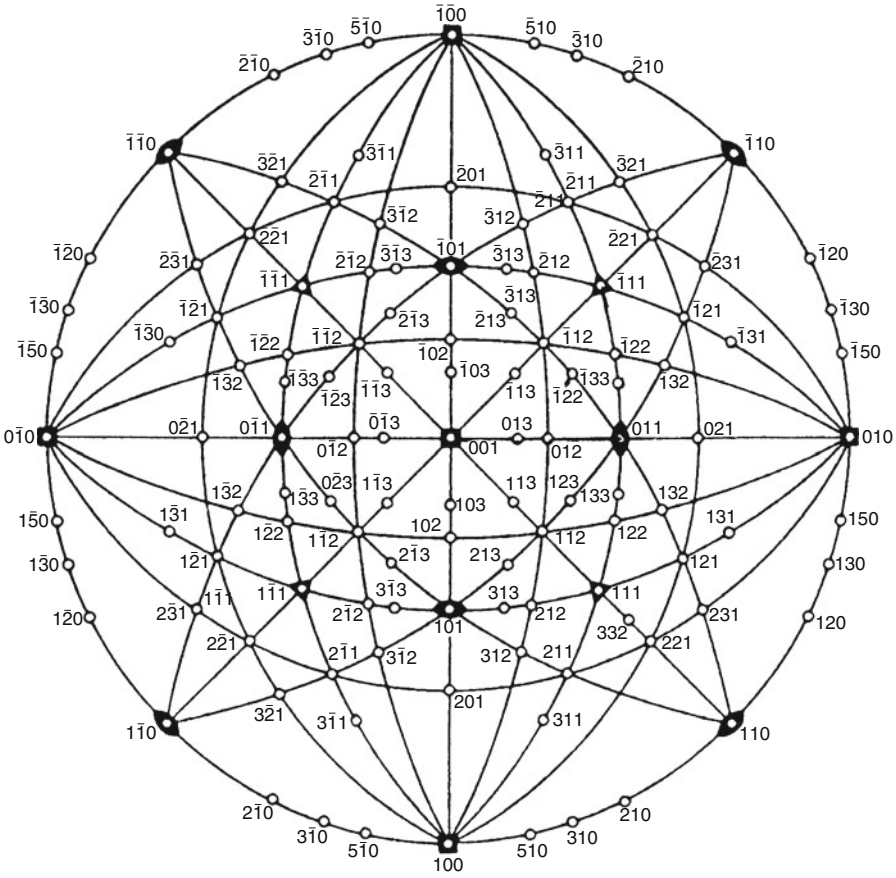


Fig. A1.8 Standard (001) projection of the poles and zone circles for a cubic crystal

Three lattice planes are *in zone* if:

$$\begin{vmatrix} h_1 & k_1 & l_1 \\ h_2 & k_2 & l_2 \\ h_3 & k_3 & l_3 \end{vmatrix} = 0 \tag{A1.6}$$

A1.2.7 Twinning

A *twin* is a polycrystalline structure formed by putting together two or more pieces of material of the same crystallographic structure, assembled according to well-defined laws. We distinguish between

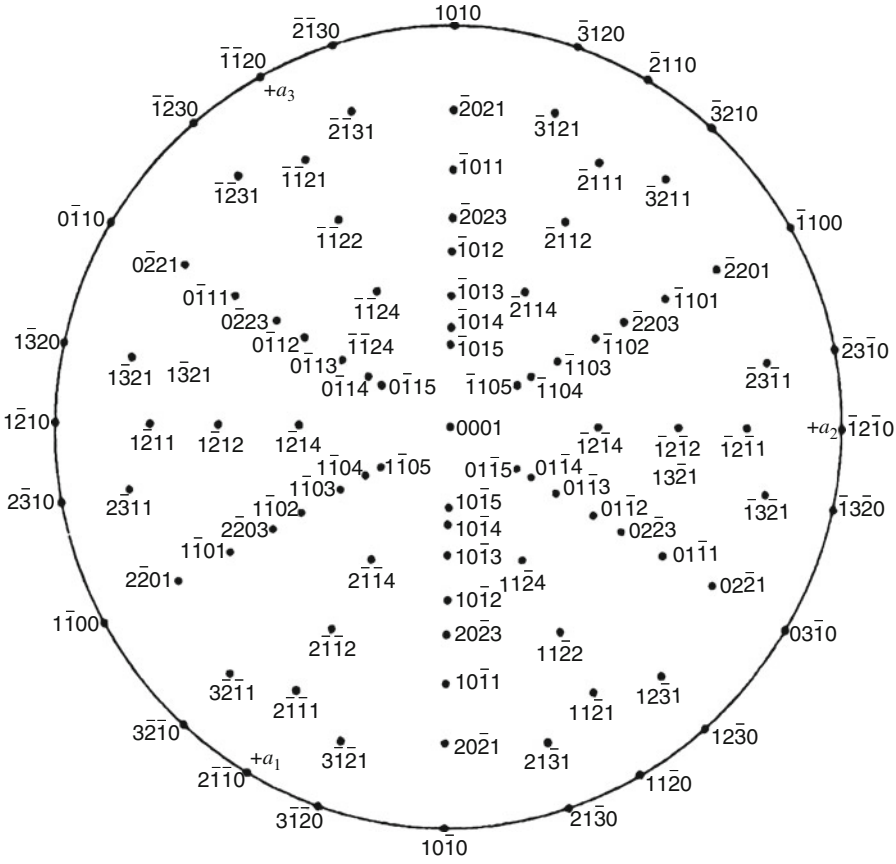


Fig. A1.9 Standard (0001) projection for CPH zinc (c/a = 1.86)

– *grown-in twins* formed during solidification
– *recrystallisation* twins
– *mechanical twins* resulting from shearing.

Figure A1.10 shows the *twinning elements*, in which the two crystallographic planes K_1, K_2 are unchanged. The twinning results from a shear parallel to the direction η_1 of K_1; this plane and this direction are called the *twinning plane* and *twinning direction* respectively. The plane of the shear is normal to K_1 and contains the direction η_1, intersecting K_2 in a line whose direction is η_2. The twinning causes K_2 and η_2 to rotate to the new orientations K_2' and η_2'.

The shear is such that:

$$\gamma = s/h = 2\tan(\pi/2 - \alpha) \tag{A1.7}$$

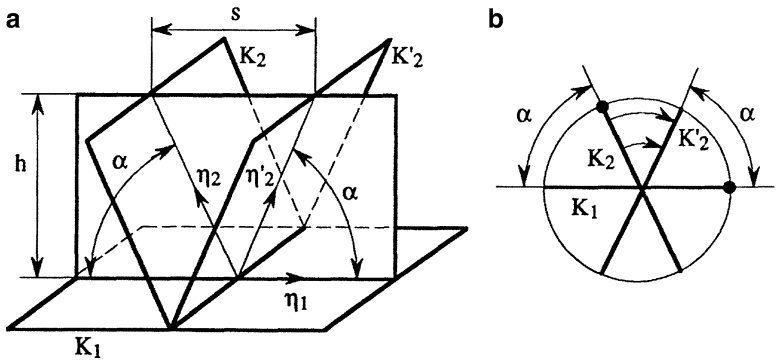


Fig. A1.10 Elements of twinning

Table A1.4 Main elements of twinning in metals

Structure	Twinning plane K_1	Twinning direction η_1	K_2	η_2	Shear *s/h*
BCC	{112}	$< 11\bar{1} >$	$\{11\bar{2}\}$	<111>	0.707
FCC	{111}	$< 11\bar{2} >$	$\{11\bar{1}\}$	<112>	0.707
CPH all	$\{10\bar{1}2\}$	$< \bar{1}011 >$	$\{\bar{1}012\}$	$< 101\bar{1} >$	Depends on c/a ratio
	$\{11\bar{2}1\}$	$< 11\bar{2}\bar{6} >$	{0001}	$< 11\bar{2}0 >$	
CPH some cases	$\{11\bar{2}2\}$	$< 11\bar{2}3 >$	$\{11\bar{2}\bar{4}\}$	$< 22\bar{4}3 >$	
Cubic diamond	{111}	$< 11\bar{2} >$	$\{11\bar{1}\}$	$< 112 >$	0.707
Tetragonal Sn β	{301}	$< \bar{1}03 >$	$\{\bar{1}01\}$	<101>	–
Orthorhombic U α	{130}	$< 3\bar{1}0 >$	$\{1\bar{1}0\}$	$< 110 >$	0.299
	Irrational	$< 312 >$	{112}	Irrational	0.228
	{112}	Irrational	Irrational	$< 3\bar{1}0 >$	0.228
	{121}	Irrational	Irrational	$< 311 >$	0.329
	Irrational	$< 5\bar{1}2 >$	$\{1\bar{1}1\}$	Irrational	0.214

Twins of the first kind are such that K_1 and η_2 have rational Miller indices. *Twins of the second kind* are such that K_2 and η_1 have rational Miller indices. To each type there corresponds a *conjugate* such that $\begin{cases} \bar{K}_1 = K_2 \\ \bar{\eta}_1 = \eta_2 \end{cases}$

NOTE The boundary between the twin and the mother crystal is not necessarily the twinning plane.

If the indices of K_1 and η_2 are (HKL), $[UVW]$, a direction $[uvw]$ becomes $[u'v'w']$ after the twinning, where $u' = u - 2U\beta$, $v' = v - 2V\beta$, $w' = w - 2W\beta$, with $\beta = (Hu + Kv + Lw)/(HU + KV + LW)$.

Similarly the plane (hkl) becomes $(h'k'l')$, where $h' = h - 2H\alpha$, $k' = k - 2K\alpha$, $l' = l - 2L\alpha$, with $\alpha = (Uh + Vk + Wl)/(UH + KV + LW)$ (Tables A1.4 and A1.5).

Table A1.5 Values of the twinning shear for CPH according to the c/a ratio

Element	Cd	Zn	Mg	Zr	Ti	Be
c/a	1.886	1.856	1.623	1.592	1.587	1.568
s/h	0.175	0.143	−0.131	−0.169	−0.175	−0.186

A1.2.8 X-Ray Diffraction

A1.2.8.1 Diffraction Conditions

The *Laue's equation*[7] states the condition that the X-rays scattered by atoms are in phase, and is therefore the condition for diffraction. It is written:

$$\left(\underline{S} - \underline{S}_0\right) \cdot \underline{a} = n\lambda \tag{A1.8}$$

where $\underline{S}$, $\underline{S}_0$ are unit vectors in the directions of the diffracted and incident rays respectively, $\underline{a}$ is the vector joining two atoms, λ is the wavelength and n is an integer.

The *Bragg's law*[8] states the condition for diffraction by a crystallographic plane. It is written:

$$2d\ \sin\theta = n\lambda \tag{A1.9}$$

where d is the distance between reflecting planes, θ is the angle of incidence and, n and λ are as before.

These relations can be written:

$$\underline{S} - \underline{S}_0 = \underline{r}^*\lambda \tag{A1.10}$$

where $\underline{r}^*$ is the vector in the reciprocal lattice $\underline{r}^* = h\underline{a}^* + k\underline{b}^* + l\underline{c}^*$

A1.2.8.2 Coherent Scattered Intensity

For an un-polarised beam the *scattered intensity of X-rays by an electron* is, in SI units:

$$I_e = \frac{I_0 e^4}{r^2 m^2 c^4} \frac{1 + \cos^2 2\theta}{2} = \left(7.934 \times 10^{-30}\right) \frac{I_0}{r^2} \frac{1 + \cos^2 2\theta}{2} \tag{A1.11}$$

[7] Max von Laue (1879–1960) was a German physicist, winner of the Nobel prize.

[8] William Henry Bragg (1862–1942) and his son William Lawrence Bragg (1890–1971) were British physicists winners together of the Nobel prize.

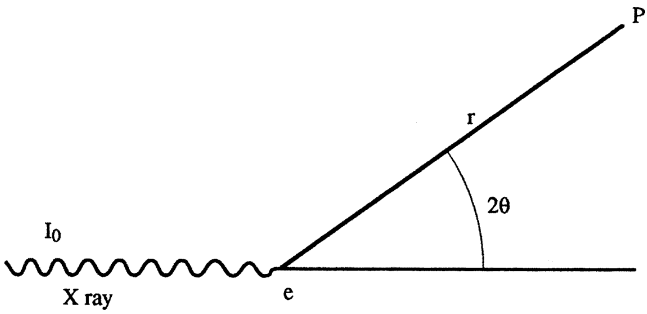


Fig. A1.11 Scattering of X-ray by an electron

where I_0 is the incident intensity, e the charge on the electron, c the velocity of light, r the distance and 2θ the scattering angle. The factor $(1 + \cos^2 2\theta)/2$ is the polarisation factor (Fig. A1.11).

Scattering by an atom. The *scattering factor f* is given by $f^2 = I_a/I_e$, the ratio of the scattered intensity for an atom to that for an electron. If θ is small the ratio tends towards the number Z of electrons in the atom.

At 0 K the atomic scattering factor is given by:

$$f_0 = \int_0^\infty U(r) \frac{\sin kr}{kr} \mathrm{d}r \tag{A1.12}$$

where $U(r)\,\mathrm{d}r$ is the number of electrons between r and $\mathrm{d}r$ from the centre of the atom, assumed spherical, and $k = 4\pi\ \sin\theta/\lambda$.

The *structure factor F* is the sum of the amplitudes of the waves diffracted by the plane (*hkl*); it is:

$$F = \sum_j f_j \exp\left[2\pi \mathrm{i}\left(hu_j + kv_j + lw_j\right)\right] \tag{A1.13}$$

where f_j is the diffraction factor for the atom j at the point (u_j, v_j, w_j) of the lattice. The diffracted intensity is F^2, given by the product of F with its complex conjugate.

If there is no diffraction, then $F^2 = 0$, which holds for centred lattices when $h + k + l$ is odd and for face-centred structures when h, k, l are not simultaneously odd or even (Table A1.6).

A1.2.8.3 Textures

We distinguish between fibre textures (threads, bars) and pole figures for more complex preferred orientations (sheets). Figure A1.12 shows the correspondence between the X-ray diagram and the stereographic projection; it follows from Bragg's law that all planes (*hkl*) that are able to diffract have their poles on a circle, called the reflection circle, at $(\pi/2 - \theta)$ from the central beam.

Table A1.6 Orders of diffracted rays for cubic structures ($Q = h^2 + k^2 + l^2 = (4a^2/\lambda^2)\sin^2\theta$)

Q	FCC	BCC
2	–	{110}
3	{111}	–
4	{200}	{200}
6	–	{112}
8	{220}	{220}
10	–	{310}
11	{113}	–
12	{222}	{222}
14	–	{123}
16	{400}	{400}
18	–	{411} {330}
19	{331}	–
20	{420}	{420}

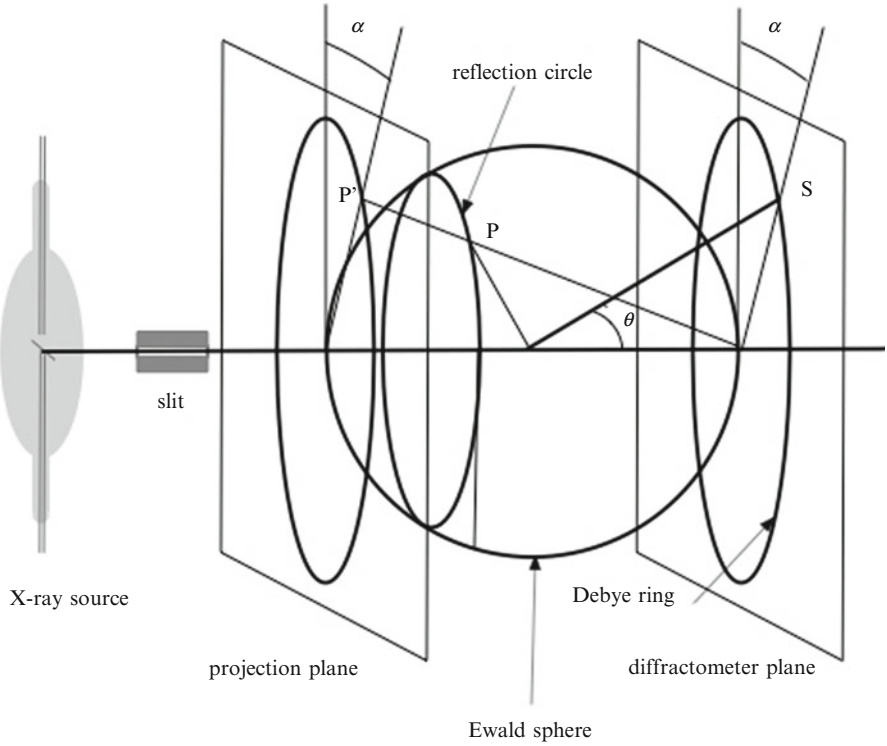


Fig. A1.12 Relation between the crystallographic plane, the diffracted beam and the stereographic projection

Consider an ideally textured fibre for which all the grains are aligned in a direction [*uvw*] with respect to the axis of the wire. Since all the poles of the planes (*hkl*) form an angle ρ with this direction they lie on the same line in the stereographic

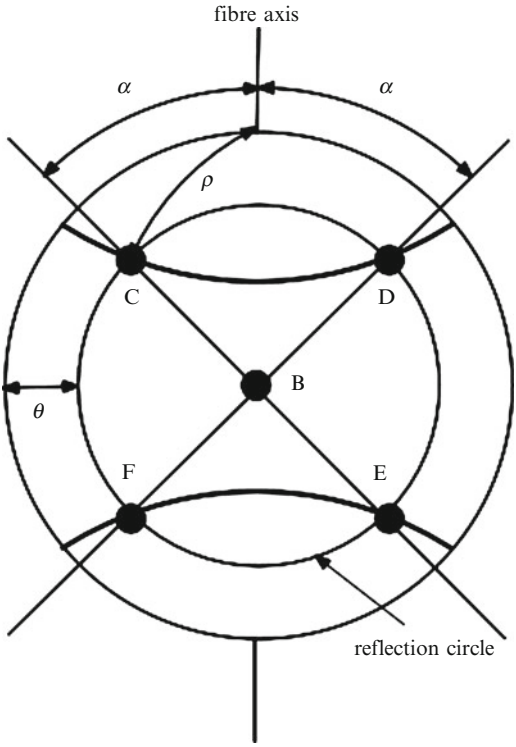


Fig. A1.13 Ideal pole figure for {111} planes in a cubic metal wire having a [111] fibre texture. The poles C, D, E, F diffract if the direction of the beam is B

projection, so the rays diffracted by these planes will be at the intersection of this line with the reflection circle: Fig. A1.13 shows the (100) reflections for the case of a [111] texture in a fibre of cubic structure. On the stereographic projection the pole figures give the density of the poles of particular (*hkl*) planes.

Figure A1.14 gives an example of the figure for the (111) poles for a brass sheet, RD being the rolling direction. The poles are seen to be grouped around the expected position Δ if the (110) plane is parallel to the rolling plane with direction $[1\bar{1}2]$ in the direction of rolling.

For fibre textures it is an advantage to use *inverse pole figures*: these give density distributions of some important direction – of the fibre axis, for example – on the stereographic projection of the crystal lattice in its standard orientation.

A1.2.8.4 Small-Angle Scattering

Any region of microscopic scale whose density differs from that of its environment will scatter X-rays in a way that is characteristic of its size, shape and number; and the values of these quantities for small diffraction angles can be deduced from the intensities of the scattered rays.

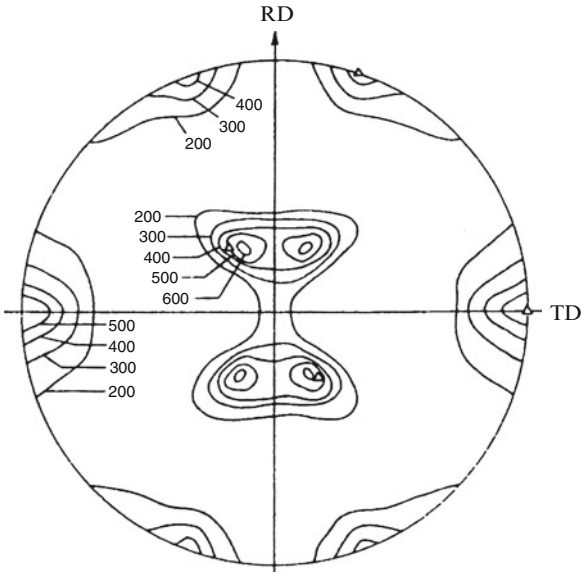


Fig. A1.14 (111) pole figure for rolled brass plate: $\Delta =$ (110) $[1\bar{1}2]$

A1.3 Polymers

A1.3.1 Chemical Structure

Polymers are constituted of macromolecules. These later are mainly constituted of linear segments resulting from the chaining of difunctional elementary (undeformable) groups. These groups can be ranged in two categories: "knee-joints" allowing rotations around skeleton bonds and "rigid rods" in which no rotation is allowed.

Branched and tridimensional polymers are formed of linear segments joined together by "crosslinks" which are groups of functionality strictly higher than 2 (Fig. A1.15).

Two important observations can be made concerning the side groups of knee-joints:

1. their electrical dissymmetry (polarity) is a key characteristic. Cohesion, refractive index, dielectric permittivity and hydrophilicity are for instance increasing functions of polarity. One can distinguish roughly three group families:
   - groups of low polarity, for instance – CH_2– and all the hydrocarbon groups
   - groups of medium polarity, for instance ethers, ketones and esters
   - groups of high polarity, especially those able to establish hydrogen bonds (alcohols, acids, amides).

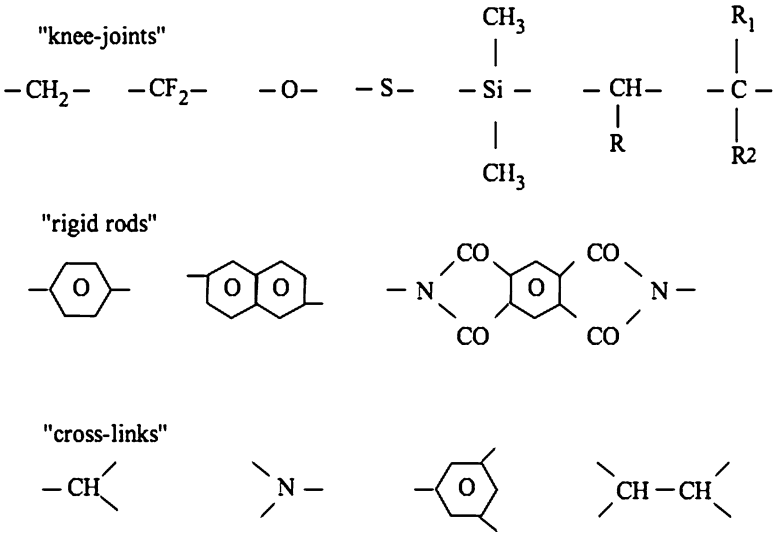


Fig. A1.15 Some common elementary groups of industrial polymers

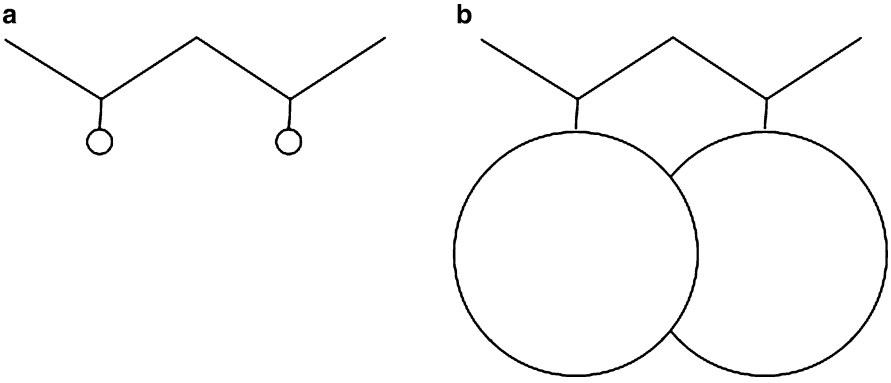
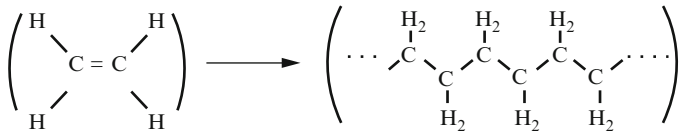


Fig. A1.16 Schematisation of a polymer having small (**a**) or bulky (**b**) side groups illustrating the steric hindrance effect

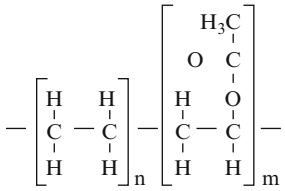
2. their bulkiness can induce steric hindrance and thus can more or less disfavor rotations of skeletal bonds (Fig. A1.16)

Macromolecules can be natural (cellulose, silk, rubber), artificial (cellulose triacetate, vulcanised rubber), or synthetic (the great majority of industrial polymers). In this case they are made of reactive molecules (monomers) able to connect one to another by covalent bonds through chains (polymerisation) or step-by-step (polycondensation) reactions.



monomer : ethene

polymer : polyethene
(polyethylene)



polymer : ethylene-vinyl acetate
(but-3-enoic; ethene)

Fig. A1.17 A homopolymer: polyethylene and a copolymer: polyethylene-vinyl acetate

Table A1.7 Energy and length of covalent bonds

	Single		Double		Triple	
Bond	Energy (kJ/mol)	Length (10^{-10} m)	Energy (kJ/mol)	Length (10^{-10} m)	Energy (kJ/mol)	Length (10^{-10} m)
C–C	334	1.54	589	1.34	836	1.21
N–N	88	1.46	267	1.25	710	1.10
O–O	146	1.49	489	1.21	–	–
C–N	234	1.47	468	1.28	673	1.16
C–H	395	1.1	–	–	–	–

A1.3.2 Structural Arrangements

A1.3.2.1 Chain Configuration

A polymer consists of chains molecules, formed by the polymerisation of one or more monomers (Fig. A1.17), the chaining resulting from the juxtaposition of covalent bonds in various groups. In a *homopolymer* there is only one type of monomer, in a *co-polymer* there are two types, possibly more. Figure A1.17 shows the homopolymer polyethylene and the copolymer polyethylene-vinyl acetate.

Table A1.7 gives *the energy and the length of covalent bonds*, which link together polymeric molecular chains.

Hydrogen bonds, with energy of 20–40 kJ/mol, hold H to electronegative atoms like O, N, F, S. The bonds can be intra- or inter-molecular.

The molecular chains are held together by van der Waals interactions, with an energy of about 2.5–4 kJ/mol.

Table A1.8 Cohesive energy density for some industrial polymers (Van Krevelen 1993)

Polymer	Acronym	d_c (MPa)
Polytetrafluorethylene	PTFE	165
Polyvinylidene fluoride	PVDF	190
Polydimethylsiloxane	PDMS	210
Polyethylene	PE	260
Polypropylene	PP	290
Poly(vinylchloride)	PVC	390
Polystyrene	PS	410
Polyoxymethylene	POM	440
Poly(etheretherketone)	PEEK	525
Poly(ethyleneterphthalate)	PET	540
Poly(vinylalcohol)	PVAL	1,100
Polyamide 6	PA6	1,110

The *cohesive energy* E_{coh} is the sum of secondary bonds in the molar volume V of the polymer constitutive repeat unit (CRV). The *cohesive energy density* d_c is defined as E_{coh}/V. Some values of d_c are given in Table A1.8. The highest cohesive energy densities are due to H bonds.

Monomers within a copolymer can be organised along the backbone in a variety of ways:

– *Alternating copolymers* possess regularly alternating monomer residues:

$$[AB\ldots]_n\text{: – A – B – A – B – A – B–}$$

– *Periodic copolymers* have monomer residue types arranged in a repeating sequence: $[A_nB_m\ldots]$, m being different from n:

– A – A – A – B – B – B – B – B – A – A – A – B – B – B – B – B

– A – A – A–

– *Statistical copolymers* have monomer residues arranged according to a known statistical rule. A statistical copolymer in which the probability of finding a particular type of monomer residue at a particular point in the chain is independent of the types of surrounding monomer residue may be referred to as a truly *random copolymer*.
– *Block copolymers* are obtained by the sequential addition of two or more homopolymer subunits linked by covalent bonds.
– *Graft or grafted copolymers* contain side chains that have a different composition or configuration than the main chain (Fig. A1.18).

Tacticity refers to the orientation of the molecular units (Fig. A1.19). In *isostatic* polymers all the substituents are oriented on the same side of the backbone molecular chain (isotactic polypropylene is the most important industrial application).

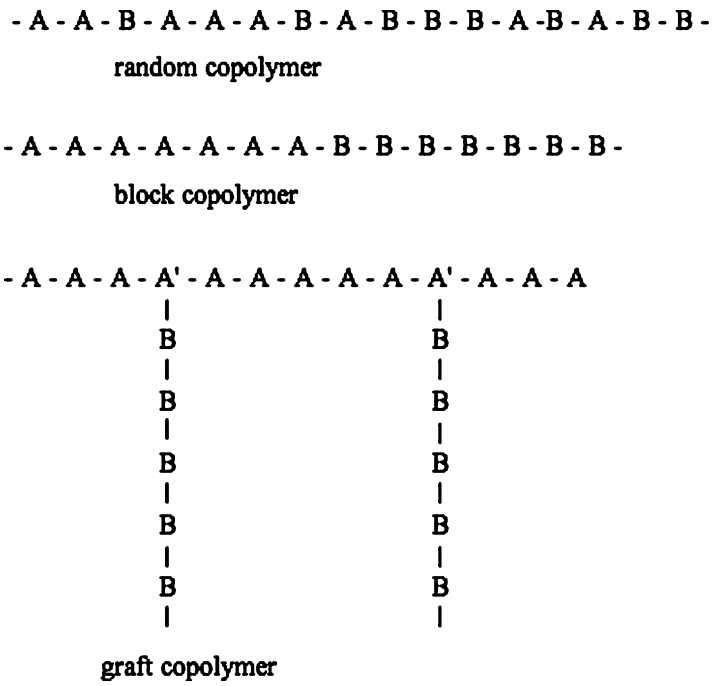


Fig. A1.18 Main types of copolymers

In *syndiotactic* polymers the substituents are alternatively on either sides of the backbone chain. In *atactic* polymers the substituents orientations on the side of the backbone chain are random.

A1.3.2.2 Architecture of the Chains

The simplest is a *linear chain*: a single backbone with no branching. In a *branched polymer* side chains are linked to the main backbone chain. A *star polymer* has branches linked together at a single end. In *brush polymers* the chains are attached to an interface. In *dendronised polymers*, dendrons, which are tree-like regularly branched chains, are attached to the main backbone chain. When the dendrons are attached together at the same end, the sphere-like polymer is a *dendrimer* (Fig. A1.20).

By creating covalent bonds between molecular chains is obtained a *cross-linked polymer.* A typical example is vulcanised rubber. Sufficiently high cross-linking can lead to the formation of an infinite network, a *gel*.

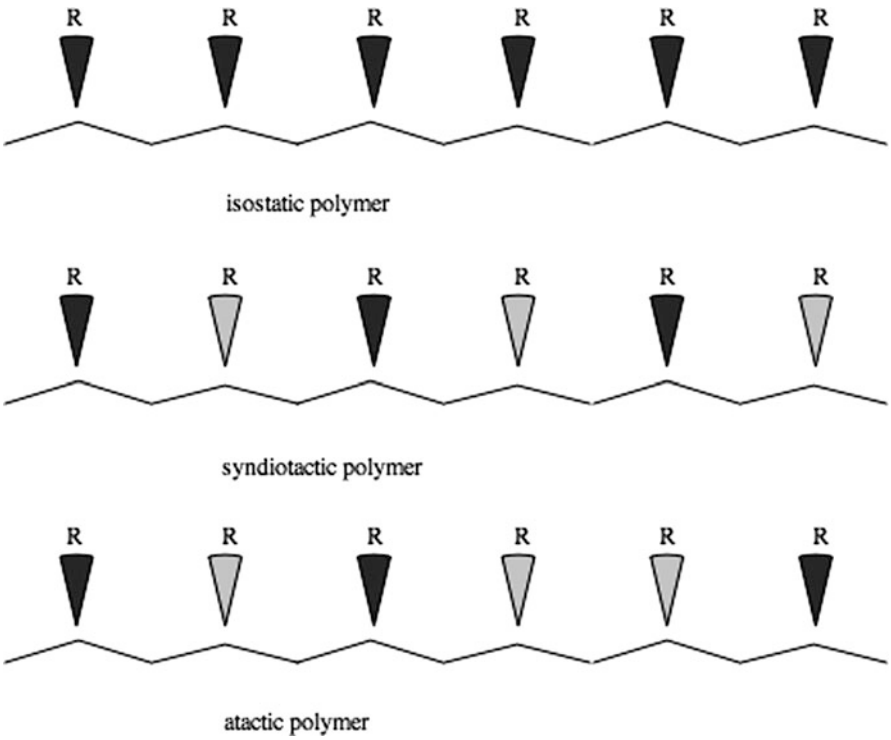


Fig. A1.19 Isotatic, syndiotactic and atactic polymers. *Dark pies* indicate a substituent pointing out to the front of the plane of the figure; *light pies* indicate substituents pointing out to the back of the plane of the figure

Networks

An ideal network is a network in which all the chains are linked. The cross-link density is often assimilated to the chain density. In a non-ideal network, there are cycles (two chains have the same extremities) or dangling chains (chains with one free end).

Entanglements in Linear Polymers

Above a critical molar mass M_c the polymers are entangled and can be considered as topological networks. The chains are able to disentangle by *reptation* (de Gennes[9] 1979). The molar mass between entanglements can be determined from the shear modulus in molten state, using the rubber elasticity theory (Chap. 2) (Fetters et al. 1999).

[9]Pierre-Gilles de Gennes (1932–2007) was a French physicist and the winner of the Nobel prize in 1991.

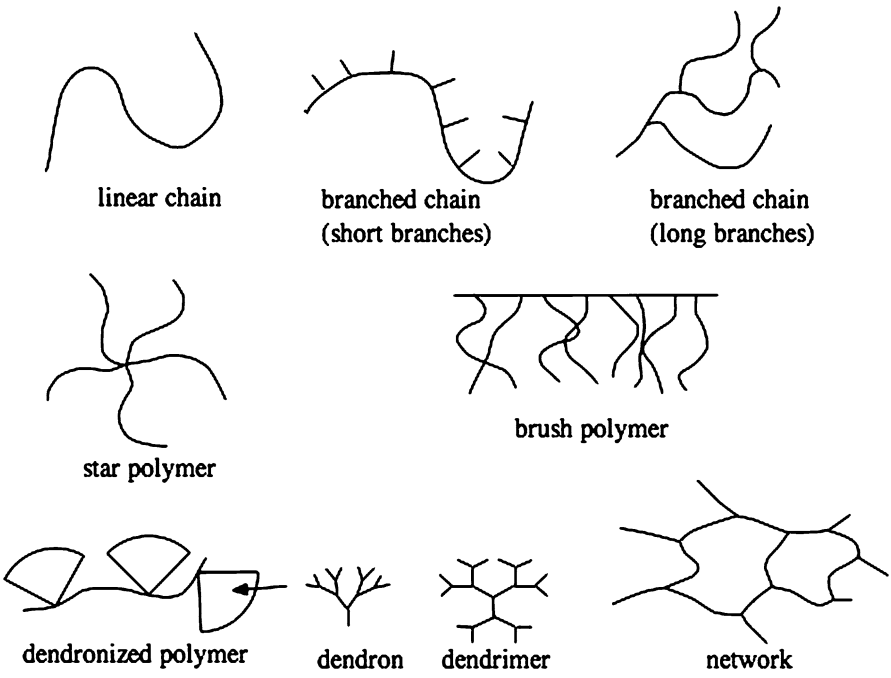


Fig. A1.20 Various architectures of polymer chains

Chain Conformation in Amorphous Polymers

Figure A1.21 shows the three conformations of a carbon-carbon bond in a vinyl polymer – $(CH_3–CHR)_n–$. The bond distances and the valence angles are fixed but the carbons can rotate more or less easily and the system displays three minima of potential energy (Fig. A1.22). In the trans conformation the bond between the carbons under consideration and the bonds linking these carbons to the next carbon chain atoms are coplanar. A projection in the plane perpendicular to the carbon-carbon bond leads to Fig. A1.22. The all trans conformation is a rigid plane zig-zag. Trans and gauche configurations can coexist. The length of trans-trans sequences is called persistence length. Three distinct situations are schematised in Fig. A1.23.

A1.3.2.3 Crystallinity

Some molecular chains can be folded or stacked together with other chains so as to form locally a crystalline structure. The polymer includes then crystalline regions within the amorphous structure. The proportion of such crystals is the *degree of crystallinity*. It can be determined from density or enthalpy measurements, or also by spectroscopy. In certain cases the polymer can be entirely crystalline.

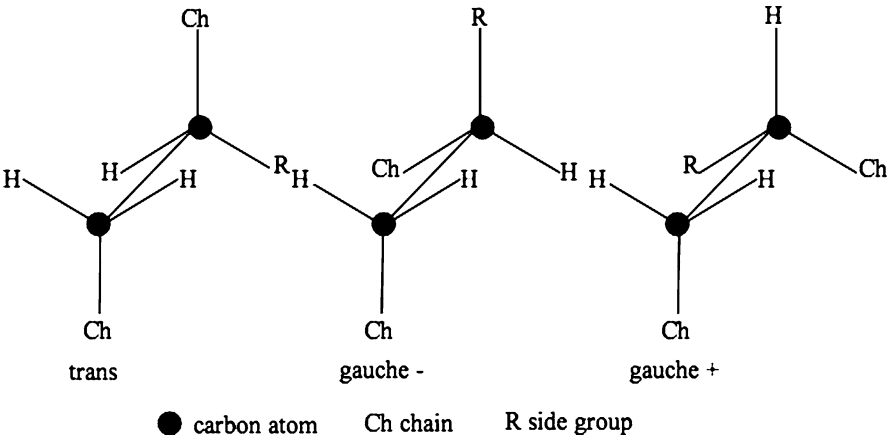


Fig. A1.21 Representation of three conformations of a carbon-carbon bond in a vinyl polymer

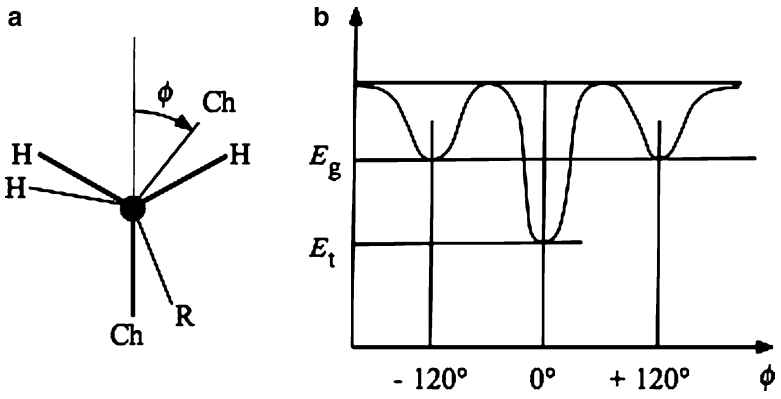


Fig. A1.22 Conformation of a carbon-carbon bond in a vinyl polymer. (**a**) Newman's representation and (**b**) shape of the variation of the potential energy with the angle ϕ

Three types of polymers can be distinguished:

1. polymers that crystallise easily. They have a symmetric structure and an aliphatic (flexible) skeleton (polyethylene – $CH_2)_n$–, polyoxymethylene – $(O–CH_2)_n$–)
2. polymers having a regular structure but with a slow crystallisation linked to the monomer asymmetry or to high chain stiffness. Their glass transition temperature is higher than room temperature and they are used as glassy polymers (poly(ethylene terephtalate) PET, poly(ether etherketone) PEEK).
3. polymers having irregular structures do not crystallise whatever the cooling rate (atactic polystyrene, atactic poly(methylmetacrylate))

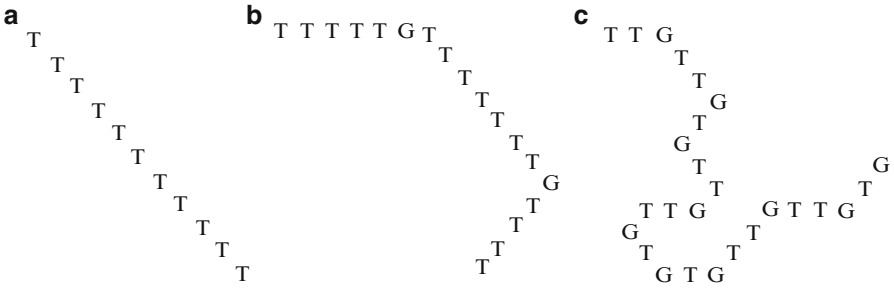


Fig. A1.23 Schematisation of chains: (**a**) all trans, (**b**) with a high persistance length and (**c**) with a low persistance length

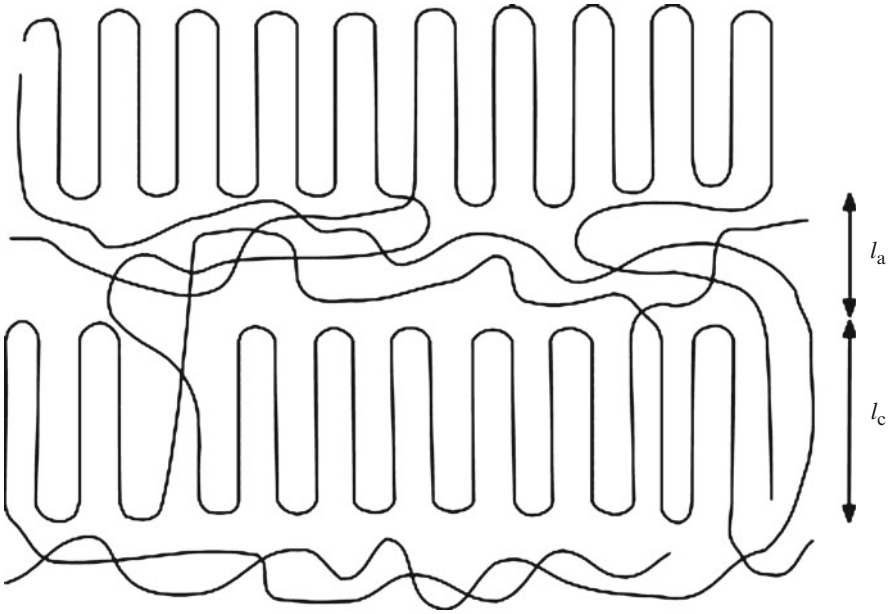


Fig. A1.24 Bidimensional schematic representation of the lamellar structure in a semi-crystalline polymer

Crystallisation proceeds generally by regular chain folding (Fig. A1.24) leading to quasi parallelipipedic lamellae of a few nanometres thickness, separated by amorphous layers of the same order of magnitude.

When the cooling rate is low enough, lamellae tend to extend longitudinally to form long ribbons growing radially from a nucleation centre to give spherulites (Fig. 1.21 in Chap. 1), which can reach the millimetric scale.

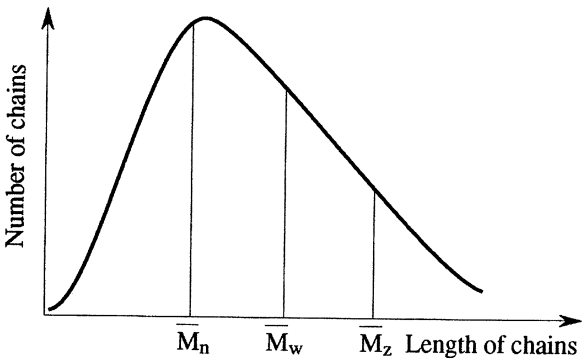


Fig. A1.25 Relative locations of the various molar mass averages of the chain lengths distribution

A1.3.2.4 Structural Parameters

The distribution of molecular weights is an important parameter in the characterisation of the structure of polymers. The average molar mass is characterised by the ratio:

$$\bar{M} = \sum_i n_i {M_i}^{\alpha} \Big/ \sum_i n_i {M_i}^{\alpha-1} \tag{A1.14}$$

where n_i is the number of molecules of weight M_i and α an integer parameter:

α = 1 yields: $\bar{M}_n = \sum_i n_i M_i / \sum_i n_{i\,i}$, the number average molar mass.

α = 2 yields: $\bar{M}_w = \sum_i n_i {M_i}^2 / \sum_i n_i M_i$, the weight average molar mass.

α = 3 yields: $\bar{M}_z = \sum_i n_i {M_i}^3 / \sum_i n_i {M_i}^2$, the Z average molar mass.

For a *homodisperse* polymer $M_n = M_w$. For a polydisperse polymer, $IP = M_w/M_n$ is the *polydispersity index* (Fig. A1.25).

The viscosity average molar mass is defined as:

$$\bar{M}_v = \left(\sum_i n_i {M_i}^{1+\alpha} \Big/ \sum n_i M_i \right)^{1/\alpha} \tag{A1.15}$$

The number average molar mass can be determined by *osmometry*:

$$\frac{\pi}{RTc} = \frac{1}{\bar{M}_n} + Ac + \ldots \tag{A1.16}$$

where π is the osmotic pressure of a solution of polymer, R the perfect gas constant and c the concentration.

The weight average molar mass can be determined by measurements of *scattered laser light*:

$$\frac{\pi K c}{R_0} = \frac{1}{\bar{M}_w} + Bc + ... \tag{A1.17}$$

where R_0 is the intensity of the light scattered in the direction of the axis of the incident beam.

The most common methods to determine molecular weight distributions are variants of high pressure liquid *chromatography* (size exclusion chromatography SEC also called gel permeation chromatography GPC).

The viscosity average molar mass can be determined by *viscosimetry*; the Z average molar mass by *sedimentation* in ultracentrifugation.

The *composition* of polymers can be determined by Fourier transform infrared spectroscopy (FTIR), Raman spectroscopy or nuclear magnetic resonance (NMR); their *crystalline structure* by wide or small angle X-ray scattering or by small angle neutron scattering.

Density, Packing Density

Polymer density depends first on atomic composition. If the monomer unit of molar mass M_m contains N_m atoms, an average atomic mass can be defined as $M_a = M_m / N_m$. M_a ranges from about 4.7 g/mol (hydrocarbon polymers) to about 16.7 g/mol (polytetrafluorethylene) for industrial polymers. The density of amorphous phases at ambient temperature varies approximatively as $\rho_a = k_a M_a^{2/3}$, where $k_a \approx 31{,}000 \pm 1{,}000\ \mathrm{kg^{1/3}\ m^{-3}\ mol^{2/3}}$.

The density is also under the second-order influence of cohesion (it increases with the cohesive energy density), crystallinity ($\rho_c \approx 1.117\ \rho_a$ in average according to Bicerano (2002)) and several other factors. For certain authors as van Krevelen (1993), the key factor is the packing density $\rho^* =$ van der Waals volume / molar volume $= V_W / V = \rho V_w / M$. Packing densities of glassy polymers at ambient temperature can vary from about 0.63 (non polar polymers such as polystyrene) to about 0.72 (highly polar, hydrogen bonded polymers such as poly(vinylalcohol)). Authors have suggested that ρ^* could be structure independent at 0 K or at T_g, but this is contradicted by experimental data.

Free Volume

The free volume concept starts from the idea that in condensed state, a given molecule displays restricted mobility because the surrounding molecules form a "cage" limiting or hindering its motion. It has been decided to define free volume as the volume needed by large amplitude cooperative segmental motions responsible for viscoelasticity in rubbery state. It could be imagined, in a first approach, that the free volume corresponds to the unoccupied volume $(1 - \rho^*) \sim 0.37 \pm 0.05$ or that the free volume is the volume excess of amorphous phase relatively to crystalline one

(∼0.12 in average), but no suitable prediction can be made from these hypotheses. A more fruitful approach starts from the observation that the volumic expansion coefficient is higher in rubbery state ($\alpha_l \sim (5–10)10^{-4}$ K^{-1}) than in glassy state ($\alpha_g \sim (1–4)10^{-4}$ K^{-1}). The free volume would be then the volume excess created by dilatation in rubbery state: $v_f = \alpha v_g(T\text{-}T_g)$ where v_f is the free volume per mass unit, v_g is the specific volume at T_g and $\alpha = \alpha_l - \alpha_g$ is called expansion coefficient of free volume.

This concept has been refined considering that a certain mobility remains in a short temperature interval below T_g and the definitive definition of free volume is:

$$f = \frac{v_f}{v_g} = f_g + (\alpha_l - \alpha_g)(T - T_g) \tag{A1.18}$$

where v_g is the specific volume at the glass transition temperature T_g.

Free volume theory applied to miscible blends. Let us consider a miscible mixture of a polymer (T_{gp}, α_p) with a solvent ($T_{gs} < T_{gp}$, α_s), which can be a true solvent (for instance absorbed water), an additive (for instance a plasticiser), another polymer (provided it is miscible) or even a random comonomer. Let us call respectively v and $(1 - v)$ the volume fractions of the solvent and of the polymer. The simplest version of the free volume theory starts from two hypothesis:

1. the free volume fractions are additive: $f = (1-v)f_p + vf_s$
2. the free volume fraction at T_g is a universal value (classically $f_g = 0.025$)

Combining both hypotheses, the following relationship is obtained:

$$T_g = \frac{(1-v)\,\alpha_p T_{gp} + v\alpha_s T_{gs}}{(1-v)\,\alpha_p + v\alpha_s} \tag{A.19}$$

Combining with the Simha-Boyer rule: $\alpha T_g =$ constant, one obtains:

$$\frac{1}{T_g} = \frac{1}{T_{gp}} + A_s v \tag{A.20}$$

where $A_s = 1/T_{gs} - 1/T_{gp}$.

The effect by which a compound of low T_g induces a decrease of the glass transition temperature of a polymer matrix to which it is mixed is called *plasticisation*. Additives used to decrease T_g are called *plasticisers*.

Figure A1.26 shows the scale of sizes of basic structural elements of polymers.

A1.3.3 Main Polymers

There are two types of polymerisation: *condensation*, in which a chemical reaction takes place (Fig. A1.27), with the elimination of a small molecule such as water or an alcohol; and *addition*, in which nothing is eliminated (Fig. A1.28) (Table A1.9).

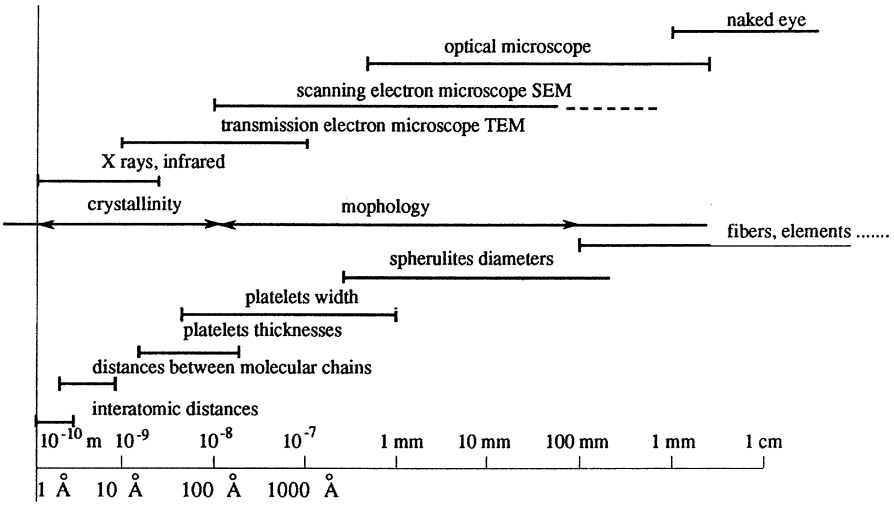


Fig. A1.26 Scale sizes of basic structural elements of polymers, with methods of observation

A1.4 Amorphous Materials

A1.4.1 Glasses

Figure A1.29 is a schematic indication of the difference between a crystalline and an amorphous solid. It is possible for the dispositions of nearest neighbours in the crystalline form to be preserved, as is the case for silica glass.

As there is no periodic structure the phenomena of X-ray diffraction by a crystal are not seen, but as the distribution of interatomic distances is not entirely random there are observable angular variations in the diffracted intensity. Fourier analysis can give the probability of a volume element being occupied by an atom, as a function of the distance from a given atom.

Oxides forming glasses are SiO_2, B_2O_3, GeO_2, P_2O_5.

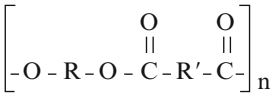
Al_2O_3, BeO_2 are intermediate glasses; they form glasses when combined with others: aluminosilicate, aluminoborate, aluminophosphate.

Mg^{++}, Zn^{++}, Ca^{++}, Sn^{++}, Pb^{++}, Ba^{++}, Li^{+}, Na^{+}, K^{+}, Cs^{+} ions are modifiers.

A1.4.2 Amorphous Metals

Amorphous metals are super-cooled liquids. They were first obtained by very fast cooling; other techniques are available like mechanical alloying, vapour deposition.

polyesters

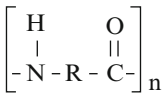


e.g. Dacron

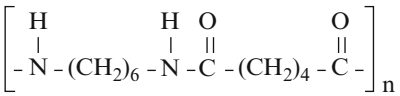
R = CH_2CH_2

R'=

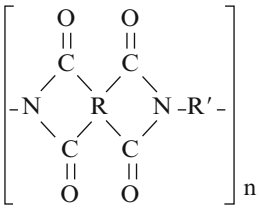
polyamides



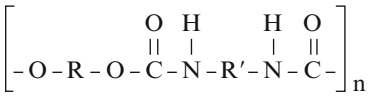
e.g. Nylon 66



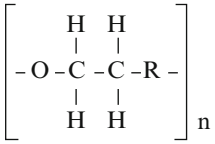
polyimides



polyurethanes



polyethers



silicones

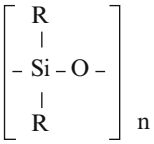


Fig. A1.27 Examples of condensation polymers

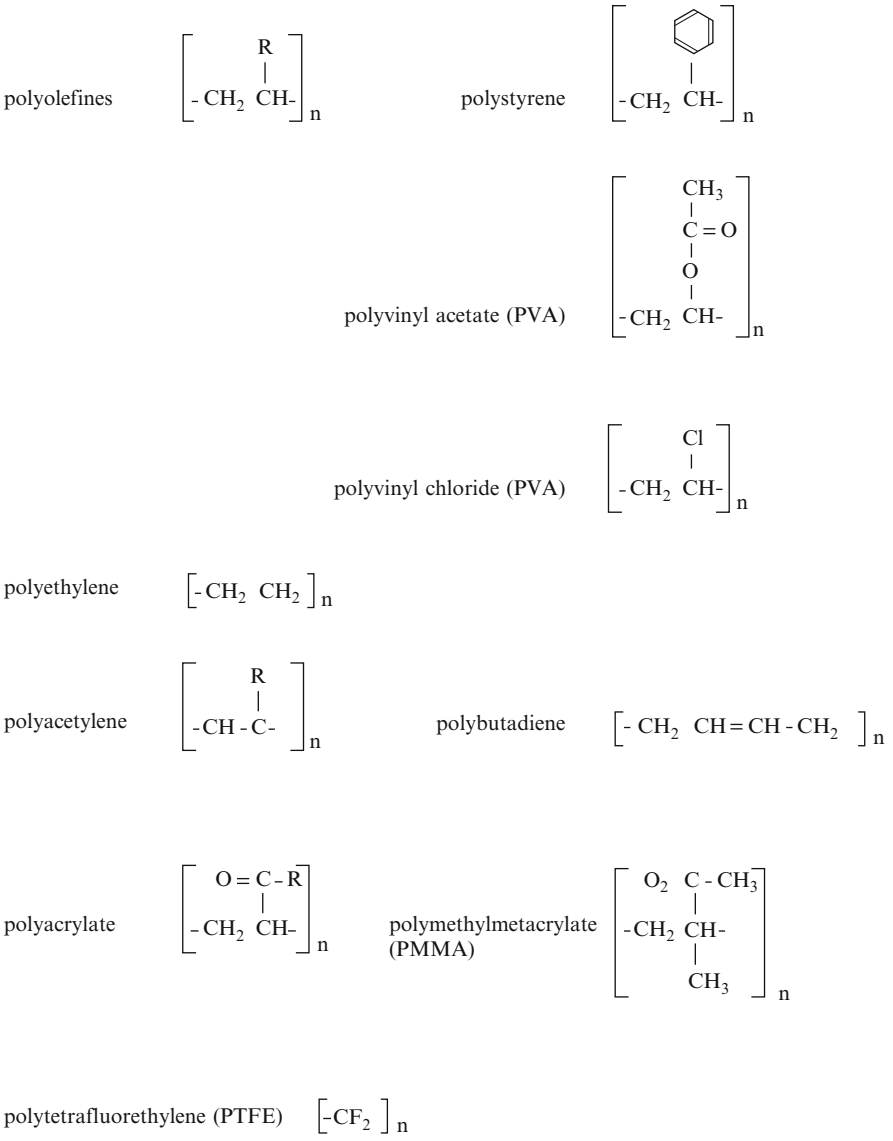


Fig. A1.28 Examples of addition polymers

A1.5 Exercises

1. Calculate the repulsion potential of NaCl, given that the lattice parameter is 3.96 x 10^{-10} m and the binding energy is 777.9 kJ/mol.
2. Which {110} planes contain the direction [111]?
3. Is the direction [123] in the plane (111)? – in the plane $(11\bar{1})$?

Table A1.9 Densities and melting characteristics of some technologically important semi-crystalline polymers

Polymer	Acronym	ρ_a (kg/m^3)	ρ_c (kg/m^3)	T_m (K)	H_f (10^{-3} J/kg)
Polyethylene	PE	850	1,000	413	285
Polypropylene	PP	850	950	440	238
Polytetrafluorethylene	PTFE	2,000	2,350	604	59
Polyoxymethylene	POM	1,250	1,540	460	237
Polyamide 6	PA6	1,080	1,230	496	195
Polyamide 11	PA11	1,010	1,180	463	227
Poly(ethyleneterephthalate)	PET	1,330	1,460	540	120

ρ_a is the density in the amorphous state; ρ_c the density in the crystalline state; T_m the melting temperature; H_f the enthalpy of fusion

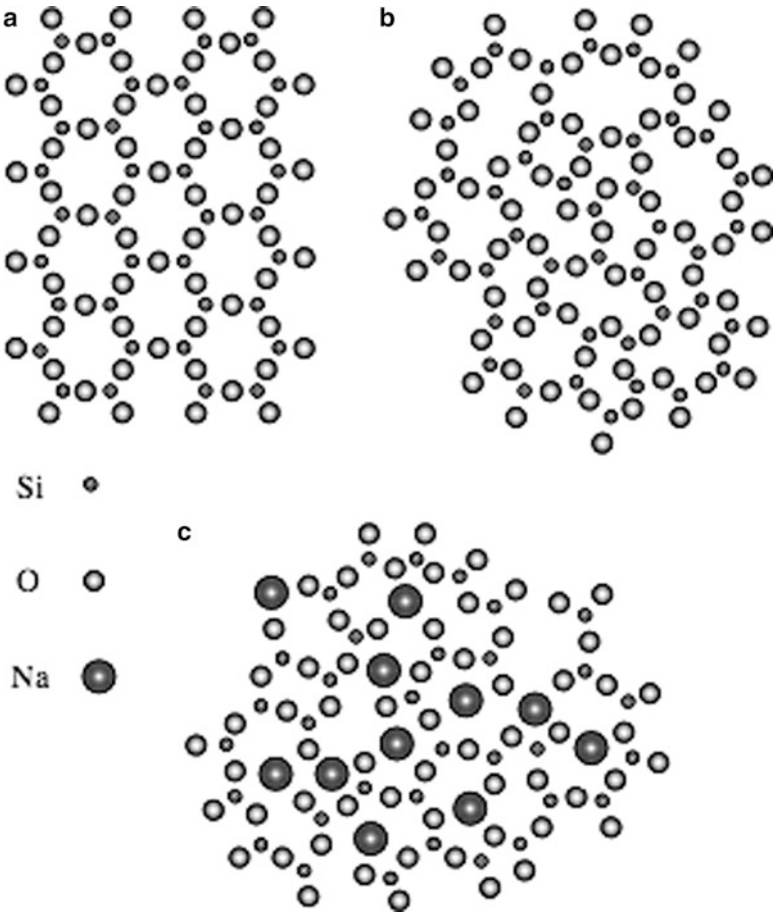


Fig. A1.29 Sketches showing (**a**) crystalline silica, (**b**) amorphous silica and (**c**) silica glass

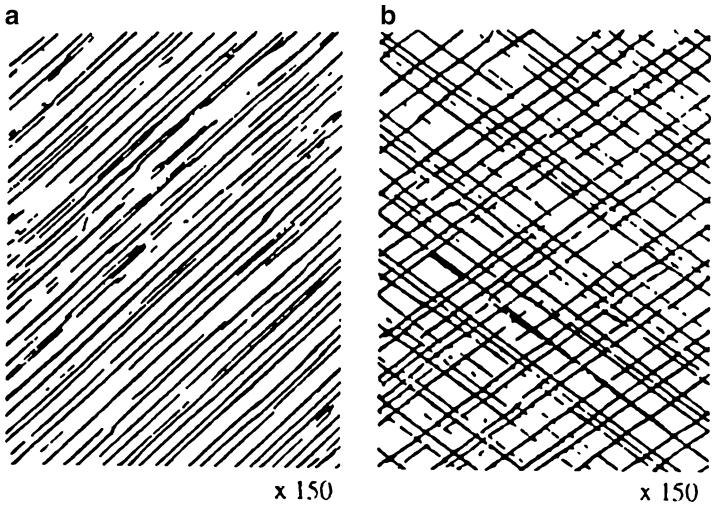


Fig. A1.30 Slip lines in aluminium (**a**) single slip and (**b**) two slip systems

4. Find the densest planes and directions in FCC, BCC and CPH structures.
5. For the three structures of Exercise 4
   - what is the number of atoms in the unit cell?
   - find the ratio of the atomic radius to the unit cell volume
   - find the dimension of the insertion sites
   - what are the reciprocal lattices?
   - how many (a) nearest (b) second nearest neighbours are there?
6. What is the value of c/a for close packing of spheres in the hexagonal system?
7. Find the angle between the directions
   (a) [123] and [110],
   (b) [111] and [122] in cubic systems.
8. Find the condition that must be satisfied by the indices *h, k, l* for planes in the zone whose direction is [*u v w*].
9. Using the stereographic projection, find the possible orientation of the grains on which the slip lines of Fig. A1.30 are seen.
10. Explain the shape of the Laue spots in Fig. A1.31.
11. Explain the Laue transmission photograph of a steel sheet (Fig. A1.32).
12. Interpret the diffraction spectrum of Fig. A1.33
13. Find the change in the diffraction angles for a polycrystal subjected to a tension of $E/10^5$, where E is Young's modulus. Show how X-ray diffraction can be used to measure an applied tension.
14. Prove the parity rules for the indices of reflecting planes for BCC and FCC structures.

Fig. A1.31 Laue transmission photograph of a thin crystal of alpha-iron (Mo radiation)

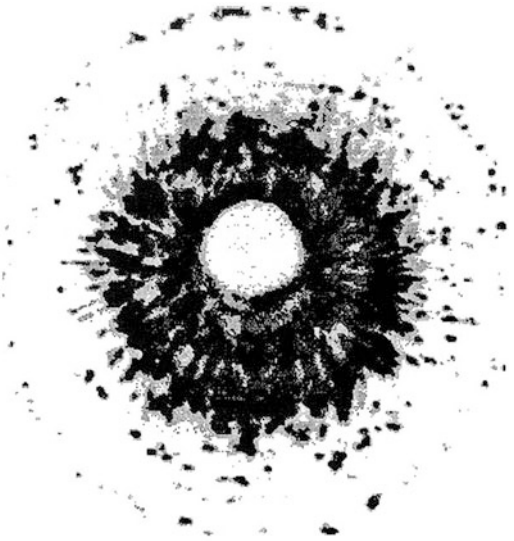


Fig. A1.32 Laue transmission photograph for an annealed steel sheet. The Debye rings are due to the diffraction of Mo Kα

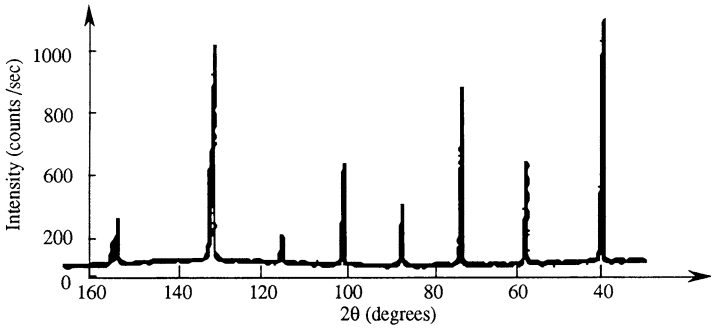


Fig. A1.33 Powder diffraction spectrum (filtered radiation of Cu Kα)

A1.6 References

Barrett C, Massalski TB (1964) Handbook of chemistry and physics. The Chemical Rubber Company, Cleveland

Barrett C, Massalski TB (1988) Structure of metals. Pergamon Press, Oxford

Bicerano T (2002) Prediction of polymer properties. Marcel Dekker, New York

De Gennes PG (1979) Scaling concepts in polymer physics. Cornell U. Press, Ithaca

Fetters LG, Lohse DJ, Graesley WW (1999) Chain dimensions and entanglement spacing in dense macromolecular systems. J Polym Sci Polym Phys 37: 1023–1033

Hermann C (1949) Kristallographie in Räumen beliebirger Dimensionszahl. I. Die Symmetrie-operationem. Acta Crystallogr 2:139–145

Steuer W (1993) Elements of symmetry in periodic lattices, quasi crystal. In: Gerold V (ed) Material science and technology, vol 1. VCH, Weinheim, pp 3–60

Van Krevelen (1993) Properties of polymers. Elsevier, Amsterdam

Appendix B
Annex 2: Phase Transformations

A2.1 Introduction

The term *phase* covers two different concepts:

- in the thermodynamic sense it refers to a defined volume of matter, characterised by particular values of a number of thermodynamic potentials, notably the free enthalpy, or Gibbs free energy, $G(P, T, X_i)$ where P is the pressure, T the temperature and X_i the concentrations of the constituents; the same phase can exist at different temperatures.
- in the crystallographic sense it refers to a distinct crystal structure; there are many examples of phases being given particular names, for example, in the case of steels, ferrite, austenite, martensite and others.

A consideration of different aspects is necessary for the understanding of transitions between phases:

- *thermodynamic*: Gibbs free energies of the phases concerned and the chemical potentials of their constituents
- *crystallographic*: crystal structures, orientations, nature of the interfaces
- *kinetic*: rates of transformations.

There are two main types of phase transformation:

- *homogeneous*: brought about by continuous processes that involve all the relevant atoms simultaneously; for example, spinodal decomposition, order-disorder transformation
- *heterogeneous*: brought about by discontinuous, localised processes such that at any instant only a limited number of atoms are passing from the initial to the final state; for example, diffusion-controlled transformations.

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

Table A2.1 Types of phase transformations

–	–	Short range diffusion	Order-disorder Allotropic Recrystallisation Vapour phase deposition
–	Thermoactivated growth	Long range diffusion	Continuous precipitation and solution Proeutectoid Eutectoid Discontinuous precipitation
Heterogeneous	–	With heat flow	Solidification Fusion
–	Athermal growth	Athermal martensitic	–
–	–	Isothermal martensitic	–
Homogeneous	Spinodal decomposition Order-disorder transformation	–	–

Table A2.1 gives a classification of the phase transformations, to which the following definitions relate:

order-disorder: change from a solid solution state in which the solute atoms are distributed at random (disorder) to one in which they occupy specified sites (order); example Au Cu, $AuCu_3$ (see Annex 1, Sect. A1.2).

thermo-activated growth: strongly influenced by the time for which a given temperature is maintained; this is the case for all diffusion-controlled transformations.

athermal growth: in general, not dependent on time; for example, martensitic transformations, which most often depend only on the temperature. For a comprehensive treatment of phase transformations see Haasen (1991).

A2.2 Equilibrium Diagrams

A2.2.1 The Nature of Equilibrium

Equilibrium between phases is established at the interface under the effect of thermal agitation and extends into the volume by diffusion and by movement of the interface; as a general rule, diffusion is involved. Complete equilibrium is reached

only after a time that is greater, the lower the temperature; thus for example the iron-carbon equilibrium diagram is used in the metastable form Fe-Fe_3C, since only certain slowly-cooled melts can have a microstructure composed of iron and graphite. The same applies to the *martensitic structure* characteristic of quenched steels, which only under *heat-treatment annealing* can evolve, by diffusion and precipitation, towards the stable state of the diagram.

A2.2.2 Thermodynamics of Equilibrium

When two phases are in equilibrium the atoms at the interface can move freely between the two, the bonds that a given atom has in Phase A being replaced by those in Phase B where the structure is different. There is a difference in binding energy between the two phases, measured by the change in *enthalpy* ΔH. The number of bonds in an ordered crystal lattice is greater than in a disordered state (liquid or gaseous, for example), which favors the ordered state. In contrast, thermal agitation favors the disordered state, the disorder being expressed by the *configuration entropy S*. These two opposing effects are brought together in the *free enthalpy* or *Gibbs free energy G*:

$$G = H - TS \tag{A2.1}$$

H (<0) is smaller, the stronger the bonds; S (>0) is greater, the greater the disorder, that is, the greater the number of possible configurations.

Phases A and B are in equilibrium if $\Delta G_{A\to B} = 0$; a system is in equilibrium if its Gibbs free energy, the centre of mass of the free enthalpies of its constituent phases, is minimal (Fig. A2.1).

Enthalpy of a solid solution. Suppose there are initially two phases consisting of atoms A and B with bonds AA, BB respectively (Fig. A2.2), and that a solid solution in which there are AB bonds can form by diffusion. If n_{AB} is the number of these bonds the changes in enthalpy and entropy are

$$\begin{aligned} \Delta H &= n_{AB}\left[H_{AB} - \left(H_{AA} + H_{BB}\right)/2\right] \\ \Delta S &= \mathrm{k}\log\Omega = \mathrm{k}\log\left(C_n^{n_A} + C_n^{n_B}\right) \end{aligned} \tag{A2.2}$$

Here $n = n_A + n_B$ is the total number of sites, n_A, n_B are the numbers of atoms of A and B respectively and C_m^r is the number of combinations of r objects taken from a set of m. Thus if $C_A = n_A/n$ is the atomic concentration of A

$$\Delta S = -n\mathrm{k}\left[C_A \log C_A + (1 - C_A)\log(1 - C_A)\right] \tag{A2.3}$$

The change in entropy will be maximum for $C_A = C_B = 1/2$, when $\Delta S = n\mathrm{k}$ log2, or: $\Delta S = R\log 2$ per mole of solution. This is possible only if the solid solution exists in continuous form between the phases A and B.

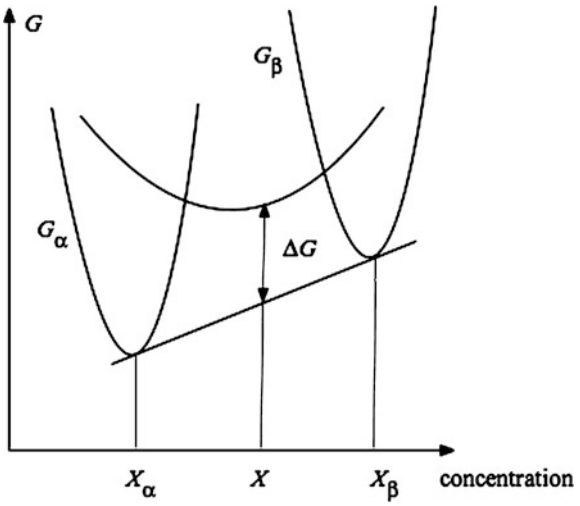


Fig. A2.1 A binary alloy, composition X, is in equilibrium as two phases α and β of compositions X_α , X_β. The diagram shows the change in free energy ΔG

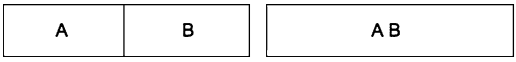


Fig. A2.2 Evolution towards a solid solution AB

The overall balance ΔG will depend on the relative positions of ΔH and ΔS (Fig. A2.3).

If $H_{AB} > ½(H_{AA} + H_{BB})$ there are two situations:

- at the higher temperature T_2, the alloy remains as a single phase, α.
- at the lower temperature T_1, ΔG goes through a local minimum and in this domain the solid solution is a mixture of two solutions rich in A and B, respectively. The more stable state is on the lowest tangent. The corresponding equilibrium shows a *miscibility gap*. The proportions of the phases present are found by the "lever rule": thus:

$$m\,(\alpha_1)\,/m\, = \mathrm{MM_2}\,/\mathrm{M_1\,M_2} \quad (\text{phase } \alpha_1)$$
$$m\,(\alpha_2)\,/m\, = \mathrm{MM_1}\,/\mathrm{M_1\,M_2} \quad (\text{phase } \alpha_2) \qquad \text{(A2.4)}$$

A2.2.3 Multi-phase Equilibria – Equilibrium Phase Diagrams

If metals A, B cannot form a continuous solid solution – for example, because their crystal lattices are too different – we have to compare the stabilities of two solid solutions α, β and possibly a third, liquid, phase (Fig. A2.4).

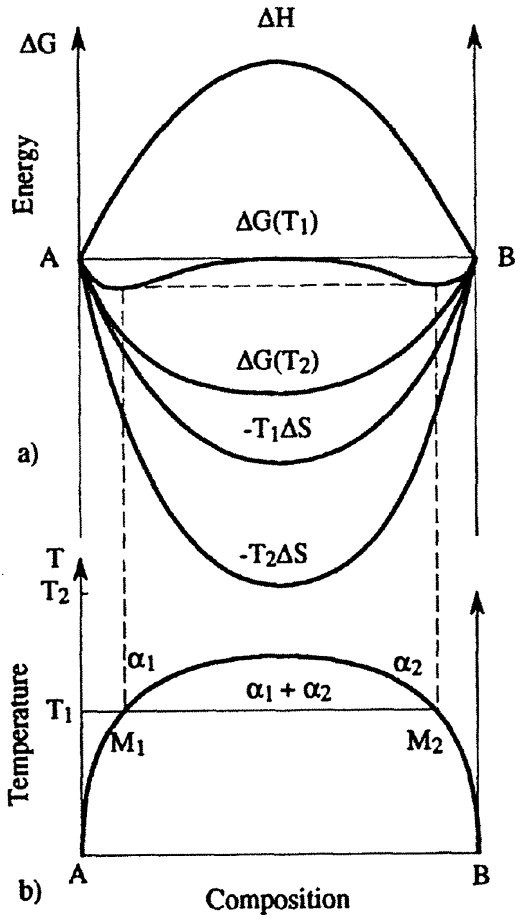


Fig. A2.3 (**a**) Free energies in a solid solution at temperature T_1, T_2; (**b**) equilibrium phase diagram in the solid domain

NOTES

1. Whilst in general equilibrium diagrams are established experimentally, in certain cases thermodynamic models of solutions will enable them to be constructed theoretically.
2. There are precise rules – the Hume-Rothery rules (Hume-Rothery 1955) that enable the conditions under which certain phases can exist to be predicted; the two most important are:

   (a) *solid solutions by substitution*. The solubility of B in A can be high only if the sizes of the A and B atoms do not differ by more than 15%.

   Example: silver ($r_{Ag} = 0.159$ nm) is only weakly soluble in copper but copper has unlimited solubility in nickel ($r_{Cu} = 0.141$ nm, $r_{Ni} = 0.138$ nm).

   (b) *formation of intermediate phases*. Certain crystal structures are always stable for particular electronic concentrations (number of electrons per atom)

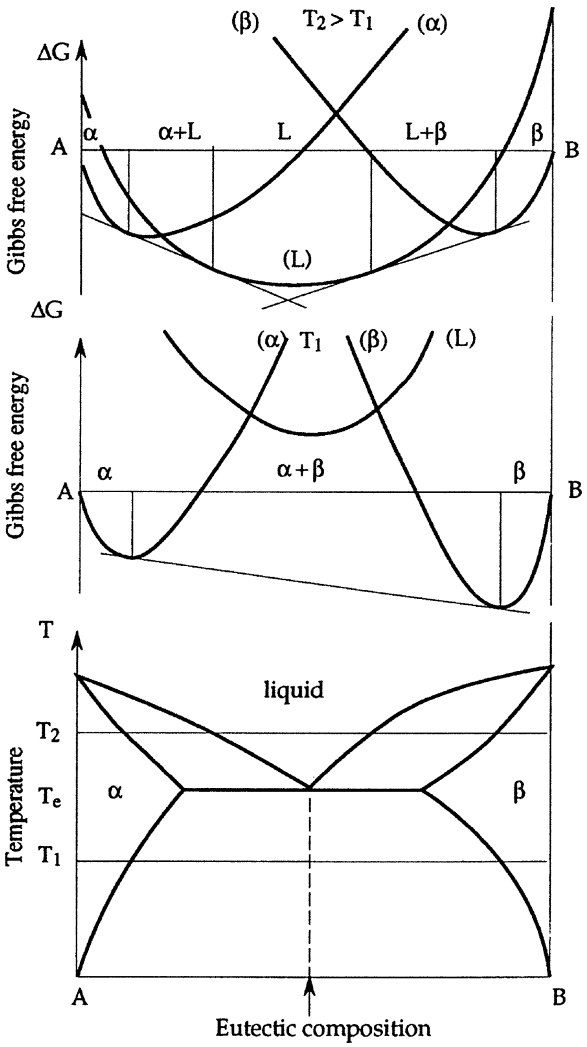


Fig. A2.4 Equilibrium between two solid phases α and β and a liquid L. The three phases equilibrium (common tangent to the 3 phases) can occur only at one precise temperature T_g corresponding to the formation of a eutectic (pressure is kept constant)

Table A2.2 Example of stability of phases in the Cu Zn system

Material	CuZn	Cu_5Zn_8	$CuZn_3$
Electrons per atoms	3/2	21/13	7/4
Structure	BCC	Brass (g)	HCP

Example: Cu, valence 1 and Zn, valence 2; the following table gives the corresponding data (Table A2.2).

Equilibrium diagrams are composition-temperature diagrams from which we can find the *composition* and the relative proportions of the different phases present at equilibrium at any given temperature. They do *not* give any structural information, or anything concerning the possible existence of metastable phases or components.

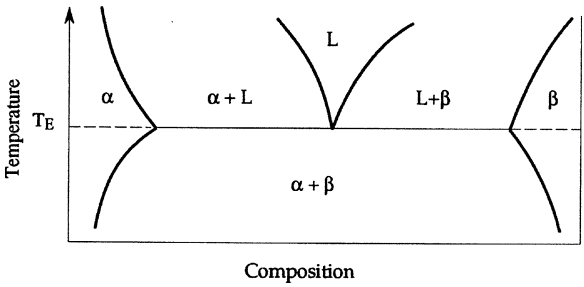


Fig. A2.5 Equilibrium diagram for a binary alloy

The phase rule. The *variance v* of a system (the number of independent parameters in the equilibrium) is:

$$v = N + 2 - \varphi \tag{A2.5}$$

where N is the number of independent constituents and φ the number of phases. Since most usually the pressure is fixed at atmospheric this will be reduced to:

$$v = N + 1 - \varphi \tag{A2.6}$$

Thus in the binary system ($N = 2$) of Fig. A2.5 there are 1-phase regions (solid solution, liquid) in which T and C_B can be chosen independently, 2-phase regions in which, when the temperature is chosen, the concentrations of the two phases present are fixed, and the 3-phase points at temperature T_E at which the compositions of all three phases are fixed.

Ternary or quaternary diagrams are much more complex and more difficult to use because of the difficulty of representation.

Examples of equilibrium diagrams

In the following pages we give a few examples of diagrams for binary alloys, as a help to understanding the heat treatments applied to a number of industrial alloys. Because of the importance of equilibrium phase diagrams the reader should attempt the exercises at the end of this Annex.

A2.2.3.1 Ferrous Alloys

Diagram 1: Iron-Carbon (Fig. A2.6)
Diagram 2: Iron – Nickel (Fig. A2.7)
Diagram 3: Iron–Chromium (Fig. A2.8)

A2.2.3.2 Light Alloys

Diagram 4: Aluminium–Silicon (Fig. A2.9)
Diagram 5: Aluminium–Copper (Fig. A2.10)

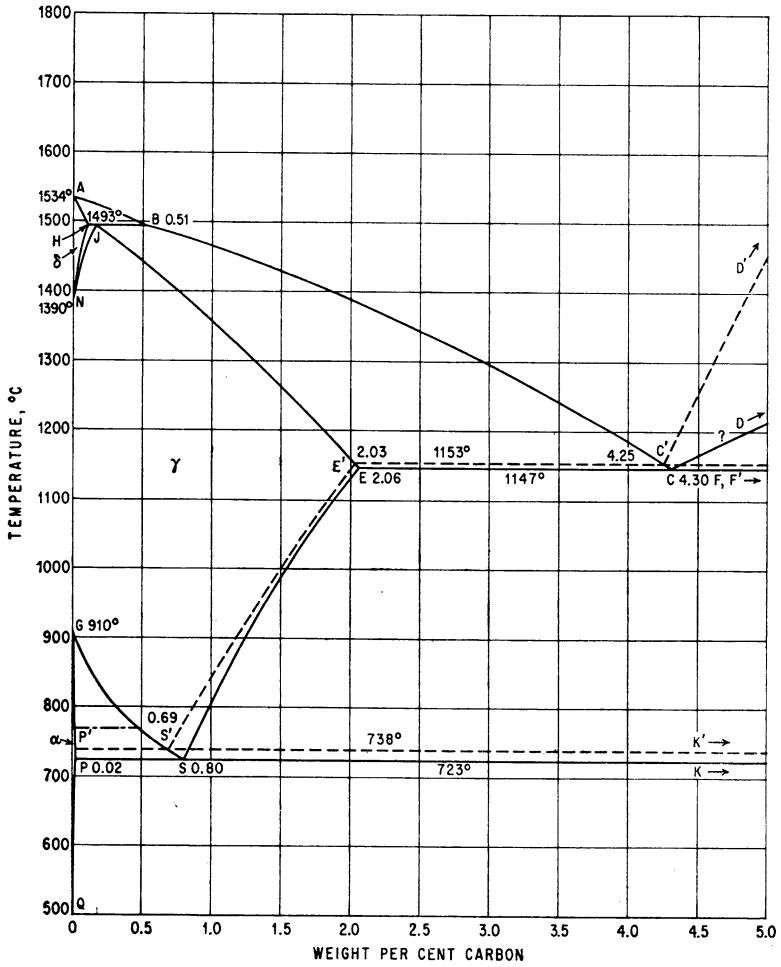


Fig. A2.6 Iron−Carbon diagram. γ: Austenite FCC; δ: Ferrite, high temperature BCC, α: Ferrite, low temperature BCC; Fe_3C: Cementite orthorhombic. Full lines: metastatable diagram for Fe – Fe_3C, corresponding to transformations in steels that have been cooled rapidly. *Dotted lines*: equilibrium diagram for Fe – C (graphite), in practice, this corresponds only to melts that have been cooled slowly (Hansen, Constitution of binary alloys, McGraw-Hill 1958)

NOTES

– *Diagram 1*. Fe_3C is not shown in this diagram; it would appear as a vertical line at 6.67% C (by mass). Notice the difference in solubility of carbon between the α and γ phases: this property forms the basis of practical heat-treatments of steel.
– *Diagram 2*. Notice that adding nickel enables the γ-domain to be considerably extended: Ni is said to be a *gammagenic* element. However, significant departures from this diagram are observed in practice; especially for the low-temperature

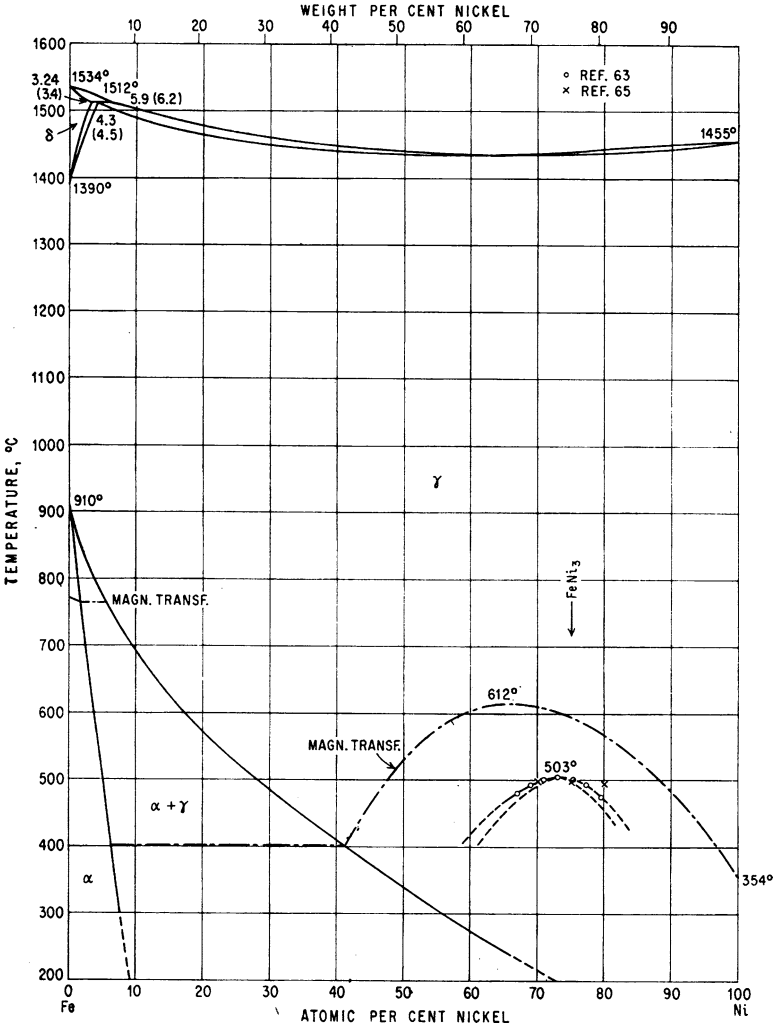


Fig. A2.7 Iron—Nickel diagram (Hansen, Constitution of binary alloys, McGraw-Hill 1958)

transformations, which are important for heat-treatments. The $\gamma \rightarrow \alpha'$ transformation at sufficiently high nickel content (over 20%) is martensitic in nature.

– *Diagram 3*. This shows the very *alphagenic* character of Cr; also the presence of an ordered σ phase. This latter appears, under certain conditions, in stainless steels and certain superalloys; it can cause serious thermal embrittlement.
– *Diagram 4*. The eutectic corresponds to an alloy Al – 13Si: this is Alpax, much used for castings, for example for automobile engine cylinder heads. It is the aluminium analogue of the Fe – 4.3 C steel used for engine cylinder blocks.

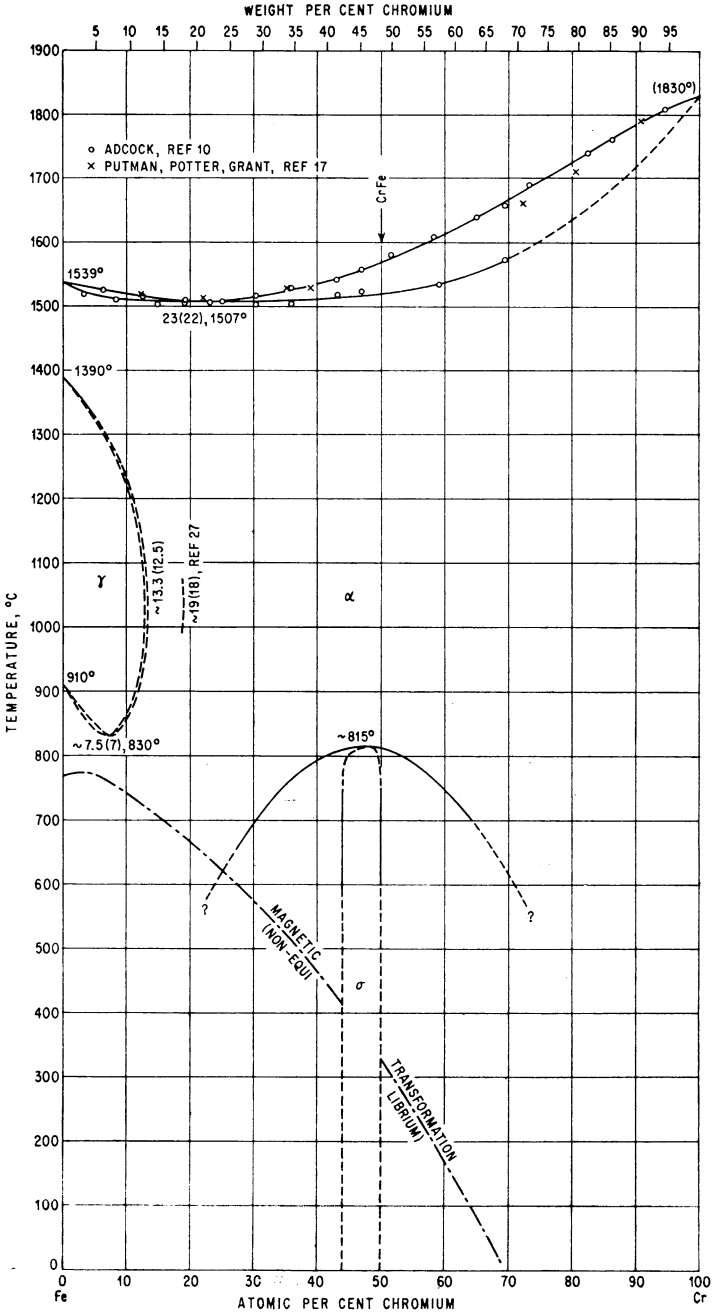


Fig. A2.8 Iron−Chromium diagram (Hansen, Constitution of binary alloys, McGraw-Hill 1958)

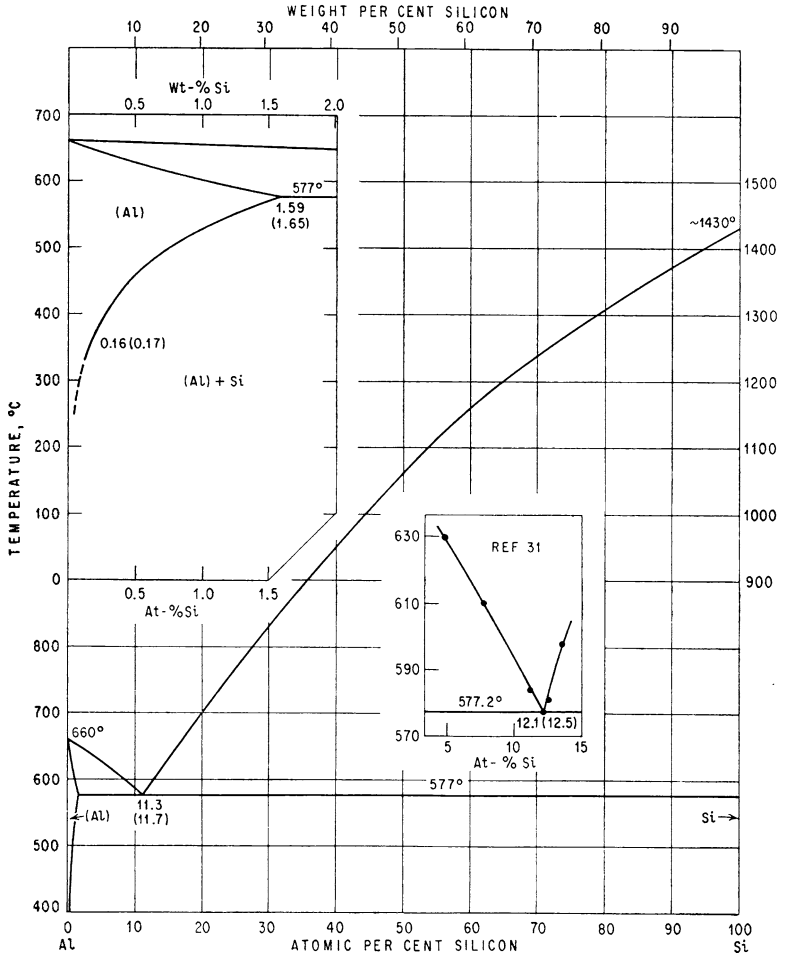


Fig. A2.9 Aluminium−Silicon diagram (Hansen, Constitution of binary alloys, McGraw-Hill 1958)

– A study of the phase diagram explains the importance of Alpax. In the Al − Si system the fusion point reaches its lowest value at the composition corresponding to the eutectic, and solidification occurs at a precisely-defined temperature rather than over a significant interval. Further, the coupled solidification of two phases, silicon platelets in the aluminium, causes the liquid metal to flow easily and hence makes possible the production of complex shapes.
– *Diagram 5.* At high aluminium contents the solubility of copper begins to fall rapidly at 548°C: this is the basis of heat-treatments of light alloys containing copper. In alloys sufficiently rich in copper the phase Al_2Cu (θ) is formed.

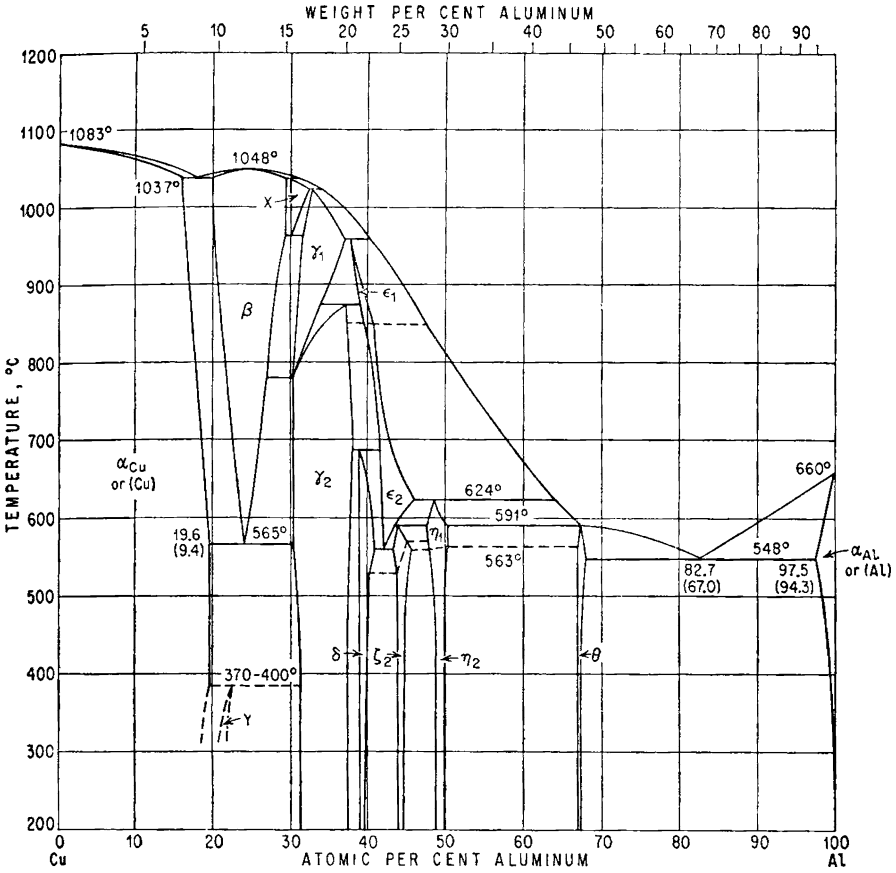


Fig. A2.10 Aluminium—Copper diagram (Hansen, Constitution of binary alloys, McGraw-Hill 1958)

A2.3 Kinetics – Diffusion

Diffusion is the phenomenon of transport of atoms from one site to another, the displacements of the individual atoms being related to the thermal agitation. It is a fundamental phenomenon, controlling the evolution of the chemical composition of the phases present and the growth of new phases by precipitation or solidification. Diffusion enables the equilibrium predicted by the diagram to be reached.

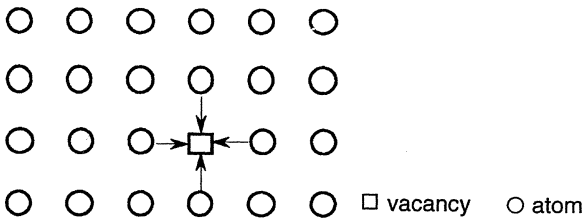


Fig. A2.11 Mechanism of self-diffusion by vacancies

Table A2.3 Orders of magnitude of vacancies concentrations in copper at various temperatures

$\Delta H_f = 1$ eV and $\Delta S_f = k$ (approx.)			
T/T_f	0.22	0.591	0.961
C	10^{-17}	6.10^{-7}	$1.3.10^{-4}$

A2.3.1 Basic Diffusion Mechanisms

A2.3.1.1 Self-diffusion (Fig. A2.11)

In a crystal lattice an atom cannot change its position unless there is a vacant site in its immediate neighbourhood, that is, a vacancy. Diffusion thus depends on the number of vacancies in the lattice, and this increases with increasing temperature; at absolute temperature T the equilibrium concentration of vacancies C is:

$$C = \exp\left(\frac{T\Delta S_f - \Delta H_f}{kT}\right) \tag{A2.7}$$

where ΔS_f, ΔH_f are the entropy and enthalpy of formation respectively (Table A2.3).

In self-diffusion the mobility of the vacancies, and therefore of the atoms, is governed by the Maxwell-Boltzmann distribution giving the frequency Γ of atom vacancy interchanges:

$$\Gamma = \bar{\nu} \exp\left(-\frac{\Delta G_f + \Delta G_m}{kT}\right) \tag{A2.8}$$

where $\bar{\nu}$ is the mean frequency of vibration of the atoms

ΔG_f is the free energy of formation of the vacancies
ΔG_m is the free energy of migration of the vacancies.

A2.3.1.2 Hetero-Diffusion

Solid solutions by substitution (Fig. A2.12).

This again depends on the movement of vacancies. The rate differs from that of self-diffusion, and is slower, the greater the diameter of the foreign atom than that of the atoms of the solid solution.

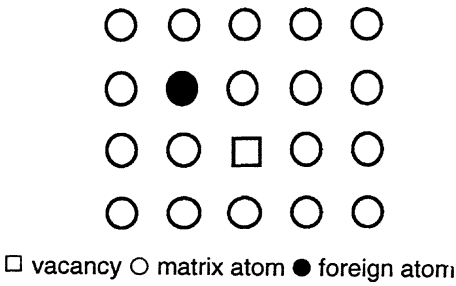


Fig. A2.12 Basic mechanism for diffusion by substitution in a solid solution

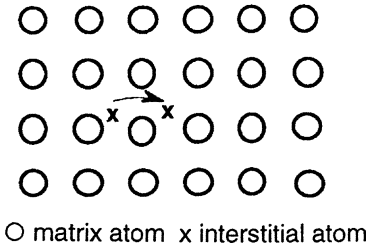


Fig. A2.13 Mechanism for diffusion by insertion in a solid solution

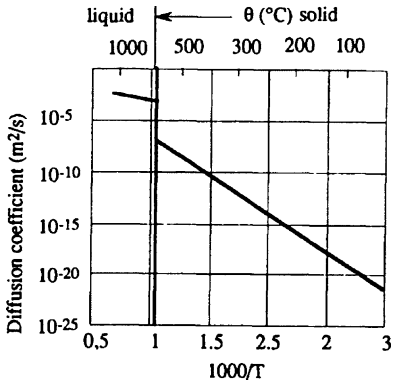


Fig. A2.14 Diffusion coefficient for Cu in Al as a function of temperature, in the solid and liquid state

Solid solution by insertion (Fig. A2.13)

In this case vacancies are not necessary for the movement of the foreign atoms. The speed of diffusion is, in general, much greater than in the two previous cases and the activation energy is less than that required for diffusion by vacancies.

NOTES

1. Since diffusion rate is a function of bond energy and thermal agitation it is greater, the lower the fusion temperature T_f and the latent heat of fusion.
2. Diffusion is slowest in compact lattices, *e.g.* FCC and CPH; and is faster in liquids than in solids (Fig. A2.14).
3. When diffusion into the volume is slow – the effect of large size, low temperature, etc. – it occurs preferentially along defects in the crystal lattice such as grain boundaries and dislocations, which form diffusion *short circuits*. The activation energy is then lower, as Fig. A2.15 shows.

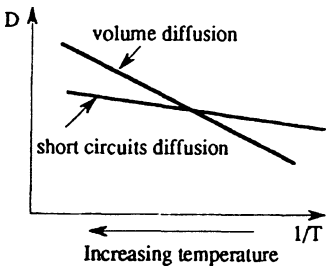


Fig. A2.15 Diagram showing how, at high temperature, volume diffusion overtakes short-circuit diffusion

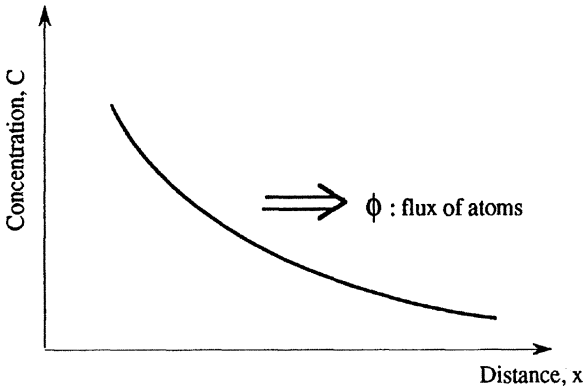


Fig. A2.16 Variation of concentration with distance and associated atomic flux

A2.3.2 The Diffusion Laws

A2.3.2.1 Fick's First Law

The atomic flux ϕ is given by:

$$\phi = -D\frac{\partial C}{\partial x} \tag{A2.9a}$$

in one dimension (Fig. A2.16) and by

$$\phi = -D\,\mathrm{grad}C \tag{A2.9b}$$

in three dimensions.

The variation of the diffusion coefficient D with temperature is given by Arrhenius's law:

$$D = D_0 \exp\left(-Q_\mathrm{d}/\mathrm{k}T\right) \tag{A2.10}$$

Table A2.4 Diffusion coefficients and activation energies

	Solvent	Solute	D_0 10^{-4} m^2/s	Q_d kJ/mol
Self-difffusion	Fe α	Fe	1.9	240
	Fe γ	Fe	0.8	269
	Cu	Cu	0.2	197
	Ag	Ag	0.4	184
	Si	Si	0.032	410
Diffusion of atoms in solution by substitution	Cu	Al	0.045	165
	Al	Cu	2.3	140
	Cu	Zn	0.34	191
	Fe γ	Ni	0.77	280
	Si	P	0.39	201
Diffusion of atoms in solution by insertion	Fe α	C	0.02	84
	Fe γ	C	0.2	134
	Fe α	N	0.003	76
	Fe γ	N	0.001	13.4

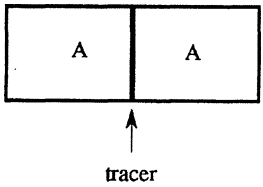


Fig. A2.17 Thin layer sandwiched between two samples of material A

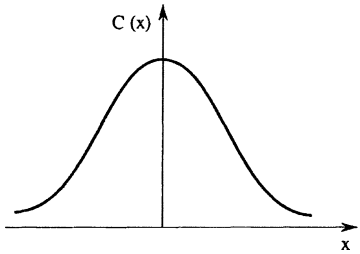


Fig. A2.18 Variation of concentration with distance

where D_0 is a constant, 0.2–2.10^{-4} $m^2\,s^{-1}$ approximately, and Q_d is the diffusion activation energy. For diffusion by vacancies $Q_d = \Delta G_f + \Delta G_m$, in general 100−200 kJ/mol. More precise values are given in Table A2.4.

A2.3.2.2 Fick's Second Law

$$\partial C/\partial T = D\left(\partial^2 C/\partial x^2\right) \tag{A2.11a}$$

in one dimension;

$$\partial C/\partial t = D\nabla^2 C \tag{A2.11b}$$

in three dimensions.

We deduce from the solution that the distance L diffused in time t is given by:

$$L = \sqrt{Dt} \tag{A2.12}$$

Example: Thin layer in an infinite sample (often of a radioactive tracer) (Fig. A2.17)

If Q is the quantity of matter per unit area in the layer, the solution of Fick's equation with the appropriate boundary conditions is:

$$C\,(x,t) = \left(Q\sqrt{2\pi Dt}\right)\exp\left(-x^2/Dt\right) \tag{A2.13}$$

At any given time t the curve of C against distance x is a gaussian centred on the interface (Fig. A2.18). This method is widely used in experimental studies of the variation of D with temperature.

NOTES

1. In a complex system such as a solid solution with a number of constituents the diffusion coefficient D is no longer simple and a number of diffusion coefficients may have to be defined.
2. Diffusion enters into many applications, for example:
   - homogenisation of alloys (Exercise A2.11.6)
   - heat treatments
   - cementation of steels (Exercise A2.11.5)
   - welding and brazing
   - oxidation of metals
   - doping of semi-conductors
   - chemical modification of glasses
   - sintering

A2.4 Nucleation

Theories of nucleation are very useful for the understanding of phase changes. They apply quite well to liquid systems, solidification being a good example (*cf.* Sect. A2.7). Application to the solid state may need more care, at least if quantitative results are required; even so, they can give important information in such cases.

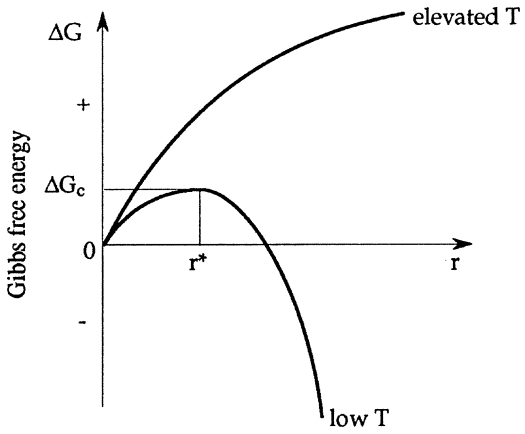


Fig. A2.19 Relation between radius r of an embryo and change in free energy ΔG

A2.4.1 Free Energy Associated with Variations in the Configuration

We consider a system that is undergoing a change of state, for example from liquid to solid or an allotropic change in the solid state. Suppose an *embryo* consisting of the new phase (β) is formed from the initial phase (α); taking the embryos to be spheres of radius r the change in free energy is:

$$\Delta G = \frac{4}{3}\pi r^3 \Delta G_v + 4\pi r^2 \gamma \tag{A2.14}$$

where $\Delta G_v = G_\beta - G_\alpha$ is the difference in free energy between the two phases; ($\Delta G_v < 0$) and γ is the interface energy.

It is easy to show that ΔG will have a maximum if the supersaturation, measured by ΔG_v, is great enough, and therefore if the temperature is low enough; this is shown in Fig. A2.19. When the embryos have reached a critical size r^* we say we are dealing with *nucleii*, and then:

$$r^* = -2\gamma / \Delta G_v$$

$$\Delta G_c = \frac{16\pi}{3} \frac{\gamma^3}{\Delta G_v{}^2} \tag{A2.15}$$

The size of the nucleii increases with decreasing temperature, but they are difficult to see, as Fig. A2.20 shows for the solidification of copper: at 30°C undercooling the critical radius is 0.01 μm, at 0.3°C it is 1 μm.

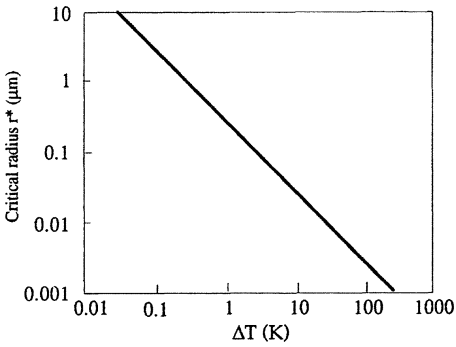


Fig. A2.20 Variation of critical nucleation radius r^* with under-cooling, for copper ($\Delta T = T_m - T$)

The height of the potential barrier ΔG_v that the embryos have to overcome is, to a first approximation, proportional to $(\Delta T)^2$; this determines the rate of formation of the nucleii, which, from the Becker-Doring theory (Russel 1980), is given by:

$$I = I_0 \exp(-\Delta G_v / kT) \tag{A2.16}$$

where I_0 is constant to first approximation. This shows that for a high density of nucleii $|T - T_E|$ must be large, where T_E is the equilibrium temperature.

Solid phases give rise to two complications. The first, purely mechanical in nature, is due to the precipitates not occupying exactly the same volume as the matrix, resulting in a distortion of the lattice. This requires a second, *volumetric*, term to be included in the energy balance, the effect of which is to require still higher undercooling and therefore still greater departures from equilibrium. The second difficulty is related to the possibility of changes in composition, when it becomes more difficult to treat the problem rigorously; it does however enable the existence of the metastable phases that are often encountered to be explained, for example the θ' and θ'' phases in the Al–Cu system (Fig. A2.21).

A2.4.2 Heterogeneous Nucleation

The height ΔG_v of the energy barrier to be crossed can be reduced by reducing the surface energy γ. This can be done, under certain conditions, by *heterogeneous nucleation* on preferential sites such as, in the solid phase, grain boundaries or dislocations: Fig. A2.22 relates to the grain boundary case; it gives the resulting ratio of $\Delta G_c^{\,j}$ (heterogeneous) to $\Delta G_c^{\,h}$ (homogeneous) as a function of the ratio: $\gamma_{\alpha\alpha}/2\gamma_{\alpha\beta}$. As this ratio tends to 1, θ tends to 0 and the embryo "wets" the boundary perfectly.

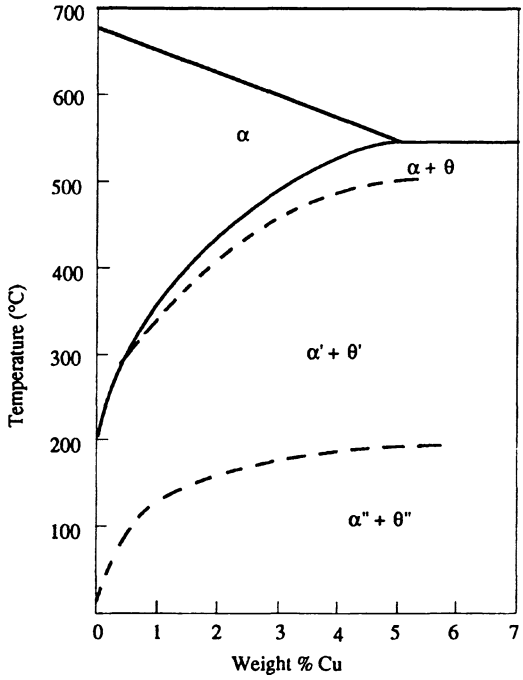


Fig. A2.21 Solubility curves for the stable and metastable states of Al—Cu alloys

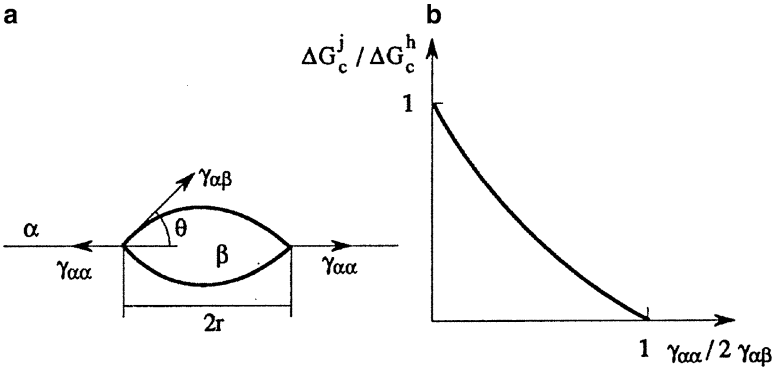


Fig. A2.22 Heterogeneous nucleation on a grain boundary; this is easier, the better the embryo "wets" the boundary

A2.5 Thermally-Activated Growth

We are concerned here with the growth of a phase β in a phase α. Figure A2.23 shows the two extremes for the concentration profiles that we have to consider.

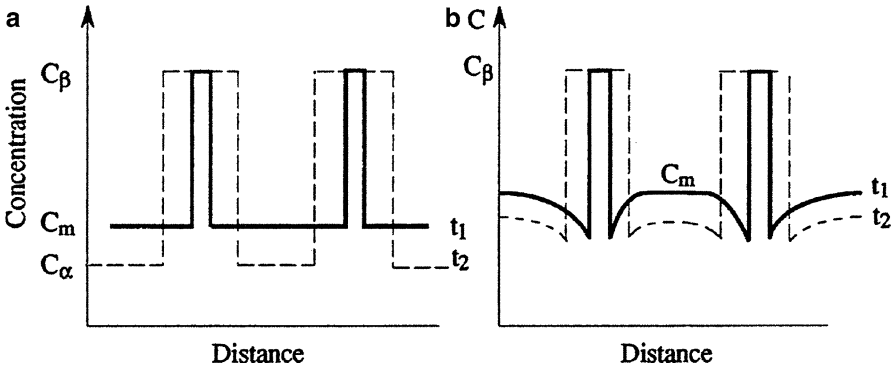


Fig. A2.23 (**a**) reaction at the interface and (**b**) growth by diffusion. The *full curves* are the concentration profiles at t_1, the *dotted curves* at $t_2 > t_1$

A2.5.1 Growth Governed by a Reaction at the Interface

The governing feature is the kinetics of atoms sticking to the interface α/β; if R is the size of the β zone at a given temperature, the growth rate dR/dt is constant. This has been adequately verified for allotropic transformations, but difficulties arise if there are impurities at the interfaces.

A2.5.2 Growth Governed by Diffusion: Zener's Theory

From Fick's equations and the conservation of mass at the interface, with some simplifying assumptions, it can be shown (Zener 1952) that the size of the growing particles is given by: $R = v\sqrt{Dt}$, where v becomes constant after a long enough time. It follows that the rate of growth is inversely proportional to $\sqrt{t}$. However, there is relatively little experimental evidence to support this except for the case of steels; Fig. A2.24 gives results for the growth of ferrite.

A2.5.3 Coalescence

After the formation of the precipitated new phase β the system can reach a quasi-equilibrium, with the concentration C_m of the solute in the matrix having reached the value C_E corresponding to the equilibrium diagram (Fig. A2.25); but the system is not yet in true equilibrium. It tends to reduce the total surface separating the β particles from the α matrix, and the particles continue to grow, the larger at the expense of the smaller: this is called *coalescence* or *ripening*. This phenomenon is

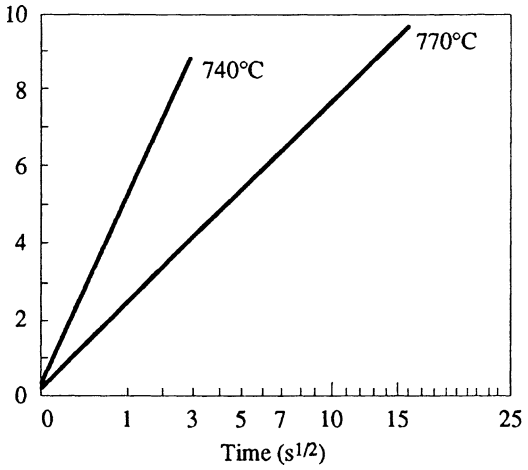


Fig. A2.24 Thickness of ferrite transformed zones as a function of (time)$^{1/2}$; 0.11% carbon steel at 740°C and 770°C

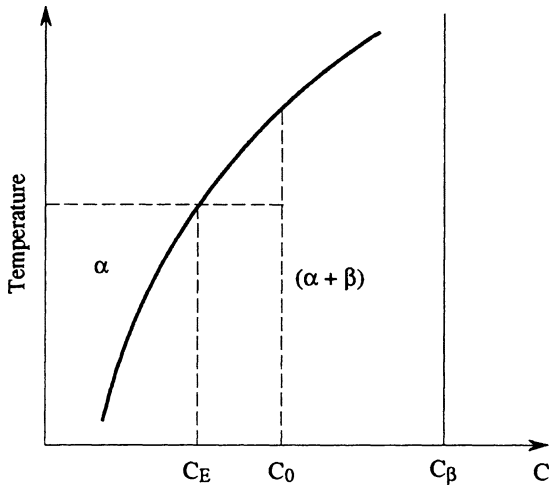


Fig. A2.25 Decomposition of an alloy of initial concentration C_0 of phase β, precipitated in the α matrix

particularly important in the study of microstructural stability of metallic materials, for example of the hardening phase γ' (Ni_3TiAl) of nickel-based superalloys.

A2.5.3.1 The Gibbs-Thomson Equation

For a given particle size r the equilibrium concentration $C(r)$ at the interface does not have exactly the value C_E^{∞} corresponding to an interface of infinite radius of curvature. Theory gives

$$C(r) = C_E^{\infty}\left(1 + \frac{2\gamma\Omega}{\mathrm{k}Tr}\right) \tag{A2.17}$$

where Ω is the atomic volume.

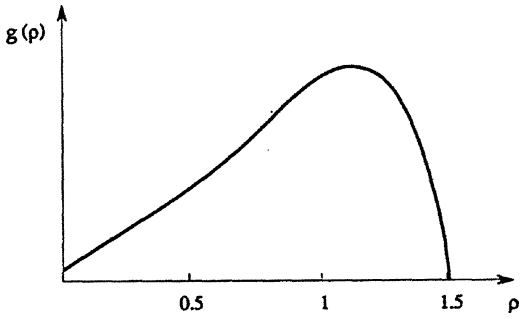


Fig. A2.26 Distribution of particle size in a system in process of ripening

Thus there is a difference in concentration between two particles of different sizes and therefore a flow of material, such that the smaller one becomes dissolved, to the advantage of the larger. Thus a system with a large number of particles ends in an *asymptotic state*, described by the theory of Lifschitz, Slyozov and Wagner (LSW), see Lifschitz and Slyozov (1961), Wagner (1961) or Wagner and Kampmann (1991).

A2.5.3.2 Kinetics of Ripening (LSW Theory)

1. The variation of mean particle size with time is given by:

$$\bar{r}^3 - \bar{r}_0^3 = \frac{64\gamma D\, C_{\mathrm{E}}^{\infty}\, \Omega^2}{9\mathrm{k}T}\,(t - t_0) \tag{A2.18}$$

   The time t_0 is that at which the system begins to coalesce, when the mean particle size is $\bar{r}_0$. In strongly supersaturated systems $C_0 > C_E^{\infty}$ it is often found that t_0 and r_0 are both zero, and it is then possible to predict the rate of coalescence of the particles. The relation has been well verified for many systems, for example for precipitates of γ' in nickel-based alloys – see Exercise A2.11.8.
2. Particle size distribution

   Writing $\rho = r\,/\bar{r}$, the distribution function is:

$$g\,(\rho) = \rho^2(3+\rho)^{-7/3}\,(3\,/2-\rho)^{-11/3}\,\exp\left(-\frac{\rho}{3\,/2-\rho}\right)$$

$$g\,(\rho) = 0 \quad \rho \geqslant 3\,/2 \tag{A2.19}$$

   Observations of real systems show that a much wider spread of particle size ($\rho > 3/2$) often occurs than the theory would predict (Fig. A2.26).

NOTE A system can continue to evolve if there are applied stresses (see Exercise A2.11.9). Differences in the crystallographic parameters and elastic constants for the two phases lead to the presence of a term expressing the interaction between the

mechanical loading and the system. This interaction energy can be calculated from Eshelby's theory of inclusions and heterogeneities (see Chap. 2 in this volume), from which zones of *stability of shape* can be defined for the particles.

A2.6 Phenomenological Theories of Kinetics and Phase Changes

So far we have described the different stages through which a transformation goes – nucleation, growth, ripening. We must now attempt to predict the overall kinetics of the transformation, that is, the way the fraction transformed develops with time. In general it is difficult to derive this from the basic mechanisms of nucleation and the growth laws; consequently we often have to turn to the phenomenological laws that we now describe. We treat these in two groups, for isothermal and non-isothermal transformations respectively.

A2.6.1 Isothermal Transformations

Let ζ be the rate of transformation at any instant. If the growth is controlled by reaction at the interface, the associate growth law is linear and the overall kinetics is described by Avrami's law:

$$\zeta = 1 - \exp\left(-kt^{\alpha}\right) \tag{A2.20}$$

where k is constant at a given temperature and $3 \leq \alpha \leq 4$ (see Table A2.5 below).

If the process is controlled by diffusion, the associated growth is parabolic and in the overall kinetics equation (A2.20) the exponent α is equal to 5/3 or 3/2 according to the assumptions made concerning the saturation of the nucleation sites**.** The following lists give the values corresponding to different mechanisms, from which it is seen that the value of the exponent alone is not sufficient to identify the mechanism (Christian[10] 1965).

A2.6.2 Non-isothermal Transformations

Transformations often take place during cooling, for example with steels and light alloys. To represent these we use CCT – Continuous Cooling Transformation – curves, of which Fig. A2.27 is an example, for a steel, trajectory 1 corresponding to

[10]Christian JW (1965) The theory of transformations in metals and alloys. Pergamon Press, Amsterdam.

Table A2.5 Values of the exponent in the Avrami's law for various mechanisms

Interface reaction	
Nucleation rate constant	4
instantaneous	3
increasing	>4
decreasing	3–4
Nucleation at triple junctions after saturation	2
at boundaries after saturation	1
Diffusion	
Nucleation rate constant	5/2
instantaneous	3/2
increasing	>5/2
decreasing	3/2–5/2
Growth of particles of initially appreciable size	1–3/2
Rods and platelets small relative to separation	1
Thickening of long rods	1
Thickening of very large platelets	1
Segregation at dislocations	2/3

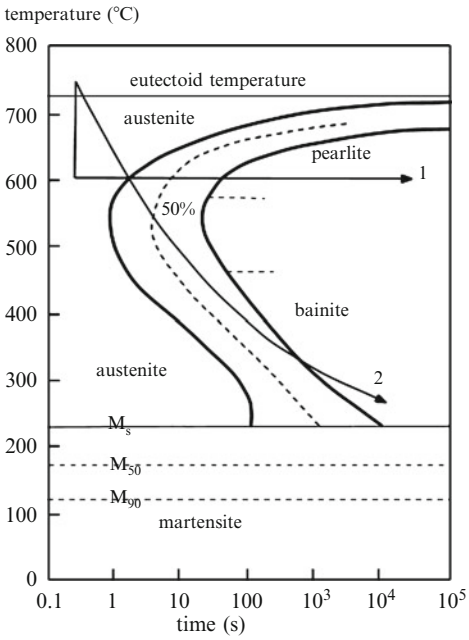


Fig. A2.27 Transformation of a steel on cooling

isothermal heat treatment. The TTT – Transformation, Temperature, Time – curve, obtained in isothermal conditions, can also be used. Trajectory 2, in contrast, is a cooling curve.

Derivation of non-isothermal transformation curves from isothermal runs up against a number of difficulties. The main one arises from the independently varying rates of nucleation and growth with temperature. The problem can be treated fairly

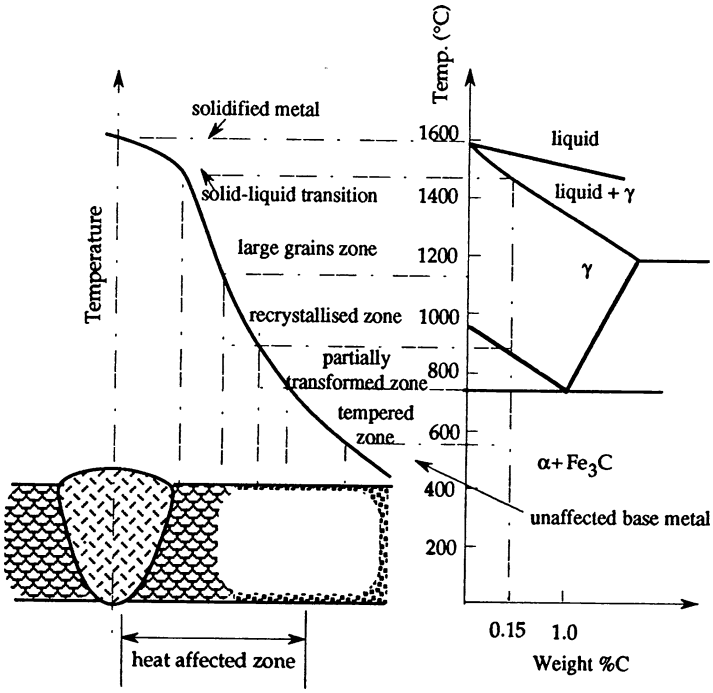


Fig. A2.28 Phase changes in the heat-affected zone in welding of a 0.15% carbon steel

simply provided that the instantaneous transformation rate is a specific function of the quantity transformed ζ and the temperature T, in which the variables are separated, that is, of the form:

$$\mathrm{d}\zeta / \mathrm{d}t = h(T) g(\zeta) \tag{A2.21}$$

Such a transformation is said to be *isokinetic*. The principle of *additivity according to linear cumulation* can then be applied, and in general we can write

$$\int_0^t \mathrm{d}t / t_\mathrm{a}(T) = 1 \tag{A2.22}$$

where $t_\mathrm{a}(T)$ is the time needed to reach an amount transformed ζ_a according to an isothermal diagram and t is the time to reach the same amount in non-isothermal conditions.

In general, this does not apply to cases where the parameter k in the Avrami equation (A2.20) can itself vary with time. However, in certain conditions, in particular when the nucleation rate is very high, it can give acceptably correct results. It is therefore very useful for treating particular problems, such as very rapid cooling after welding, as shown schematically in Fig. A2.28.

A2.7 Solidification

Solidification of a pure metal or an alloy is a phase change to which the theories developed in the preceding sections apply well. The practical importance of the microstructures produced in solidification is such as to justify a detailed study of the phenomenon.

A2.7.1 Nucleation in the Solid Phase

Using the theory of nucleation it is easy to evaluate the quantity $\Delta G_V = G_S - G_L$ as a function of the temperature difference $\Delta T = T_m - T$ at equilibrium. At $T = T_m$ we have:

$$\Delta G_V = G_S - G_L = 0 \quad \text{that is} \quad H_S - H_L = T_m (S_S - S_L) \tag{A2.23}$$

from which, assuming that near T_m both ΔH and ΔS are independent of T, it follows:

$$\Delta G_V = (H_S - H_L)(1 - T/T_m) = (H_S - H_L)\Delta T/T_m \tag{A2.24}$$

and hence the critical size of the embryos is $r^* = 2\gamma_{SL} T_m / L_m \Delta T$.

An accurate value of $(H_S - H_L)$ can be derived from the measurement of the latent heat of solidification L_m.

Homogeneous nucleation, initiated by embryos of the solid phase, can occur only at several hundred degrees of under-cooling, and such temperatures have not often been reached except in very special cases, such as with very fine, very pure droplets. Solidification usually takes place from only a few degrees of under-cooling, with *heterogeneous* nucleation, with the solid phase growing on foreign particles or outwards from the walls of the mould. It is clearly important to be able to control heterogeneous nucleation in practice; for light alloys it can be initiated by adding various elements at the moment of solidification, or Mg in the case of cast iron with spheroidal graphite.

A2.7.2 Growth of the Solid Phase

The equilibrium curves of Fig. A2.29 show the formation of solid, which in general is less rich than the liquid in the added element B. The liquid becomes richer in B and the solid does so in turn so as to remain in equilibrium. Finally there is liquid with composition $C_1{}^f$ in equilibrium with solid of composition C_0.

The theoretical condition for homogeneity of the liquid is quite well satisfied in practice since diffusion is fast and it is favoured by convection. The solid can remain homogeneous only by diffusion of B from the interface, which requires the cooling to be *very slow indeed*. In actual conditions, the diffusion is insufficient

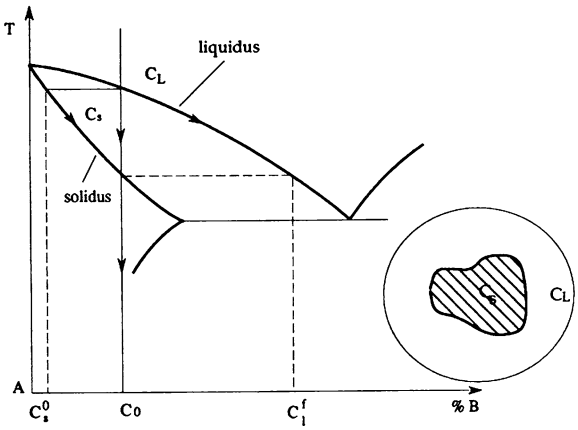


Fig. A2.29 Equilibrium diagram for solidification

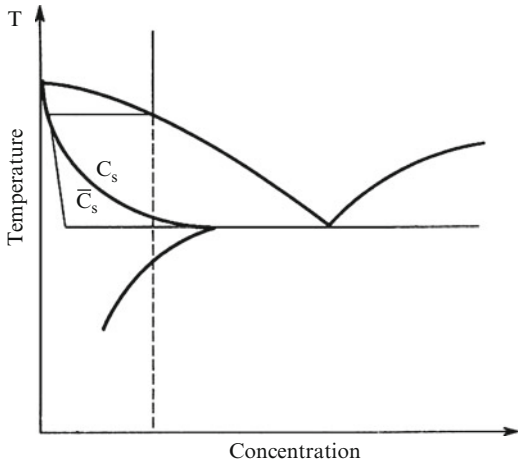


Fig. A2.30 Solidification phase diagram: equilibrium and out-of-equilibrium

and the solid is heterogeneous with an average composition that is a mean between its initial value C_S^0 and the interface value $\bar{C}_S$. Thus its composition follows a *real* solidus, as in Fig. A2.30, rather than the equilibrium solidus.

At the end of a real solidification, when $C_S = C_0$, the temperature is lower than the equilibrium value. Very often a small quantity of a eutectic is formed (as shown in Fig. A2.30), even if this would not be predicted by the equilibrium diagram; this has many and important practical consequences:

– a heterogeneous solid results, which has to be homogenised by diffusion at high temperature
– rolling and forging temperatures have to be reduced, otherwise there could be local re-melting resulting from melting of the eutectic ("burning the alloy")
– much segregation of poisonous elements, such as sulphur and phosphorous.

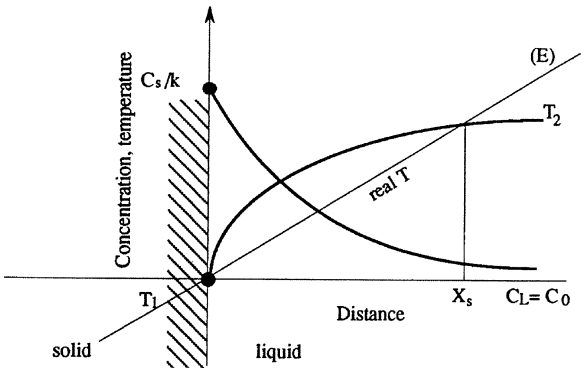


Fig. A2.31 Representation of structural under-cooling

Reducing the rate of solidification does not in practice give any significant improvement; in fact, it results in both a coarser structure – the effect of the nucleation – and diffusion distances that do not compensate for the expected effect of increasing the homogeneity of the material.

A2.7.3 Morphology of the Solid Phase

Ahead of the advancing front of solidification, the liquid is richer than the solid; at the interface the concentration in the liquid is C_S/k, where k is the *partition coefficient* between the solid and liquid phases. Thus we might say that if the concentration at a distance x from the front is $C(x)$, this will correspond to a liquidus temperature $T_2 > T_1$, where T_1 is the temperature corresponding to the interface (Fig. A2.31).

However, the actual temperature of the liquid, controlled by the rate of loss of heat from the mould containing the alloy, is given by the line E in the figure and is below T_2 over the distance X_S: we say there is a zone of *structural under-cooling*. This distance X_S is determined by the cooling conditions; it is small if the temperature gradient in the liquid is high, large if this is low.

The distance over which this structural under-cooling extends has a marked influence on the morphology of the solid formed: it causes alloys to form a *dendritic* structure on solidifying. As will be readily understood, any protuberance from the solidifying front is unstable; it will grow quickly up to the limit of the under-cooled zone, displacing solute laterally which can generate new instabilities, secondary and tertiary in order: this is shown diagramatically in Fig. A2.32. The main axis along which the dendrite grows is in the direction of the heat extraction and its length is of the order of X_S. The axes and arms of the three-dimensional dendritic structure are in the *rapid crystalline growth directions*, which for FCC metals are the <100> directions.

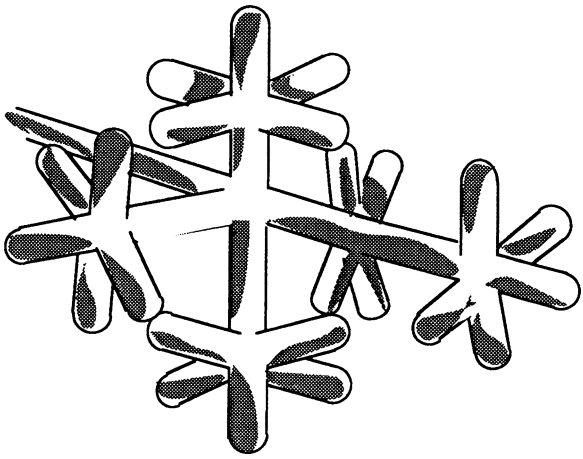


Fig. A2.32 A dendrite: secondary and tertiary branches develop from the primary stem, enabling heat to be exchanged more easily between solid and liquid

Circulating between the dendrites is a liquid rich in solute, from which there is what is called *minor segregation*. The circulation through the dendritic forest is slow, and the difficulty of maintaining a supply of fresh liquid can lead to the formation of micro-porosities.

To avoid instabilities of the dendritic type, and to ensure regular growth, a controlled, high thermal gradient – *cf.* line (E) in Fig. A2.31 – must be imposed, and a slow solidification achieved by gradual reduction of this. Turbine blades with improved characteristics, in which the solidification is oriented, are produced in this way, Fig. A2.33 showing the principle of fabrication. The solid-liquid interface is effectively planar; eliminating certain grains and initiating the solidification with a well-oriented nucleus can produce single-crystal blades.

A2.7.4 Solidification of Eutectics

This type of solidification, described in Figs. A2.34 and A2.35, is another illustration of the phenomena of diffusion in the liquid phase, of great practical importance in low-melting alloys and in the fabrication of oriented structures by unidirectional, controlled solidification.

When the liquid reaches the eutectic composition after depositing α (called the *primary*) from the solid solution it follows the α-liquidus, curve 1 in Fig. A2.34. The diagram shows that the under-cooling relative to the β-liquidus (curve 2) increases rapidly, all the more easily as the liquid at the interface becomes richer in B. There is then nucleation of β, which reduces the concentration of B locally – the passage from (1) to (2) – and results in a large under-cooling relative to α. Thus nucleation and growth of α alternates with that of β, as shown in Fig. A2.35. The growth of the

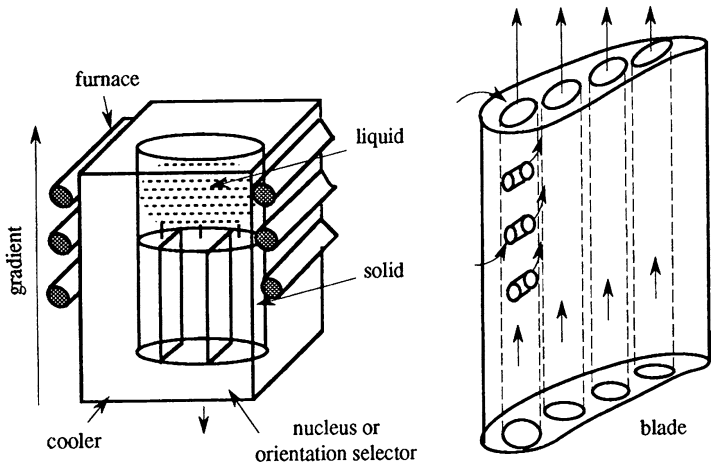


Fig. A2.33 Controlled solidification process for manufacturing turbine blades

Fig. A2.34 Solidification of a eutectic alloy

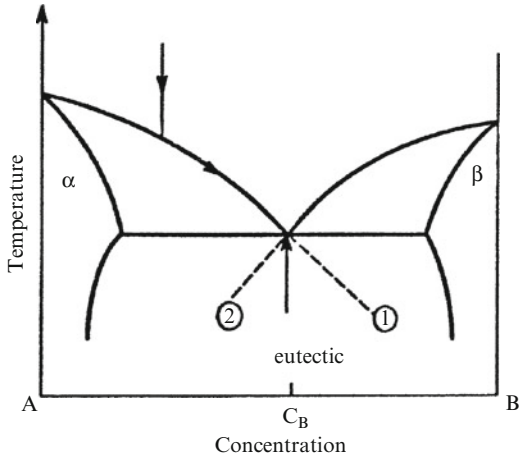
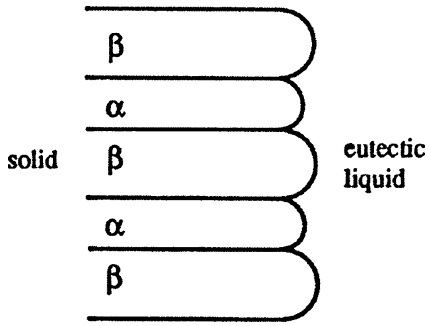


Fig. A2.35 Structure of a eutectic alloy



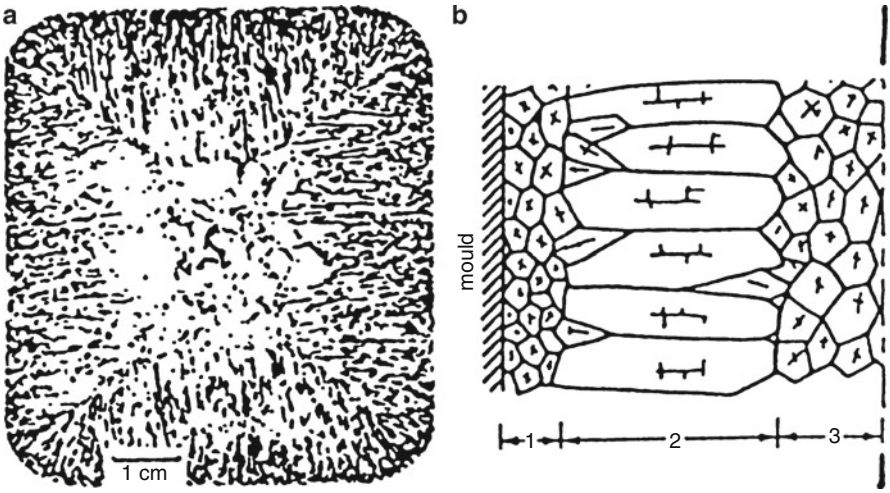


Fig. A2.36 (**a**) Transverse section through a steel ingot, showing solidification structures; (**b**) diagram showing the solidification zones from the surface in contact with the mould to the centre of the ingot. The axes of the dendrites are shown in the *grains*; zones are (1) skin, (2) basaltic, (3) equiaxial

eutectic implies that the liquid in front of α, which will be rich in B, exchanges, by diffusion, the elements A and B with the liquid in front of β, which will be rich in A.

The rate of solidification determines the diffusion length in the liquid and through this the details of the microstructure of the eutectic. If this rate is high the diffusion length is small and the structure is fine; if low, the length is greater and the structure is coarser.

A2.7.5 Structure of Solidified Material

In general there are three distinct zones in a piece of solidified material: skin, basaltic and equiaxial (Fig. A2.36).

A2.7.5.1 Skin Zone

When the molten metal comes into contact with a mould whose temperature is much lower than the liquidus, the under-cooling is very great. Nucleation occurs quickly and a dense network of dendritic crystals forms in the metal, randomly oriented. The structural under-cooling distance in this zone is small and the temperature gradient in the liquid is high.

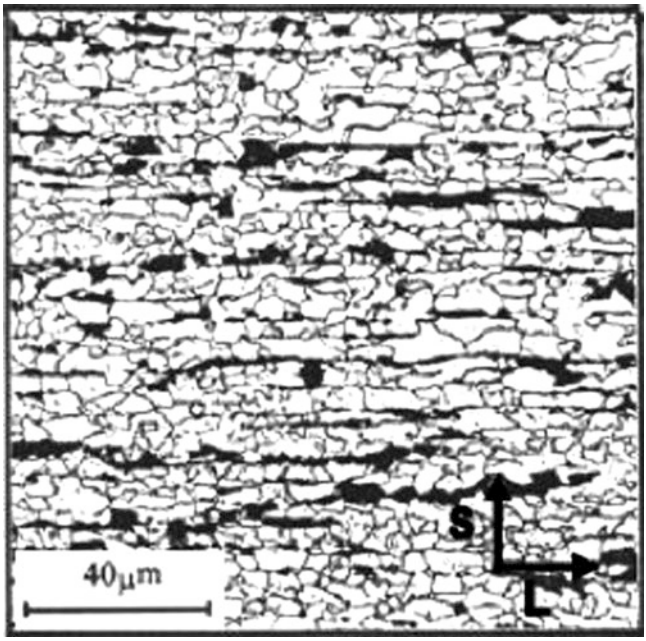


Fig. A2.37 Micrograph of a rolled X65 pipeline steel (0.09 C, 1.59Mn, 0.29Si, 0.05 V, 0.025Nb) with a yield strength of 440 MPa in the longitudinal (L) direction and 500 MPa in the transverse (T) direction. Notice banded structure of ferrite and pearlite due to the segregation of the alloy elements during solidification. S corresponds to the short transverse direction

A2.7.5.2 Basaltic Zone

The structural undercooling distance increases and dendritic growth continues, with preferential development of dendrites whose axes are in the direction of the heat flow. The length of the basaltic grains can reach several centimetres.

A2.7.5.3 Equiaxial Zone

At the end of the solidification the development of the basaltic zone is constrained by the nucleation of many crystals in the body of the remaining liquid, which is everywhere above melting point. This nucleation, often occurring on dendritic debris, is heterogeneous and the grains in this zone are randomly oriented. The zone is very rich in elements of the alloy: we have *major segregation*.

The major and minor segregations are responsible for the local compositional changes, which explain the variations in the microstructure, even in material that is afterwards rolled: Fig. A2.37 shows an example. It is very difficult to eliminate this segregation, even the minor, without some very lengthy homogenisation treatment.

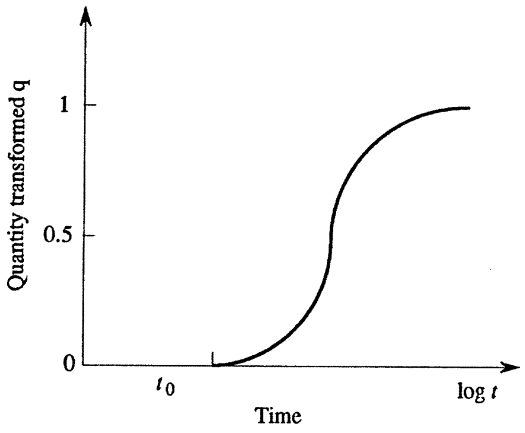


Fig. A2.38 Kinetic isotherm for precipitation; q is the fraction precipitated, t_0 is the incubation time required for the nucleation

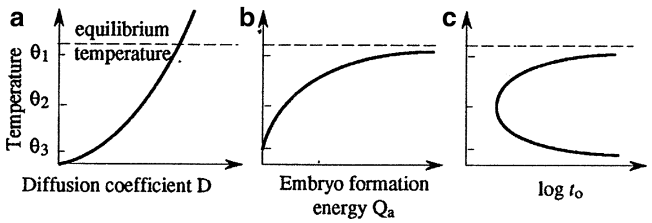


Fig. A2.39 Variation of incubation time t_0 with temperature, for isothermal precipitation. At a high temperature θ_1, the mobility of the solute atoms is high (**a**) but the incubation time is very long because the embryo formation energy is high. (**b**) At a low temperature θ_3 it is long again because the atomic mobility is low. The minimum time occurs at some intermediate value $\theta_2(c)$

A2.8 Precipitation

As in the liquid phase, precipitation in the solid phase is due to nucleation and growth. Its rate is controlled by diffusion in the solid phase. Embryos of the precipitates begin to form as ageing starts in the supersaturated matrix, their size being smaller the further the temperature of the alloy is below equilibrium (see Sect. A2.4). The amount precipitated increases with time in the manner shown in Fig. A2.38. The curve is of the Avrami type described in Sect. A2.6; t_0 in the figure is an incubation time, necessary for the nucleation.

The mechanism of isothermal precipitation by nucleation and growth depends essentially on two factors: the energy of formation of the nucleii and the mobility of the atoms. The lower the formation energy ΔG_c for an embryo of critical radius r^*, the shorter is the incubation time; but the rates of nucleation and growth increase with the atomic mobility, which itself increases exponentially with increasing temperature. The overall kinetics of the precipitation is thus determined by two opposing effects, as indicated in Fig. A2.39; this gives a qualitative explanation of the C-shape characteristic of isothermal transformation curves.

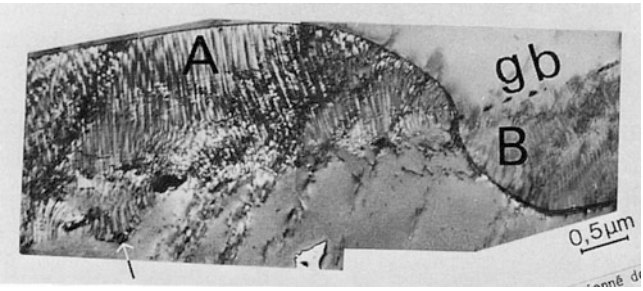


Fig. A2.40 Fe-Cr-Ni-Al aged for 10 min at 600°C. The two-phase (Fe-NiAl) zones A, B developed from the α/β boundary. The platelets of the NiAl phase are oriented with respect to the B grain (By courtesy of R. Taillard)

A2.8.1 The Two Types of Precipitation

We distinguish between *continuous* and *discontinuous precipitation*. In the first, the concentration of the solute in the mother phase falls continuously until the equilibrium value is reached.

Discontinuous precipitation occurs when, during the transformation, the crystal becomes divided into regions of two types. In those of one type the transformation is complete; whilst in those of the other type the supersaturated solid solution remains unchanged. This type of precipitation requires the nucleation and growth of cells of 2-phase products. The reaction progresses by the advance of the boundary separating the part that has been transformed from the part that has not, and this advance is shown by a sudden change in the properties of the matrix (composition, crystallographic parameters) – hence the name *discontinuous*. Discontinuous precipitation occurs generally in association with a recrystallisation, more precisely with a migration of a boundary; it is also referred to as a *cellular* reaction. Figure A2.40 shows an example, precipitation of the NiAl phase in a Fe-Cr-Ni-Al steel. This type of reaction commonly occurs when the level of supersaturation is low. The precipitates tend to make the material brittle, as we indicated in Chap. 1 (Fig. 1.45) in relation with Al-Li alloys, and steps are therefore usually taken to avoid it.

A2.8.2 Coherency Between the Precipitates and the Matrix

In continuous precipitation the individual precipitates grow larger by draining of solute atoms. The crystallographic orientation relations between the mother phase and the product determine the energy associated with the α/β interface and the elastic energy stored in the precipitate and the matrix. Products of three types can be envisaged, as in Fig. A2.41:

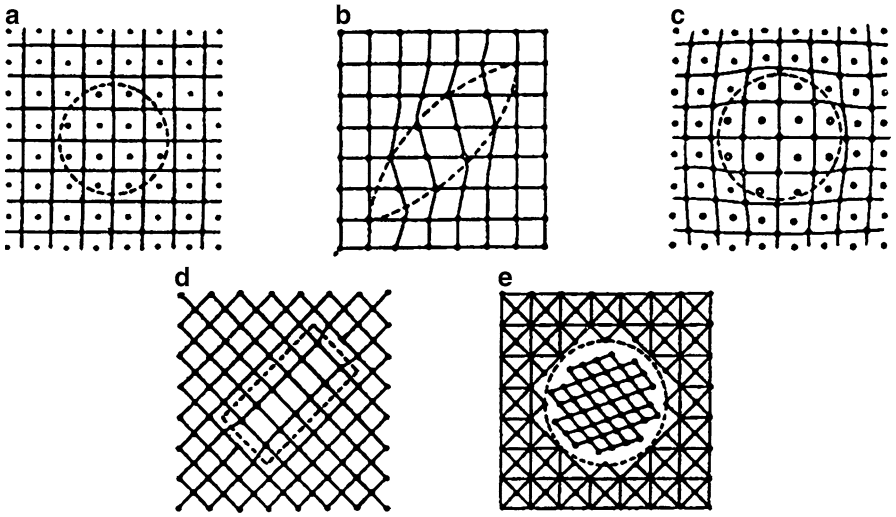


Fig. A2.41 Types of precipitation in the solid phase

1. *coherent*, in general having the same crystal structure as the matrix, cases (a) and (c) in the figure; or a different structure, case (b), that can be obtained from the mother phase by a shear.
2. *semi-coherent*, with a number of linear defects in the interface between the precipitate and the matrix: case (d).
3. *incoherent*, with a structure so different from that of the matrix that no coherent interface can be found: case (e).

Precipitates of the first two types often belong to the class of *metastable phases*. It seems that there is always a significant level of well-defined coherency between the precipitates and the matrix in the first stages of the process, and that this diminishes as the reaction advances; thus the first stages are so coherent that we should speak of *pre-precipitation* rather than precipitation. The clusters so formed, which are of very small size, about 10 nm, are called *Guinier-Preston zones* (Guinier 1938, 1939) in the case of light alloys; their shape can vary greatly according to size and to the electronic structures of the solvent and the solute: for example, spheres in Al-Mg, ellipsoids in Al-Zn, discs in Al-Cu.

Taking as an example the widely studied Al-Cu alloy (duralumin), for which we gave the equilibrium diagram in Sect. A2.2, and the metastable diagram in Fig. A2.21, the precipitation sequence is:

$$\text{Solid solution} \rightarrow \text{GP zones} \rightarrow \theta'' \rightarrow \theta' \rightarrow \theta\,(Al_2Cu)$$

Three metastable states have to be traversed before equilibrium is reached. The GP zones and the θ'' and θ' phases are formed on the {100} planes of the matrix, θ'' and θ' having a tetragonal structure. θ'' is coherent with the matrix and θ' is semi-coherent.

NOTES

1. Very advanced techniques are now available, which make it possible to follow the precipitation sequences in great detail: high-resolution (0.2−0.3 nm) electron microscope, neutron diffusion, etc. (see Chap. 1).
2. The problems of gaining a detailed knowledge of the decomposition of a solid solution are still very alive; they arise in many contexts, not only with metals but also with glasses.
3. Such knowledge is essential for the understanding of the relations between the mechanical properties on the one hand, such as monotonic and cyclic plasticity, damage (strength, stress corrosion), and on the other the microstructure of a material at any instant.

A2.9 Martensitic Transformations

In contrast to the phase changes described so far, martensitic transformations occur *without diffusion*. They exist in many metallic systems − steels, Fe-Ni and Cu-Al alloys, alloys of titanium and of cobalt, etc. To describe their general features we consider first the martensitic $\gamma \rightarrow \alpha'$ transformation in steels.

A2.9.1 General Features of Martensitic Transformations in Steels

Figure A2.42 is the TTT curve for a eutectoid (0.8 %C) steel. We see from this that if, starting from the stable austenitic (γ) domain, the steel is cooled very quickly, transformation of a new type, different from those that involve diffusion (pearlitic and bainitic transformations), begins to occur at the temperature M_S at the start of the phase change. This is the *martensitic transformation* in which the austenite is transformed into a new structure with the same composition, martensite.

In this transformation the carbon does not have time to diffuse so as to partition itself between the two phases corresponding to the equilibrium diagram, ferrite and cementite; it remains in place during the allotropic transformation γ (FCC) $\rightarrow \alpha'$ (BCC) of the austenite. In carbon steels the martensite is not in fact strictly centred cubic, but centred quadratic as in Fig. A2.43, where we show the particular location of the carbon atom. In a eutectoid steel with 0.8% carbon 7.2% of the cells have a carbon atom in this position and the c/a ratio for the lattice is 1.04. Martensite is a ferrite so highly supersaturated with carbon that these atoms bring about a large distortion of the centred cubic lattice. The resulting microstructure is very characteristic (Fig. A2.44), the martensite appearing in the form of *needles, laths* or *plates.*

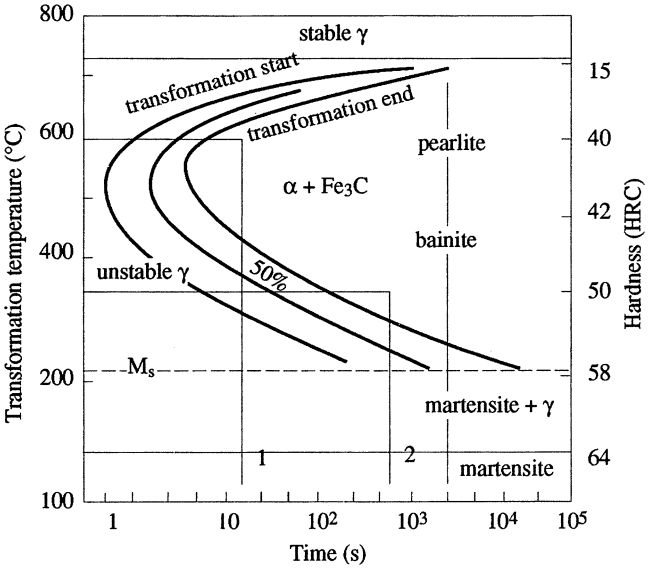


Fig. A2.42 Isothermal transformation curves for a eutectoid steel (0.8% C); austenising 30 min. at 850°C

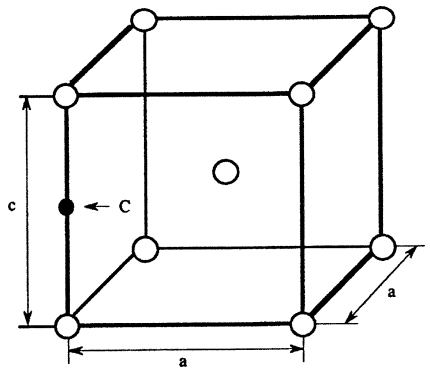


Fig. A2.43 Centred quadratic martensite cell for a carbon steel, showing the location of the carbon atom

Mainly because of the large volume change (1−4%) accompanied by the transformation from austenite to martensite, the latter is greatly deformed and includes a large number of defects (dislocations and twins); its hardness is mainly a function of the carbon content of the austenite (Fig. A2.45).

The rate of the martensitic transformation is very high, approaching that of sound in the material, around 10^3 m/s. Unlike precipitation reactions, it involves not diffusion but a "co-operative" movement of the atoms, which does not require thermal activation. Thus the fraction transformed is, in general, independent of time and depends only on the temperature, which must be less than M_S (at which the phase change starts).

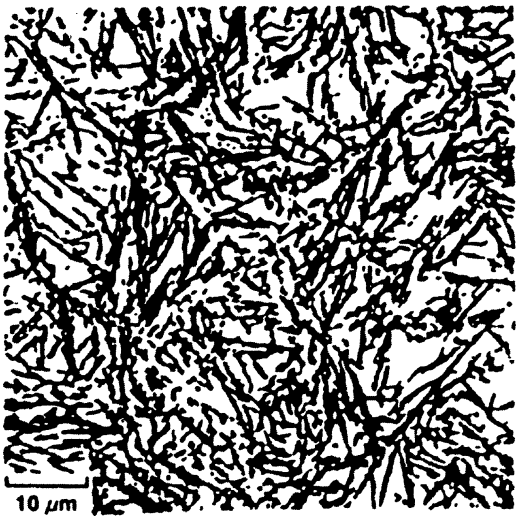


Fig. A2.44 Optical micrograph showing the martensite structure in a 0.8 %C steel

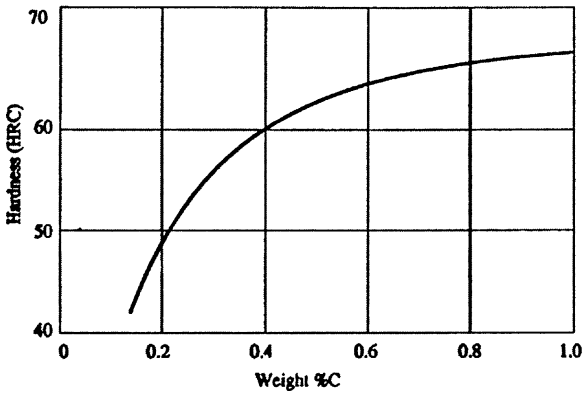


Fig. A2.45 Dependence of the hardness of quenched martensite on the carbon content of the steel

The passage from the crystalline structure of the γ phase to the martensitic phase is brought about a crystallographic mechanism of shear type, which moves the atoms from their positions in the initial lattice to those in the lattice of the transformed product; for carbon steels this shear stress is of the order of 20%. Support for the belief that such a mechanism is in action is given by the following:

1. the change of shape, or the relief effects seen on what were plane surfaces in the mother phase (Fig. A2.46).
2. the displacement of fine scratches by platelets of martensite (Fig. A2.47).
3. the crystallographic relations found to exist between the austenite and the martensite: in steels with a carbon content exceeding 1.5%; the Kurdjumov-Sachs orientations relationships are often found to hold, that is: $(111)_\gamma // (011)_{\alpha'}$ and $\left[0\bar{1}1\right]_\gamma // \left[1\bar{1}1\right]_{\alpha'}$.

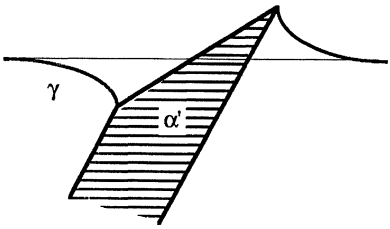


Fig. A2.46 Homogeneous shear associated with the $\gamma \rightarrow \alpha'$ transformation, producing a relief effect on the surface

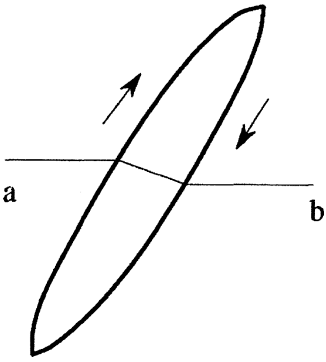


Fig. A2.47 Displacement of lines a, b by shear associated with the martensitic transformation

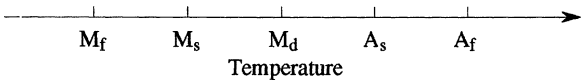


Fig. A2.48 Defining temperatures for the martensitic transformations

A2.9.2 Critical Points of the Transformation, and a Note on Thermodynamics

A2.9.2.1 Temperatures

On cooling, if the transformation is *purely athermal* (meaning no time effects) it can be described by the two temperatures M_s, at which it starts, and M_f, at which it ends. On heating, provided that this is rapid enough, the inverse transformation $\alpha' \rightarrow \gamma$ can itself be martensitic. This is difficult to achieve in the case of carbon steels because of the speed with which the carbon precipitates; it is less difficult in other systems, such as Fe-Ni and Co-Ni. Correspondingly, this inverse transformation can be defined by its starting and finishing temperatures A_s, A_f (Fig. A2.48)

In steels the difference $M_s - A_s$ is considerable, from 100°C to 400°C, whereas in other systems, such as *thermoelastic martensites*, it is only a few degrees. It expresses a large departure from the thermodynamic equilibrium of the two phases, which latter can be reached by raising the temperature above M_s − if that can be done without initiating any diffusional transformation − and applying a plastic deformation. This leads to defining a temperature M_d above which it is not possible to initiate the martensitic transformation by plastic deformation. As a general rule

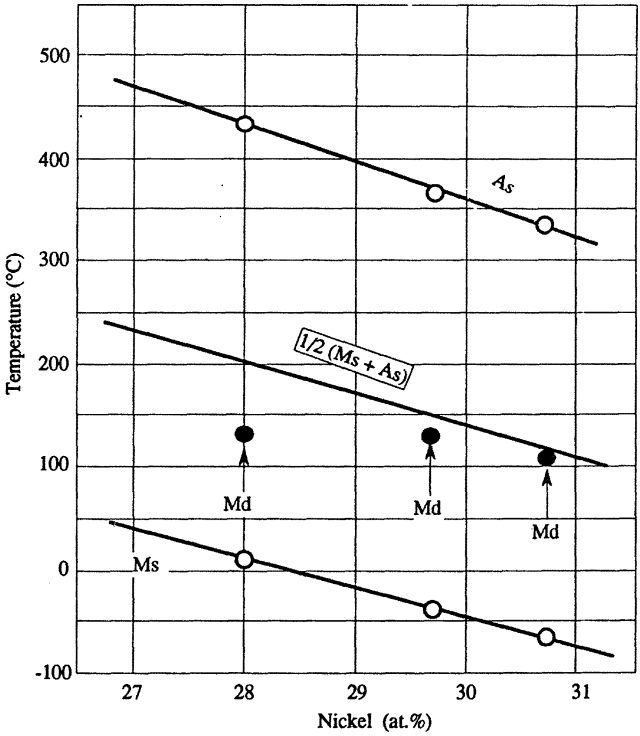


Fig. A2.49 Transformation points for the spontaneous martensitic transformation: M_s cooling, A_s heating, M_d deformation-induced for Fe-Ni alloys

we can take $M_d = ½(M_s + A_s)$ – see Fig. A2.49 – and M_d is approximately the equilibrium temperature for the phases γ and α′, that is, for which $\Delta G_{\gamma\to\alpha'} = 0$

Empirical relations have been found which enable M_S to be calculated from the known composition by weight of the austenite; the Stevens-Haynes formula, often used for steels, is:

$$M_s = 561 - 474\,(\%C) - 33\,(\%Mn) - 17\,(\%Ni) - 17\,(\%Cr) - 21\,(\%Mo) \quad (\text{in}^\circ C)$$

A2.9.2.2 Thermodynamics

For transformations that involve diffusion and some chemical change to the phases present the phase equilibrium condition is expressed by the equality of the chemical potentials of the various constituents in their various phases. The situation is different in the case of martensitic transformations, since in these there are no changes of composition and the equilibrium is expressed by the equality of the free energies of the two phases. For ferrous alloys the variation in free energy is:

$$\Delta G_{\gamma\to\alpha'} = G_\gamma - G_{\alpha'} \tag{A2.25}$$

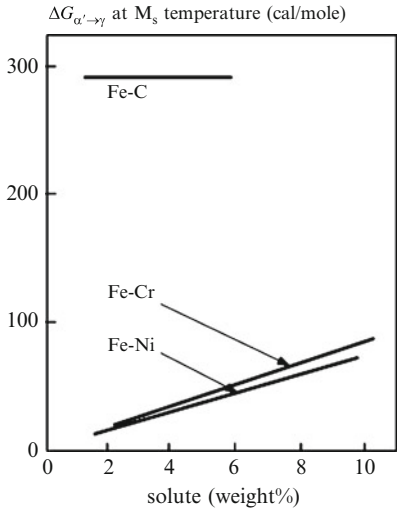


Fig. A2.50 Variation of free energy $\Delta G_{\alpha' \to \gamma}$ with temperature M_s for alloys Fe−C, Fe−Ni, Fe−Cr

The necessary condition for the transformation is that ΔG is negative. In general, for a ferrous alloy containing a dissolved element A (*e.g.* C, Ni, Cr, Mn, Si):

$$\Delta G_{\gamma \to \alpha'} = (1 - X)\,\Delta G^{\mathrm{Fe}_{\gamma \to \alpha'}} + X \Delta G^{\mathrm{A}_{\gamma \to \alpha'}} + \Delta G_{\mathrm{m}}{}^{\gamma \to \alpha'}$$

where $\Delta G^{\mathrm{Fe}_{\gamma \to \alpha'}}$ is the variation in free energy for pure iron.

$\Delta G^{\mathrm{A}_{\gamma \to \alpha'}}$ is the variation in free energy for pure A.
$\Delta G_{\mathrm{m}}{}^{\gamma \to \alpha'}$ is the variation in free energy for the mixture.

Only the first term in this relation is known. Kaufman and Cohen (1958) suggested the following relation: $\Delta G_{\alpha \to \gamma} = 1{,}202 - 2.63.10^{-3}\,T^2 + 1.54.\ 10^{-6}\,T^3$ (cal/mol)

The two other terms cannot always be found experimentally and have to be deduced from models of the solid solution.

For steels, ΔG associated with martensitic transformations is large, around 300 cal/mol (Fig. A2.50). This is a consequence, as we indicated earlier, of the considerable distortions that accompany the transformation.

A2.10 Further Reading

Ashby MF, Jones DRH (1986) An introduction to microstructures. Processing and design. Pergamon Press, Oxford

Barralis J, Maeders G (2002) Précis de métallurgie, Elaboration, Structures – Propriétés, Normalisation. Nathan, Paris

Bhadeshia HKDH (2001) Bainite in steels, 2nd edn. Institute of materials, London

Bhadeshia HKDH, Honeycombe RWK (2006) Steels – microstructure and properties, 3rd edn. Butterworth-Heinemann, Boston
Cahn RW, Haasen P, Kramer EJ (eds) (1992) Constitution and properties of steels. Materials science and technology. A comprehensive treatment. VCH, Weinheim, New York, Bassel, Cambridge
Cahn RW, Haasen P, Kramer EJ (eds) (1991) Phase transformations in materials. Materials science and technology. A comprehensive treatment. VCH, Weinheim, New York, Bassel, Cambridge
Christian JW (1965) The theory of transformations in metals and alloys. Pergamon Press, Oxford
Desré P (2010) Thermodynamique des matériaux : équilibre de phases et métastabilité. EDP Sciences, Les Ulis
Dieter GE (1988) Mechanical Metallurgy. Mc Graw-Hill Book Company, New York
Easterling KE (1983) Introduction to the physical metallurgy of welding. Butterworth, London
Easterling KE (1992) Introduction to the physical metallurgy of welding, 2nd edn. Butterworth-Heinemann, Oxford
Frost HJ, Ashby MF (1982) Deformation mechanism maps. Pergamon, Oxford
Haasen P (1996) Physical metallurgy. Cambridge University Press
Kostorz G (2001) Phase transformations in materials. Wiley-VCH
Kurz W, Fisher DJ (1986) Fundamentals of solidification. Trans. Tech. Publications, Switzerland
Lesoult G (2010) Thermodynamique des matériaux. Presses polytechniques et universitaires romandes, Lausanne
Massalski T (1990) Binary phase diagrams. ASM International, Materials Park.
Phase Transformations (1968) American Society for Metals, Materials Park.
Porter DA, Easterling KE (1992) Phase transformation in metals and alloys, 2nd edn. Chapman & Hall, London
Porter DA, Easterling KE, Sherif MY (2009) Phase transformations in metals and alloys. CRC Press
Sims CT, Hagel WC (1972) The superalloys. Wiley, New York, London, Sidney, Toronto

A2.11 Exercises

A2.11.1 Equilibrium Diagram, Purification by Zone Melting

A beam of an alloy AB, whose equilibrium diagram is given below, is placed in a furnace, which is moved slowly so that a moving liquid region is formed. Use the diagram to explain the redistribution of the solute B between the liquid and solid phases. Can the operation be repeated? if so, in what circumstances? This is the method used to purify Ga and Si semiconductors and to produce certain high-purity metals (Fig. A2.51).

A2.11.2 Steel Microstructures

Steels of various carbon contents are cooled slowly from the austenitic state. Figure A2.52 gives micrographs of the structures. What is the carbon content of each? Use the Fe–Fe_3C part of the iron-carbon equilibrium diagram to justify your answers.

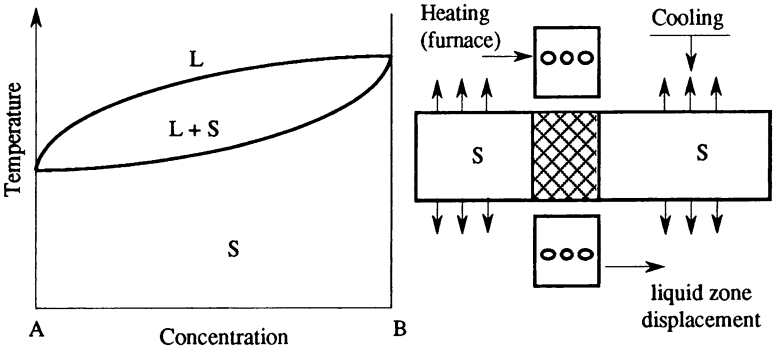


Fig. A2.51 Melting zone purification

Stage of treatment	% residual austenite	
	Alloy 12−5	Alloy 12−7
A	0	0
B1	5	23
B2	0	8
C	25	4.7

A: austenitising at 1,030°C, quenched in air
B1 tempered at 700°C for 1 h + recooled in air
B2 as B1, recooled in liquid nitrogen
C tempered at 600°C for 5 h after B2

A2.11.3 Martensitic Transformation

Two steels with 12% Cr and 5% and 7% Ni respectively are thermally treated. They are quenched from 1,000 °C: what can be predicted from the Fe−Cr−Ni equilibrium diagram given in Fig. A2.53? (martensite laths are formed in both cases). They are then subjected to a repeated annealing: 1 h. at 700°C − cooling in air − cooling to −196°C − 5 h at 500°C, after which the residual austenite content is measured at ambient temperature. The data are given in the table. Explain what has happened; would you expect an increase in the nickel content in the austenite? Electron microscopy shows that the γ phase is distributed in bands: how do you explain this? Figure A2.54 gives the results of tensile tests to determine the yield strengths of the two specimens: how do you explain the rapid variation of this with temperature?

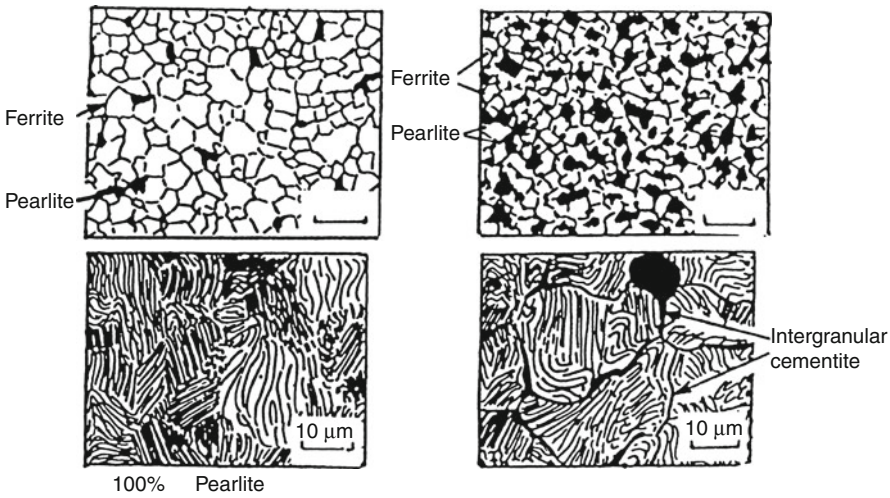


Fig. A2.52 Micrographs of structure of 4 steels

A2.11.4 Steel Microstructures

Various heat treatments have been applied to a Cr–Mo steel (0.39C, 1.5Cr, 0.5Mo), the TTC curve for which is given in Fig. A2.55. This steel is much used in mechanical engineering for items such as gears. Why is such care taken over the conditions for austenitising (maintaining in the γ phase) and the resulting grain size of the austenite? Four structures resulting from various cooling rates are shown in the figure. Identify these microstructures.

A2.11.5 Surface Treatment: Cementation

Carbon cementation. Cementation is a thermo-chemical treatment, the aim of which is to increase the carbon content of the sample at the surface. This enrichment is achieved by keeping the surface in contact with a carbon-donating material in powder, paste or liquid form. The diffusion process is always followed by the thermal treament of quenching – why is this? The quenching creates residual compressive stresses in the surface – why?

Suppose the operation takes place in the gaseous phase (Fig. A2.56); show that the solution to Fick's second equation is:

$$\frac{C - C_0}{C_S - C_0} = \text{erf}\left(\frac{x}{2\sqrt{Dt}}\right)$$

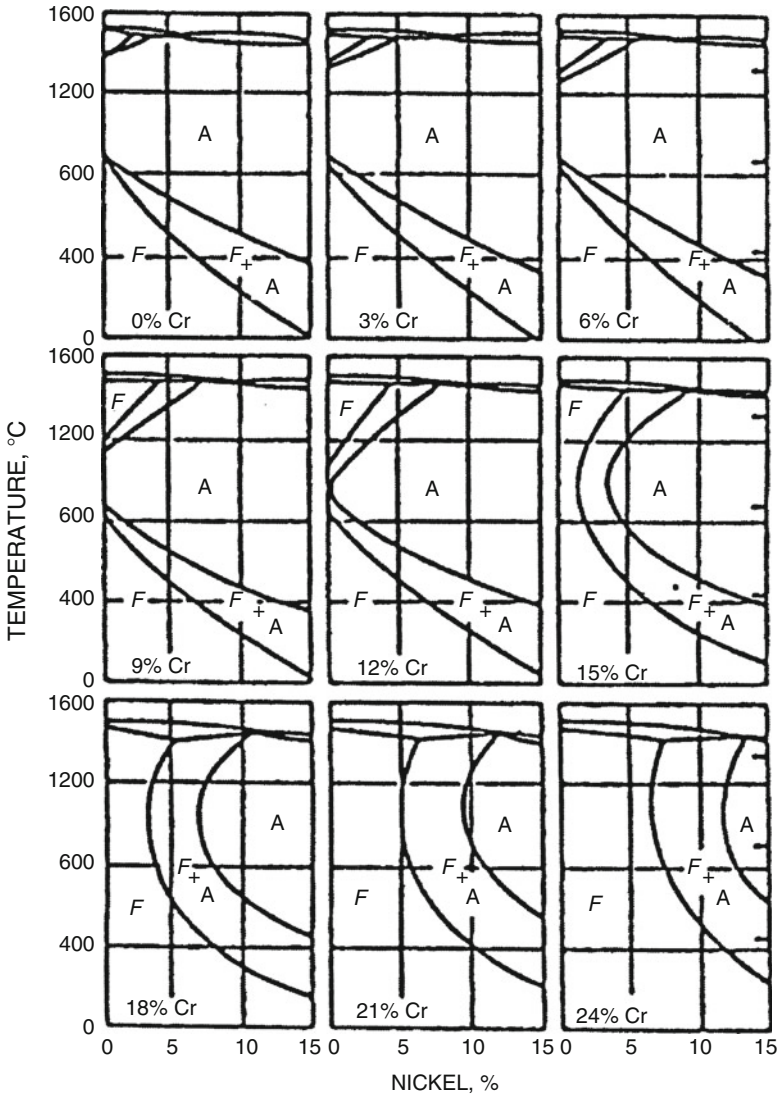


Fig. A2.53 Equilibrium diagram for the ternary system Fe−Ni−Cr (A = austenite, F = ferrite)

where $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-u^2\right) \mathrm{d}u$. C_0 is the initial carbon content of the steel, C_S the saturation value for the γ phase at the temperature of the treatment.

Figure A2.56 gives carbon concentration profiles for a steel with $C_0 = 0.15\%$, $C_S = 1.3\%$ at 925 °C. Use these to find the diffusion coefficient of carbon in the γ phase

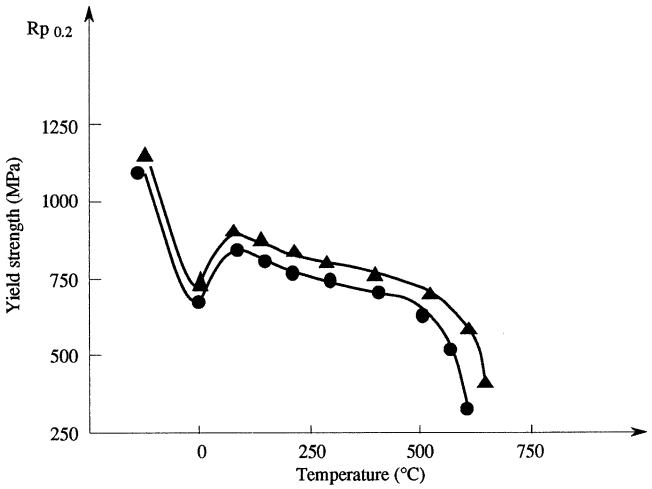


Fig. A2.54 Variation of yield strength with temperature

and compare your value with that given in the tables in Sect. A2.3.2 (Table A2.4). What do you think of the result? Why is it necessary to conduct this process at a temperature above that of the eutectoid?

A2.11.6 Solidification

Solidification of a Fe−C−Ni alloy has resulted in segregations of the additive elements C, Ni. The distances over which these segregations occur depend on the conditions during the solidification and on the growth of dendrites. Taking the data in Table A2.4 concerning the diffusion coefficients for these elements in Fe_γ, show that there is a possibility of homogenising this alloy by holding it in the austenitic phase for a long time. What quicker methods can you suggest for the process, bearing in mind that band structures such as those shown in Fig. A2.37 are to be avoided?

A2.11.7 Hardening by Precipitation and Coalescence of the Precipitates in a Ferritic Stainless Steel

Experiments have been made with steels of the following percentage composition:

(a) Fe − 19.6Cr − 2.03Ni − 0.97Al (b) Fe − 19.6Cr − 4.15Ni − 1.87Al

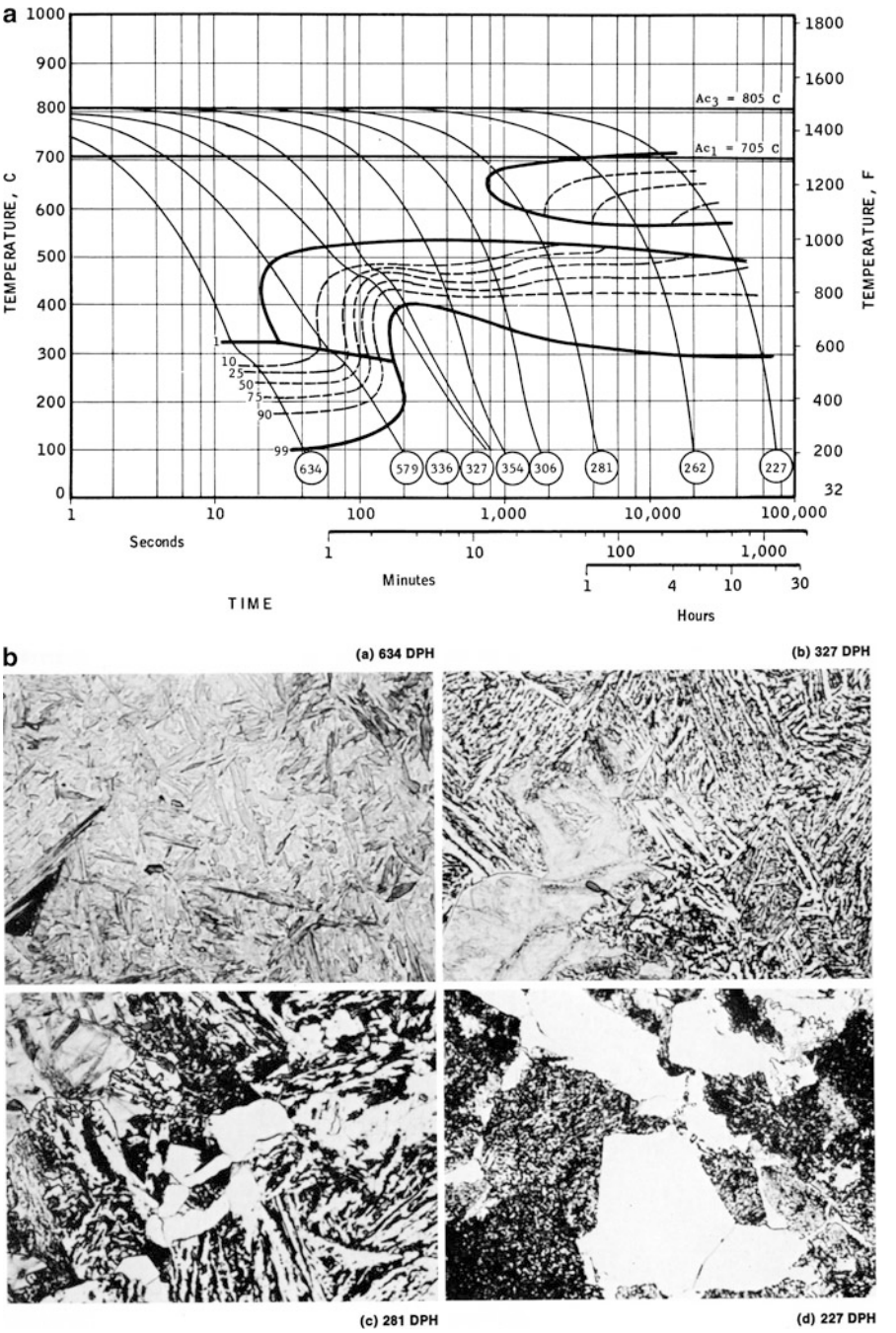


Fig. A2.55 (**a**) Continuous cooling transformation diagram for a steel containing 0.39% C, 0.37% Si, 1.45% Mn and 0.49% Mo austenised at 835°C for 20 mn; (**b**) Typical microstructure at various rates) (From W.W. Cias. Phase transformation kinetics and hardenability of medium-carbon alloy steels, Climax Molybdenum Company)

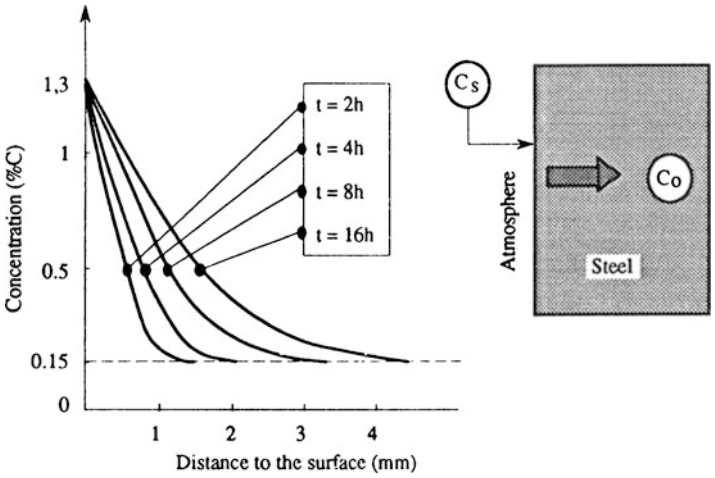


Fig. A2.56 Carbon concentration profiles after gaseous cementation for various periods of time

1. Explain the choice of these compositions, knowing that the hardening phase produced by the precipitation has the composition NiAl (structure B_2, CsCl, see Annex 1). Is it normal for these steels to have a BCC structure?
2. Samples of the steels have been held at 1,150°C for 1 h; what structure will they then have? After this they are cooled quickly and then aged at various temperatures. The curves of Fig. A2.57 give the variations of hardness with time of ageing and the micrographs of Fig. A2.58 show the precipitation of the NiAl phase.
    1. Is the form of the hardness curves what you would expect?
    2. Comment on the legend on the micrographs. Why is there a diffraction spot of structure of the type (001) in the [001] section?

A2.11.8 Precipitation

Figure A2.59 shows the variation of precipitate size with time at various temperatures, suggesting a growth law of the form:

$$A = \frac{64\gamma D\, C_{\mathrm{E}}^{\infty}\, \Omega^2}{9\mathrm{k}T}$$

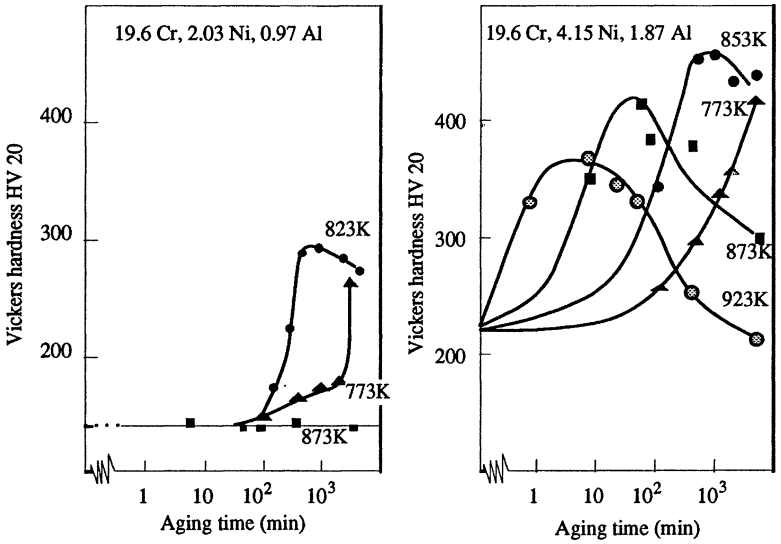


Fig. A2.57 Effect of time and temperature of ageing on hardness of two steels (By courtesy of R. Taillard)

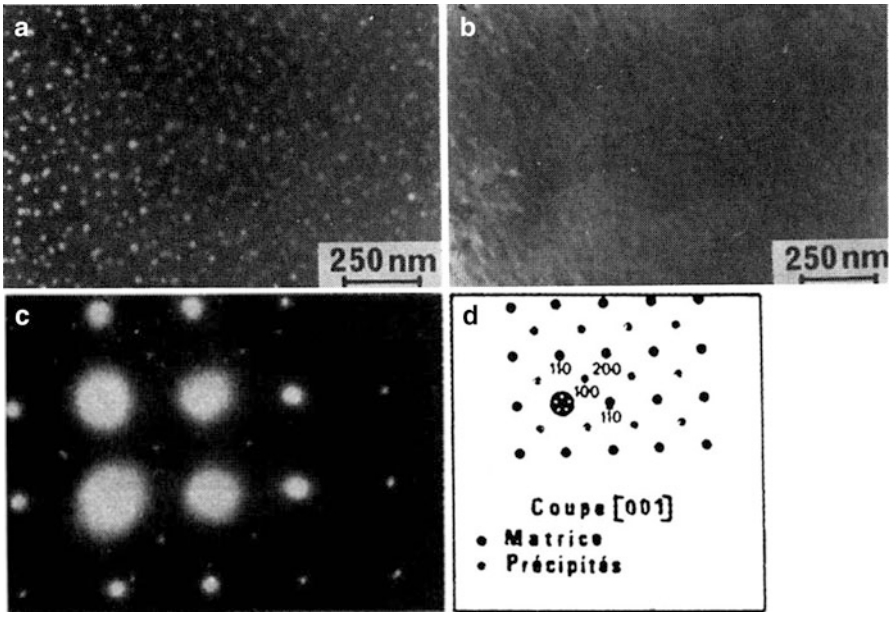
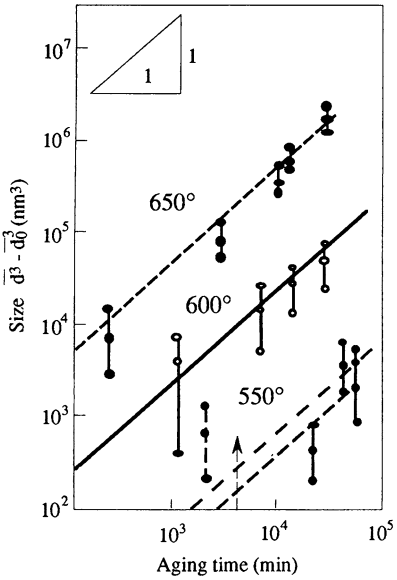


Fig. A2.58 Electron micrographs of NiAl precipitates: (**a**) light field (**b**) dark field (**c**) and (**d**) are diffraction conditions (By courtesy of R. Taillard)



Temperature °K	C_E^{∞} (T) mol/m³
823	$1.36\ 10^3$
873	$1.70\ 10^3$
923	$2.65\ 10^3$

Fig. A2.59 Variation of size of NiAl precipitates in Fe-Cr-Al alloy (By courtesy of R. Taillard)

Deduce the value of the apparent activation energy Q_d for diffusion of Al and Ni in α−iron (Fe−20Cr), assuming that γ_S and Ω are independent of temperature. The variation of the solubility has been determined − how could this be done? – and the results tabulated; discuss the values found for Q_d.

A2.11.9 Effect of Applied Stress on the Morphology of Precipitates in Nickel-Based Alloys

As Fig. A2.60 shows, the morphology of single-crystal materials containing a large volume fraction of precipitates ($f_{\gamma'}$ about 50–60%) is considerably changed by the application of stress. In this case (the alloy Udimet 700), tension leads to a rafting process and a redistribution of precipitates, initially cubic and distributed isotropicaly, into layers perpendicular to the direction of the stress. In contrast to this, compression rearranges them into rods aligned parallel to the stress. These changes have an important effect on the behaviour of the material in creep.

Give a qualitative explanation of these changes, bearing in mind the following:

– the precipitates are coherent with the matrix
– any departure from coherence, measured by the relative difference in the values of the parameters for the γ' phase and the matrix, $\delta = (a_{\gamma'} - a_{\gamma})/a_{\gamma}$, can be positive, negative or zero.

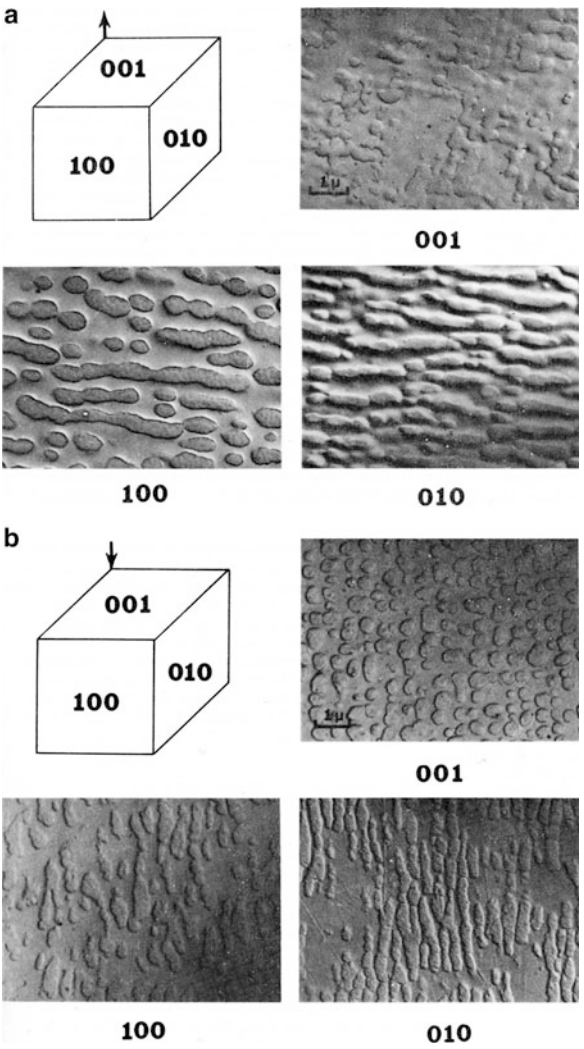


Fig. A2.60 Morphology of coalescence in single crystals of U−700, $T = 954°C$, $\sigma = 155$ MPa, $t = 100$ h; (**a**) tension; (**b**) compression (Tien JK, Copley SM (1971) Metall Trans 2:215–219)

– the elastic constants, Young's modulus in particular, are different for the precipitates and the matrix.

Further information can be found in:

Pineau A (1976) Influence of uniaxial stress on the morphology of coherent precipitates during coarsening-Elastic energy considerations. Acta Metall 24:559
Nabarro FRN (1996) Rafting in superalloys. Metall Mater Trans 27A:513–530

Appendix C
Annex 3: Continuum Mechanics: Basic Concepts and Equations

Main Formulae

NB: all refer to small perturbations

Strains

$$\underline{\underline{\varepsilon}} = \frac{1}{2}\left(\mathrm{grad}\,\underline{u} + \mathrm{grad}^{\mathrm{T}}\underline{u}\right)$$

Rectangular cartesian coordinates: $\varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$

$$\mathrm{d}s^2 - \mathrm{d}s_0^2 = 2\mathrm{d}\underline{x}.\underline{\underline{\varepsilon}}.\mathrm{d}\underline{x}$$

$$e_{\underline{u}} = \underline{u}.\underline{\underline{\varepsilon}}.\underline{u}$$

$$\Delta V/V = \mathrm{Tr}(\underline{\underline{\varepsilon}})$$

Compatibility: $\varepsilon_{22,33} + \varepsilon_{33,22} = 2\varepsilon_{23,23}$ etc.

$$\varepsilon_{23,31} + \varepsilon_{31,23} = \varepsilon_{12,33} + \varepsilon_{33,12} \text{ etc.}$$

Boundary conditions: $\underline{u}(P) = \underline{u}^{\mathrm{g}},\ P \in \partial V,\ \underline{u}^{\mathrm{g}}$ given on ∂V

Stresses

$$\underline{T}(M, \underline{n}) = \underline{\underline{\sigma}}(M).\underline{n} = \sigma\underline{n} + \tau\underline{t} \quad (\sigma = \underline{n}.\underline{\underline{\sigma}}.\underline{n})$$

Equilibrium and motion: $\underline{\underline{\sigma}} = \underline{\underline{\sigma}}^{\mathrm{T}}$, $\mathrm{div}\underline{\underline{\sigma}} + \rho\underline{f} = \rho\underline{\gamma}$

Boundary conditions: $\underline{\underline{\sigma}}(P).\underline{n} = \underline{T}^{\mathrm{g}},\ P \in \partial V,\ \underline{T}^{\mathrm{g}}$ given on ∂V

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

Linear isotropic elasticity

$$\begin{cases} \underline{\underline{\sigma}} = \lambda \mathrm{Tr}(\underline{\underline{\varepsilon}})\underline{\underline{\delta}} + 2\mu\underline{\underline{\varepsilon}} \\ \sigma_{ij} = \lambda\varepsilon_{kk}\delta_{ij} + 2\mu\varepsilon_{ij} \\ \underline{\underline{\varepsilon}} = [(1+\nu)/E]\,\underline{\underline{\sigma}} - (\nu/E)\mathrm{Tr}(\underline{\underline{\varepsilon}})\underline{\underline{\delta}} \qquad \varepsilon_{ij} = [(1+\nu)/E]\,\sigma_{ij} - (\nu/E)\sigma_{kk}\delta_{ij} \end{cases}$$

$$E = 2\mu(1+\nu) = 3k(1-2\nu) \quad \lambda = k - 2\mu/3 = E\nu/(1+\nu)(1-2\nu)$$

$$\text{Navier equations:} \begin{cases} (\lambda+\mu)\ \mathrm{grad}\,\mathrm{div}\underline{u} + \mu\nabla^2\underline{u} + \rho\underline{f} = \rho\underline{\gamma} \\ (\lambda+\mu)u_{k,ki} + \mu u_{i,kk} + \rho f_i = \rho\partial^2 u_i/\partial t^2 \end{cases}$$

A3.1 Deformations

A3.1.1 Strain Tensor for a Solid

Let $\underline{a}(a_1, a_2, a_3)$ be the initial position of a point M of the body and $\underline{x}(x_1, x_2, x_3)$ its present position; the transformation is defined by $\underline{x}(\underline{a}, t)$; $\underline{u}(\underline{a}, t) = \underline{x}(\underline{a}, t) - \underline{a}$ is the displacement vector. The second-order transformation gradient tensor $\underline{\underline{F}}$ is defined by

$$\underline{dx} = \underline{\underline{F}}.\underline{da} \tag{A3.1}$$

A3.1.1.1 Rectangular Cartesian Coordinates

In rectangular cartesian coordinates, we have, with $\otimes$ denoting a tensorial product

$$\underline{\underline{F}} = \frac{\partial x_i}{\partial a_j}\underline{e}_i \otimes \underline{e}_j \quad \left(F_{ij} = \frac{\partial x_i}{\partial a_j}\right) \tag{A3.2}$$

This is an asymmetrical tensor; starting from this we define strain tensors in Lagrange ($\underline{\underline{\Delta}}$) or Eulerian ($\underline{\underline{A}}$) variables:

$$2\Delta_{ij}\mathrm{d}a_i\mathrm{d}a_j = |\mathrm{d}\underline{x}|^2 - |\mathrm{d}\underline{a}|^2 = 2A_{ij}\mathrm{d}x_i\mathrm{d}x_j \tag{A3.3}$$

These tensors characterise completely the deformation of the solid, that is, they enable us to find the changes in lengths, angles and volume at any point. They are symmetrical, and therefore there is a coordinate system in which they are diagonal. In cartesians

$$2\Delta_{ij} = \frac{\partial x_k}{\partial a_i}\frac{\partial x_k}{\partial a_j} - \delta_{ij} = \left(\underline{\underline{F}}^{\mathrm{T}}.\underline{\underline{F}} - \underline{\underline{\delta}}\right)_{ij} \tag{A3.4}$$

which can be expressed also in terms of the displacement

$$\Delta_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial a_j} + \frac{\partial u_j}{\partial a_i} + \frac{\partial u_k}{\partial a_i}\frac{\partial u_k}{\partial a_j}\right) \tag{A3.5}$$

With the small perturbation hypothesis, this becomes

$$\varepsilon_{ij} = \Delta_{ij}^{\text{lin}} = \frac{1}{2}\left(\frac{\partial u_i}{\partial a_j} + \frac{\partial u_j}{\partial a_i}\right) \tag{A3.6}$$

where

ε_{11} is the length variation in the direction $\underline{e}_1$
$2\varepsilon_{12}$ is the angular variation of $(\underline{e}_1, \underline{e}_2)$
$\theta = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} = \text{Tr}(\underline{\underline{\varepsilon}})$, the relative volume variation $\Delta V/V$.

A3.1.1.2 Orthogonal Curvilinear Coordinates

$\underline{M}(\theta_i)$ defines a point M in an orthonormal system of curvilinear co-ordinates (θ_i):

in cylindrical polars $(\theta_i) \rightarrow (r, \theta, z)$
in spherical polars $(\theta_i) \rightarrow (r, \theta, \varphi)$

Write $\underline{E}_i = \partial \underline{M}/\partial\theta_i$, the orthogonal coordinate system related to the parameter θ_i and $\underline{e}_i = \underline{E}_i/|\underline{E}_i|$ the associated physical orthogonal coordinate system.

The transformation takes the point $\underline{M}(\theta_i)$ to $\underline{M}'(\theta_i)$; from $\underline{\mathrm{d}M'} = \underline{\underline{F}}.\underline{\mathrm{d}M}$ we get

$$\begin{aligned} &\underline{\underline{F}} = (1/|\underline{E}_i|)\left[\partial \underline{M'}(\theta_k, t)/\partial\theta_i \otimes \underline{e}_i\right] \\ &\Rightarrow \quad 2\Delta_{ij} = \frac{1}{|\underline{E}_i|}\frac{1}{|\underline{E}_j|}\left[\frac{\partial \underline{\underline{M'}}}{\partial\theta_j}(\theta_k, t)\frac{\partial \underline{\underline{M'}}}{\partial\theta_i}(\theta_k, t)\right] - \delta_{ij} \end{aligned} \tag{A3.7}$$

which conforms with the definition

$$2\,|\underline{E}_i|\,|\underline{E}_j|\,\Delta_{ij}\mathrm{d}\theta_i\,\mathrm{d}\theta_j = \underline{\mathrm{d}M'}.\underline{\mathrm{d}M'} - \underline{\mathrm{d}M}.\underline{\mathrm{d}M} \tag{A3.8}$$

A3.1.2 Strain Tensor for a Variety (Curve or Surface)

The dimension of the strain tensor depends on the model chosen to represent the material being studied. If this is a curve the tensor has dimension 1 (but order 2), characterised by a single component, and we speak of the deformation of a thread

or cord; if it is a surface we have a symmetric tensor of dimension 2, with three independent components. In calculating the tensor we must note that a variety of dimension p ($p = 1$, 2 or 3) is parametrised by p parameters θ_i: for example, the natural parameter for a curve is its curvilinear abscissa.

Thus the displacement of a point of the variety depends on the parameters θ_i (and is not defined outside the variety). The above basic relations remain valid, provided that the θ_i define an orthogonal coordinate system (see exercises in Volume III).

A3.1.3 Eulerian Tensors for Virtual Strain and Strain Rate

If $\delta\underline{u}(\underline{x}, t)$ and $\underline{v}(\underline{x}, t)$ are the virtual displacement and velocity respectively at a point, the tensors $\delta\underline{\underline{A}}$ and $\dot{\underline{\underline{A}}}$ are *strictly* linear functions of $\delta\underline{u}$ and $\underline{v}$; for a 3-dimensional solid

$$\delta A_{ij} = \frac{1}{2}\left(\frac{\partial \delta u_i}{\partial x_j} + \frac{\partial \delta u_j}{\partial x_i}\right) \qquad \dot{A}_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \tag{A3.9}$$

In other cases, $\delta\underline{\underline{A}}$ (or $\dot{\underline{\underline{A}}}$) can be calculated by replacing $\underline{u}$ by $\delta\underline{u}$ (or $\underline{v}$) in the linearised strain tensor $\underline{\underline{\varepsilon}}$.

A3.1.4 Compatibility Equations

These express the fact that the deformation field is derived from a displacement field. In rectangular coordinates, these are written for small strains, as

$$\in_{ikl} \in_{jmn} \varepsilon_{km,ln} = 0 \tag{A3.10}$$

where $\in_{ijk}$ is zero if any two of i, j, k are equal and $+1$ or -1 according as ijk is an even or an odd permutation of 123. This gives six equations for the $\underline{\underline{\varepsilon}}$ components, of which three are independent.

A3.2 Stresses

A3.2.1 Definitions and Properties

The second-order tensor $\underline{\underline{\sigma}}(M)$, with components $\sigma_{ij}(M)$ is the *Cauchy stress tensor* at the point M in the structure, a Eulerian tensor, defined on the current configuration; it is symmetrical, $\sigma_{ij} = \sigma_{ji}$. If media (1) and (2) are separated by

an interface passing through M the stress vector exerted by (1) on (2) is

$$\underline{T}(M, \underline{n}) = \underline{\underline{\sigma}}(M).\underline{n} \tag{A3.11}$$

where $\underline{n}$ is the unit normal to the interface from (2) to (1).

$\underline{T}$ is traditionally decomposed into a normal component $\sigma = \underline{T}.\underline{n} = \underline{n}.\underline{\underline{\sigma}}.\underline{n}$ and a tangential or *shear* component τ: $\underline{T} = \sigma\underline{n} + \tau\underline{t}$.

In cartesians, taking $\underline{e}_x = \underline{n}$, it is easily established that

σ_{xx} is the normal stress on the plane with normal $\underline{e}_x$

σ_{xy} is the shear stress on this plane, acting in the direction $\underline{e}_y$

If $\sigma_1, \sigma_2, \sigma_3$ are the principal stresses, the set of couples (σ, τ) cover the area in the (σ, τ) plane common to the three circles with centres on the σ-axis and passing through the points with abscissae $\sigma_1, \sigma_2, \sigma_3$ (the Mohr circles).

A3.2.2 Field Equations

The following relations, which express the fundamental principles of mechanics (statics and dynamics), hold always:

– in the static case the *equilibrium equations*

$$\operatorname{div} \underline{\underline{\sigma}} + \rho \underline{f} = 0 \tag{A3.12}$$

where $\rho \underline{f}$ is the body force per unit of mass

– in the dynamic case the *equations of motion*

$$\operatorname{div} \underline{\underline{\sigma}} + \rho(\underline{f} - \underline{\gamma}) = 0 \tag{A3.13}$$

where $\underline{\gamma}$ is the acceleration.

$\underline{\underline{\sigma}}$ also satisfies the boundary conditions:

$$\underline{\underline{\sigma}}(P).\underline{n}(P) = \underline{F}(P) \tag{A3.14}$$

where P is a point in any part of the boundary where a surface force $\underline{F}$ is prescribed. Stress fields $\underline{\underline{\sigma}}$ that satisfy (A3.12) and (A3.14) are said to be *statically admissible* with the body and surface forces $\rho \underline{f}$ and $\underline{F}$.

When the assumption of small perturbations can be made the current configuration can be taken to be the initial configuration for expressing the boundary conditions (A3.14).

In cartesians (A3.13) is written in full as :

$$\begin{cases} \partial\sigma_{xx}/\partial x + \partial\sigma_{xy}/\partial y + \partial\sigma_{xz}/\partial z + \rho(f_x - \gamma_x) = 0 \\ \partial\sigma_{xy}/\partial x + \partial\sigma_{yy}/\partial y + \partial\sigma_{yz}/\partial z + \rho(f_y - \gamma_y) = 0 \\ \partial\sigma_{xz}/\partial x + \partial\sigma_{yz}/\partial y + \partial\sigma_{zz}/\partial z + \rho(f_z - \gamma_z) = 0 \end{cases} \qquad \text{(A3.13 bis)}$$

In orthogonal curvilinear co-ordinates div $\underline{\underline{\sigma}}$ is found by using the relation

$$\nabla \otimes \underline{\underline{\sigma}} = \frac{1}{|\underline{E}_k|}\frac{\partial}{\partial\theta_k}\left(\sigma_{ij}\underline{e}_i \otimes \underline{e}_j\right) \otimes \underline{e}_k \qquad \text{(A3.15)}$$

where k is the index with respect to which the derivation is performed. This gives a third-order tensor, from which the result follows by contracting the two last indices.

A3.3 Problems in Linear Elasticity

We restrict ourselves to quasi-static problems with simple mixed boundary conditions with the assumption of small perturbations.

Data*:* The solid occupies a volume V, with boundary ∂V

Body forces $\rho\,\underline{f}(M)$ (per unit of mass), $M \in V$

Surface forces $\underline{F}(P)$ (per unit of area), $P \in S_T$

Prescribed displacement $\underline{u}^g(P)$, $P \in S_u$

$S_T \cup S_u = \partial V$, $S_T \cap S_u = \varnothing$ (simple mixed boundary conditions)

Unknowns*:* Displacement field $\underline{u}(M)$, $\forall M \in V$ (3 unknowns)

Strain field $\underline{\underline{\varepsilon}}(M)$ (6 unknowns)

Stress field $\underline{\underline{\sigma}}(M)$ (6 unknowns)

Equations*:* (rectangular cartesians)

$$\varepsilon_{ij} = {}^1\!/_2(\partial u_i/\partial x_j + \partial u_j/\partial x_i) : \text{ 6 linear equations} \qquad \text{(A3.6 bis)}$$

$$\sum_j \left(\partial\sigma_{ij}/\partial x_j\right) + \rho f_i = 0 : \text{3 linear equations} \qquad \text{(A3.12 bis)}$$

and the 6 relations for linear homogeneous isotropic elastic behaviour

$$\sigma_{ij} = \lambda\varepsilon_{kk}\delta_{ij} + 2\mu\varepsilon_{ij} \quad \text{or} \quad E\varepsilon_{ij} = (1+\nu)\sigma_{ij} - \nu\sigma_{kk}\delta_{ij} \qquad \text{(A3.16)}$$

where λ, μ are Lamé's coefficients, ν is Poisson's ratio
E is Young's modulus.

Boundary conditions: $\underline{u}(P) = \underline{u}^g$ on S_u, $\sigma_{ij}(P)n_j = F_i$ on S_T

The problem is completely linear: with the assumption of small perturbations the regions S_T and S_u can be defined in the load-free state and remain unchanged; thus we can use the *method of superposition of solutions* and the uniqueness of the solution could be established.

A3.3.1 Navier Equations for Linear Homogeneous Isotropic Elasticity

If we take as principal unknown the displacement field $\underline{u}$ (with 3 components) the above set of equations will be satisfied if $\underline{u}$ satisfies the following vector equation, the Navier or Lamé-Clapeyron[11] equation:

$$\begin{cases} (\lambda + 2\mu)\text{grad}\,\text{div}(\underline{u}) - \mu\text{curl}\,\text{curl}(\underline{u}) + \rho\underline{f} = 0 \\ \text{or equivalently} \\ (\lambda + \mu)\text{grad}\,\text{div}(\underline{u}) + \mu\nabla^2\underline{u} + \rho\underline{f} = 0 \\ \left((\lambda + \mu)u_{j,ij} + \mu u_{i,jj} + \rho f_i = 0\right) \end{cases} \tag{A3.17}$$

This gives 3 scalar equations, to which the boundary conditions have to be added.

A3.3.2 Beltrami Equations

If we take as principal unknown the stress field $\underline{\underline{\sigma}}$ (with 6 components) and we express the compatibility equations (A3.10) in terms of stresses through the constitutive equations (A3.16), taking account of the equilibrium equations (A3.12), we find the Beltrami[12] equations

$$\begin{cases} \nabla^2\underline{\underline{\sigma}} + \dfrac{1}{1+\nu}\text{grad}\,\text{grad}(\text{Tr}\underline{\underline{\sigma}}) + \dfrac{\nu}{1-\nu}\text{div}\,\underline{f}\,\underline{\underline{\delta}} + \text{grad}\,\underline{f} + \text{grad}\,\underline{f}^{\mathrm{T}} = 0 \\ \left(\sigma_{ij,kk} + \dfrac{1}{1+\nu}\sigma_{kk,ij} + \dfrac{\nu}{1-\nu}f_{k,k}\delta_{ij} + f_{i,j} + f_{j,i} = 0\right) \end{cases} \tag{A3.18}$$

Note that these six equations are not fully independent; they have to be used together with the equilibrium equations (and the boundary equations) for the problem to be solved.

[11] Benoît Paul Emile Clapeyron (1799–1864) was a French engineer and physicist.

[12] Eugenio Beltrami (1835–1900) was an Italian mathematician.

More detailed developments, especially at finite strain, can be found *e.g.* in Salençon (2001).

Reference

Salençon J (2001) Handbook of continuum mechanics. Springer, Berlin

Appendix D
Tables

Table D1 Structural parameters of various elements and compounds

Compounds	Structure	Lattice parameters (10^{-10} m)	Interatomic distance (10^{-10} m)
Metals	–	–	–
Li	BCC	3.5093 at 78 K	3.039
	FCC	4.404	3.114
	CPH	3.111 at 78 K	3.111
		5.093	3.116
Be	CPH	2.2858	2.2858
		3.5843	2.226
Mg	CPH	3.2093	3.2093
		5.2103	3.197
Ti alpha	CPH	2.95	2.95
		4.686	2.897
Ti beta	BCC	3.3065 at 900°C	2.863
Zr alpha	CPH	3.232	3.232
		5.147	3.179
Zr beta	BCC	3.62 at 900°C	3.135
Hf	CPH	3.1967	3.1967
		5.0578	3.131
V	BCC	3.024	2.619
Nb	BCC	3.3004	2.601
Ta	BCC	3.3058	2.649
Cr	BCC	2.8839	2.497
Mo	BCC	3.1473	2.726
W	BCC	3.16469	2.741
Fe alpha	BCC	2.8665	2.482

(continued)

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

Table D1 (continued)

Compounds	Structure	Lattice parameters (10^{-10} m)	Interatomic distance (10^{-10} m)
Fe gamma	FCC	3.642 at 950°C	2.575
Co alpha	CPH	2.5071	2.5071
		4.0686	2.497
Co beta	FCC	3.5442 at T > 450°C	2.506
Ni	FCC	3.52387	2.491
Rh	FCC	3.8031	2.689
Pd	FCC	3.8898	2.750
Ir	FCC	3.8394	2.714
Pt	FCC	3.9231	2.774
Cu	FCC	3.61496	2.556
Ag	FCC	4.0862	2.889
Au	FCC	4.07825	2.883
Zn	CPH	2.6648	2.6648
		4.9467	2.913
Cd	CPH	2.97887	2.97887
		5.61765	3.293
Al	FCC	4.04958	
Sn alpha gray tin	Cubic diamond	6.4912	2.810
Beta white tin	Tetragonal	5.8197	3.022
		3.17488	
Pb	FCC	4.9505	3.500
U alpha	Orthorhombic	2.854	2.754
		5.869	
		4.955	
Covalents	–	–	
C diamond	C diamond	3.56679	1.544
C graphite	Hex	2.456	1.421
		6.696	
Si	C diamond	5.4307	2.351
Ge	C diamond	5.65735	2.450
Carbides and nitrides	–	–	
SiC	Cubic ZnS	4.348	3.074
	Hex wurtzite	3.076	3.086
		5.048	
TiC	FCC	4.3186	3.053
WC	Hex (P6mm)	2.9065	2.9065
		2.8366	2.8366
Fe_3C	Orthorhombic	5.08493	
BN h or alpha	Hexagonal		
Beta	Diamond	3.615	2.556

(continued)

Table D1 (continued)

Compounds	Structure	Lattice parameters (10^{-10} m)	Interatomic distance (10^{-10} m)
Oxides	–	–	
SiO_2 alpha quartz	Rhombohedral	7.12	
Beta quartz	Hex	4.913	2.635
		5.404	2.660
Al_2O_3 alpha	Rhombohedral	4.754	3.97
		12.982	
CaO	FCC	4.815	3.404
MgO	FCC	4.216	2.980
ZrO_2	FCC	5.1291	3.626
	Monoclinic	5.169	
		5.232	
		5.341	
TiO_2	Tetragonal	9.174	
		5.449	
		5.138	

Table D2 Molecular weight and density of various elements and compounds

Compounds	Molecular weight	Density (kg m^{-3})
Metals	–	–
Li	6.941	534
Be	9.012	1,848
Mg	24.3050	1,738
Ti	47.867	4,510
Zr	91.224	6,520
Hf	178.49	13,310
V	50.9415	6,100
Nb	92.90638	8,570
Ta	180.9479	16,400
Cr	51.9961	7,150
Mo	95.94	10,220
W	183.84	19,300
Mn	54.938	7,300
Fe	55.845	7,874
Co	58.933200	8,900
Ni	58.6934	8,902
Rh	102.90550	12,410
Pd	106.42	12,020
Ir	192.217	22,562
Pt	195.078	21,450
Cu	63.546	8,960
Ag	107.8682	10,500

(continued)

Table D2 (continued)

Compounds	Molecular weight	Density (kg m^{-3})
Au	196.96655	19,300
Zn	65.409	7,134
Cd	112.411	8,690
Al	26.981538	26,98.9
Sn	118.710	7,290
Pb	207.2	11,350
U	238.028	19,100
Covalents	–	–
C diamond	12.0107	3,508
Graphite		2,250
Si	18.09	2,330
Ge	72.61	5,323
Carbides and nitrides	–	–
SiC	40.07	3,217
TiC	59.91	4,940
WC	195.87	15,630
Fe_3C	179.5457	
BN h or alpha	24.818	2,250
C or beta		2,300
Oxides	–	–
SiO_2	50.09	2,650
Al_2O_3	149.828	3,950–4,100
CaO	56.077	3,300–3,400
MgO	40.304	3,580
ZrO_2	123.2228	5,700
TiO_2	79.8658	4,224

Table D3 Thermodynamic properties of various elements and compounds

Compounds	Enthalpy of formation (kJ/mol at 298 K)	Fusion temperature (°C)	Enthalpy of fusion (kJ/mol)
Metals	–	–	–
Li	–	180.54	4.60
Be	–	648.85	9.80
Mg	–	1277.85	9.04
Ti alpha	0		
Beta	473	1659.85	20.90
Zr	–	1851.85	23.00
Hf	–	2229.85	25.50
V	–	1886.85	17.60
Nb	–	1467.85	27.20
Ta	–	2995.85	31.40
Cr	–	1856.85	15.30
Mo	–	2616.85	27.70
W	–	3406.85	35.20

(continued)

Table D3 (continued)

Compounds	Enthalpy of formation (kJ/mol at 298°K)	Fusion temperature (°C)	Enthalpy of fusion (kJ/mol)
Mn	–	1243.85	14.40
Fe	–	1534.85	14.90
Co	–	1494.85	15.20
Ni	–	1452.85	17.60
Rh	–	1965.85	21.55
Pd	–	1551.85	17.20
Ir	–	2409.85	26.40
Pt	–	1771.85	19.70
Cu	–	1083.45	13.00
Ag	–	961.93	11.30
Au	–	1064.43	12.70
Zn	130.4	419.58	6.67
Cd	–	320.95	6.11
Al	–	660.37	10.67
Sn gray	301.2	331.97	7.20
White	0		
Pb	–	327.50	5.12
U	–	1132.35	15.50
Covalents	–	–	–
C diamond	1.92	3546.85	105.0
Graphite	0		
Si	–	1409.85	39.60
Ge	–	937.45	34.70
Carbides and nitrides	–	–	–
SiC	−71.55	2,730	–
TiC	−184.10	3,160	–
WC	–	2,870	–
Fe_3C	–	1,227	
BN	476.98	2,973	–
Oxides	–	–	–
SiO_2	−910.86	1,610	–
Al_2O_3	−1675.7	2,072	–
CaO	−635.09	2,570	–
MgO	−601.24	2,800	–
ZrO_2	−1,080	2,683	–
TiO_2	−938.72	1,853	–

Table D4 Thermal properties of various elements and compounds

Compounds	Specific heat (J/kg K)	Thermal conductivity (W/m K)	Coefficient of thermal expansion (10^{-6}/K)
Metals	–	–	–
Li	3,600	84.70	46
Be	1,820	200	11.5
Mg	1,020	156	26
Ti alpha	520	21.9	8.5
Ti beta	–		
Zr alpha	270	22.7	5.7
Zr beta	–		
Hf	140	23	–
V	490	30.7	8
Nb	260	53.7	7.2
Ta	140	57.5	6.5
Cr	450	93.7	6.2
Mo	250	138	4.9
W	130	174	4.5
Fe alpha	440	80.2	11.8
Fe gamma			
Co alpha	420	100	12.3
Co beta			
Ni	440	90.7	13
Rh	240	150	8
Pd	240	71.8	11.8
Ir	130	147	6.4
Pt	130	71.6	9
Cu	380	401	16.4
Ag	240	429	18
Au	130	317	14.4
Zn	390	116	29.7
Cd	230	96.8	30
Al	900	237	69
Sn alpha gray tin	230	66.6	23.4
Beta white tin			
Pb	130	35.3	29
U alpha	120	27.6	13.9
Covalents	–	–	
C diamond	710	990	1.2
C graphite			7.9
Si	710	148	5.1
Ge	320	59.9	6.1

(continued)

Table D4 (continued)

Compounds	Specific heat (J/kg K)	Thermal conductivity (W/m K)	Coefficient of thermal expansion (10^{-6}/ K)
Carbides and nitrides	–	–	–
SiC	1,130	18.5 (at 400°C)	4.8
TiC	560	–	–
WC	210	–	–
Fe_3C	-	–	–
BN h or alpha	1,610	20 ∥; 27⊥	2.7 ∥; 38 ⊥
Beta	1,470		
Oxides	–	–	
SiO_2 alpha quartz	670–750	–	0.77–1.4
Beta quartz			
Al_2O_3 alpha	753–850	–	5.4
CaO	750	–	–
MgO	920	–	–
ZrO_2	400–610	–	–
TiO_2	690	–	–

Appendix E
Physical Constants; Conversion Factors

$\mathbf{N_A}$	Avogadro's number	**6.022** 141 79·$\mathbf{10^{23}}$	Number of atoms in 12 g of ^{12}C
		$\pm$ 3.0·10^{16} $\mathbf{mol^{-1}}$	
u	Atomic mass unit	**1.660** 538 782(83) $\mathbf{10^{-27}}$ **kg**	1/12 the mass of ^{12}C
h	Planck's constant	**6.626** 068 96·$\mathbf{10^{-34}}$	The energy of a photon is the product of its frequency by the Planck constant
		$\pm$ 0.000 000 33 10^{-34} **J·s**	
$\hbar$	-	**1.054** 571 628·$\mathbf{10^{-34}}$	h/2π
		$\pm$ 0.000 000 053·10^{-34} **J·s**	
-	Mass of the electron	**510.998** 918 **keV·c**$^{\mathbf{-2}}$	–
		(**9.109** 382·10^{-31} **kg**)	
e	Charge on the electron	**−1.602** 176 53·$\mathbf{10^{-19}}$ **°C**	–
F	Faraday's constant	**96 485**.3399° **C·mol**$^{\mathbf{-1}}$	$N_A e$
$\boldsymbol{\lambda_0}$	Wave length of X-ray corresponding to 1 eV	**12 398**.04 $\mathbf{10^{-10}}$	–
		$\pm$ 0.12 10^{-10} **eV·m**	
c	Velocity of light	299 792 458 m·s^{-1}	–
		300 000 km·s$^{\mathbf{-1}}$	
k	Boltzmann's constant	**1.380** 650 4(24) $\mathbf{10^{-23}}$ **J·K**$^{\mathbf{-1}}$	–
		8.617 343(15) $\mathbf{10^{-5}}$ **eV·K**$^{\mathbf{-1}}$	
R	Gas constant	**8.314** 472(15) **J·K**$^{\mathbf{-1}}$**·mol**$^{\mathbf{-1}}$	$R = N_A k$
		1.985 8775(34) **cal·K**$^{\mathbf{-\&}}$**·mol**$^{\mathbf{-1}}$	
-	Melting point of ice	**273**.16 **K**	–

(continued)

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180, DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

(continued)

-	J (joule)	eV (electron-volt)	cal (calorie)
1 J =	–	**6.241** 506 48·**10**^{18}	**0.238** 845 896 63
1 **eV** =	**1.602** 177 3·**10**$^{-19}$	–	**3.826** 734 377·**10**$^{-20}$
1 **cal** =	**4.186** 8	**26.131** 939 331·**10**^{18}	–

1 J = 10^7 erg
1 J = 0.737562 ft lb = 0.000948 BTU
1 ft lb = 1.355818 J 1 BTU = 1,054 J

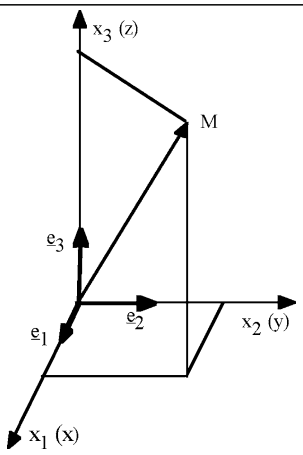
–	MPa (mega pascal)	bar	atm (atmosphere)	psi (pound per square inch)	mm Hg
1 **MPa** =	–	10	9.869	145	7,500
1 **bar** =	0.1	-	0.987	14.5	750
1 **atm** =	0.101	1.013	-	14.7	760
1 **psi** =	6.89 10^{-3}	0.069	0.068	–	51.71
1 **mmHg** =	133 10^{-6}	1.333 10^{-3}	1.316 10^{-3}	19.34 10^{-3}	–

1 MPa = 10^6 N·m^{-2}
1 m = 39.37 in.
1 in. = 0.0254 m

Appendix F
Coordinate Systems

The following tables give the expressions for the various quantities and functions used in the text, in the three main coordinate systems: rectangular cartesians, cylindrical polars and spherical polars respectively.

F.1 Rectangular Cartesian Coordinates

Definitions	 $\underline{e}_i = \dfrac{\partial \underline{M}}{\partial x_i}$
Coordinates	$\underline{OM} = x_i \underline{e}_i$
Vector field	$\underline{u}(M) = u_i(M)\underline{e}_i$

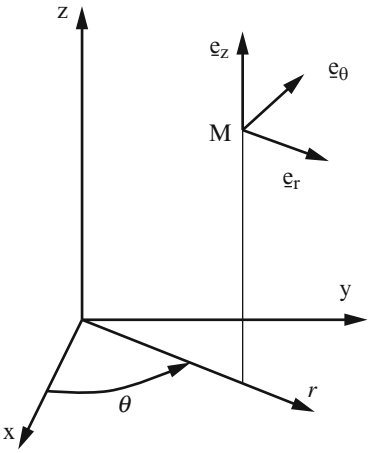
(continued)

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

Unit base vectors	$\mathrm{d}\underline{e}_i = 0$
	$\mathrm{d}\underline{M} = \mathrm{d}x_i\,\underline{e}_i$
Gradient of a scalar f $\mathrm{d}f = \underline{\underline{\mathrm{grad}}}\ f \cdot \mathrm{d}\underline{M}$	$\underline{\mathrm{grad}}\, f = \dfrac{\partial f}{\partial x_i}\underline{e}_i = f_{,i}\,\underline{e}_i$
Gradient of a vector field $\underline{u}$ $\mathrm{d}\underline{u} = (\underline{\underline{\mathrm{grad}}}\,\underline{u}).\mathrm{d}\underline{M}$	$\underline{\underline{\mathrm{grad}}}\,\underline{u} = \dfrac{\partial u_i}{\partial x_j}\underline{e}_i \otimes \underline{e}_j = u_{i,j}\,\underline{e}_i \otimes \underline{e}_j$
Deformation field $= (1/2)(\underline{\underline{\mathrm{grad}}}\,\underline{u} + \underline{\underline{\mathrm{grad}}}^{\mathrm{T}}\underline{u})$	$\underline{\underline{\varepsilon}} = \varepsilon_{ij}\underline{e}_i \otimes \underline{e}_j \quad \varepsilon_{ij} = \dfrac{1}{2}\left(u_{i,j} + u_{j,i}\right)$
Divergence of a vector field $\mathrm{div}\,\underline{u} = \mathrm{Tr}\left(\underline{\underline{\mathrm{grad}}}\,\underline{u}\right)$	$\mathrm{div}\,\underline{u} = u_{k,k}$
Laplacian of a scalar $\Delta f = \mathrm{div}\left(\underline{\mathrm{grad}}\, f\right)$	$\Delta f = \sum_i \dfrac{\partial^2 f}{\partial {x_i}^2} = f_{,ii}$
Laplacian of a vector field $\underline{\Delta u} = \underline{\mathrm{div}}\left(\underline{\underline{\mathrm{grad}}}\underline{u}\right)$	$\underline{\Delta u} = u_{i\,,kk}\,\underline{e}_i$
Second-order tensor field $\underline{\underline{T}}$	$\underline{\underline{T}} = T_{ij}\underline{e}_i \otimes \underline{e}_j$
Divergence of a symmetric second-order tensor field $\underline{\mathrm{div}}\,\underline{\underline{\sigma}}$	$\underline{\mathrm{div}}\,\underline{\underline{\sigma}} = \sigma_{ij,j}\,\underline{e}_i$
Change of coordinates for a vector $\underline{u}$	$\underline{u} = u_i\,\underline{e}_j$
Change of coordinates for a symmetric second-order tensor	$\underline{\underline{\varepsilon}} = \varepsilon_{ij}\underline{e}_i \otimes \underline{e}_j$

F.2 Cylindrical Polars

Definitions



$$\underline{e}_{\mathrm{r}} = \frac{\partial \underline{\underline{M}}}{\partial r} \quad \underline{e}_{\theta} = \frac{1}{r}\frac{\partial \underline{\underline{M}}}{\partial \theta} \quad \underline{e}_{\mathrm{z}} = \frac{\partial \underline{\underline{M}}}{\partial z}$$

(continued)

Coordinates
$\underline{OM} = r\underline{e}_{\mathrm{r}} + z\underline{e}_{\mathrm{z}}$

Vector field $\underline{u}(M)$
$\underline{u}(M) = u_{\mathrm{r}}\underline{e}_{\mathrm{r}} + u_{\theta}\underline{e}_{\theta} + u_{\mathrm{z}}\underline{e}_{\mathrm{z}}$

Unit base vectors
$$\mathrm{d}\underline{e}_{\mathrm{r}} = \mathrm{d}\theta\,\underline{e}_{\theta} \quad \mathrm{d}\underline{e}_{\theta} = -\mathrm{d}\theta\,\underline{e}_{\mathrm{r}} \quad \mathrm{d}\underline{e}_{\mathrm{z}} = 0$$

$$\mathrm{d}\underline{M} = \mathrm{d}r\,\underline{e}_{\mathrm{r}} + r\mathrm{d}\theta\,\underline{e}_{\theta} + \mathrm{d}z\,\underline{e}_{\mathrm{z}}$$

Gradient of a scalar f
$\mathrm{d}f = \underline{\underline{\mathrm{grad}}}\, f \cdot \mathrm{d}\underline{M}$
$$\underline{\underline{\mathrm{grad}}}\, f = \frac{\partial f}{\partial r}\underline{e}_{\mathrm{r}} + \frac{1}{r}\frac{\partial f}{\partial \theta}\underline{e}_{\theta} + \frac{\partial f}{\partial z}\underline{e}_{\mathrm{z}}$$

Gradient of a vector field $\underline{u}$
$\mathrm{d}\underline{u} = (\underline{\underline{\mathrm{grad}}}\, \underline{u}).\mathrm{d}\underline{M}$
$$\underline{\underline{\mathrm{grad}}}\, \underline{u} = \begin{bmatrix} \dfrac{\partial u_{\mathrm{r}}}{\partial r} & \dfrac{1}{r}\left(\dfrac{\partial u_{\mathrm{r}}}{\partial \theta} - u_{\theta}\right) & \dfrac{\partial u_{\mathrm{r}}}{\partial z} \\ \dfrac{\partial u_{\theta}}{\partial r} & \dfrac{1}{r}\left(\dfrac{\partial u_{\theta}}{\partial \theta} + u_{\mathrm{r}}\right) & \dfrac{\partial u_{\theta}}{\partial z} \\ \dfrac{\partial u_{\mathrm{z}}}{\partial r} & \dfrac{1}{r}\dfrac{\partial u_{\mathrm{z}}}{\partial \theta} & \dfrac{\partial u_{\mathrm{z}}}{\partial z} \end{bmatrix}$$

Deformation field
$\underline{\underline{\varepsilon}} = (1/2)\,(\underline{\underline{\mathrm{grad}}}\, \underline{u} + \underline{\underline{\mathrm{grad}}}^{\mathrm{T}} \underline{u})$
$$\underline{\underline{\varepsilon}} = \begin{bmatrix} \dfrac{\partial u_{\mathrm{r}}}{\partial r} & \dfrac{1}{2}\dfrac{\partial u_{\theta}}{\partial r} + \dfrac{1}{2r}\left(\dfrac{\partial u_{\mathrm{r}}}{\partial \theta} - u_{\theta}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_{\mathrm{r}}}{\partial z} + \dfrac{\partial u_{\mathrm{z}}}{\partial r}\right) \\ & \dfrac{1}{r}\left(\dfrac{\partial u_{\theta}}{\partial \theta} + u_{\mathrm{r}}\right) & \dfrac{1}{2}\left(\dfrac{\partial u_{\theta}}{\partial z} + \dfrac{1}{r}\dfrac{\partial u_{\mathrm{z}}}{\partial \theta}\right) \\ \mathrm{sym} & & \dfrac{\partial u_{\mathrm{z}}}{\partial z} \end{bmatrix}$$

Divergence of a vector field
$\mathrm{div}\,\underline{u} = \mathrm{Tr}(\underline{\underline{\mathrm{grad}}}\, \underline{u})$
$$\mathrm{div}\,\underline{u} = \frac{\partial u_{\mathrm{r}}}{\partial r} + \frac{1}{r}\frac{\partial u_{\theta}}{\partial \theta} + \frac{u_{\mathrm{r}}}{r} + \frac{\partial u_{\mathrm{z}}}{\partial z}$$

Laplacian of a scalar
$\Delta f = \mathrm{div}\,(\underline{\underline{\mathrm{grad}}}\, f)$
$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}$$

Laplacian of a vector field
$\underline{\Delta}\,\underline{u} = \underline{\mathrm{div}}(\underline{\underline{\mathrm{grad}}}\, \underline{u})$
$$\underline{\Delta u} = \left(\Delta u_{\mathrm{r}} - \frac{2}{r^2}\frac{\partial u_{\theta}}{\partial \theta} - \frac{u_{\mathrm{r}}}{r^2}\right)\underline{e}_{\mathrm{r}} + \left(\Delta u_{\theta} + \frac{2}{r^2}\frac{\partial u_{\mathrm{r}}}{\partial \theta} - \frac{u_{\theta}}{r^2}\right)\underline{e}_{\theta} + \Delta u_{\mathrm{z}}\underline{e}_{\mathrm{z}}$$

Second-order tensor field $\underline{\underline{T}}$
$$\underline{\underline{T}} = T_{\mathrm{rr}}\underline{e}_{\mathrm{r}} \otimes \underline{e}_{\mathrm{r}} + T_{\mathrm{r}\theta}\underline{e}_{\mathrm{r}} \otimes \underline{e}_{\theta} + T_{\mathrm{rz}}\underline{e}_{\mathrm{r}} \otimes \underline{e}_{\mathrm{z}} + T_{\theta\mathrm{r}}\underline{e}_{\theta} \otimes \underline{e}_{\mathrm{r}} + T_{\theta\theta}\underline{e}_{\theta} \otimes \underline{e}_{\theta} + T_{\theta\mathrm{z}}\underline{e}_{\theta} \otimes \underline{e}_{\mathrm{z}} + T_{\mathrm{zr}}\underline{e}_{\mathrm{z}} \otimes \underline{e}_{\mathrm{r}} + T_{\mathrm{z}\theta}\underline{e}_{\mathrm{z}} \otimes \underline{e}_{\theta} + T_{\mathrm{zz}}\underline{e}_{\mathrm{z}} \otimes \underline{e}_{\mathrm{z}}$$

Divergence of a symmetric second-order tensor field $\underline{\mathrm{div}}\,\underline{\underline{\sigma}}$
$$\underline{\mathrm{div}\underline{\sigma}} = \left(\frac{\partial \sigma_{\mathrm{rr}}}{\partial r} + \frac{1}{r}\frac{\partial \sigma_{\mathrm{r}\theta}}{\partial \theta} + \frac{\partial \sigma_{\mathrm{rz}}}{\partial z} + \frac{\sigma_{\mathrm{rr}} - \sigma_{\theta\theta}}{\mathrm{r}}\right)\underline{e}_{\mathrm{r}} + \left(\frac{\partial \sigma_{\mathrm{r}\theta}}{\partial r} + \frac{1}{r}\frac{\partial \sigma_{\theta\theta}}{\partial \theta} + \frac{\partial \sigma_{\mathrm{z}\theta}}{\partial z} + 2\frac{\sigma_{\mathrm{r}\theta}}{r}\right)\underline{e}_{\theta} + \left(\frac{\partial \sigma_{\mathrm{zr}}}{\partial r} + \frac{1}{r}\frac{\partial \sigma_{\mathrm{z}\theta}}{\partial \theta} + \frac{\partial \sigma_{\mathrm{zz}}}{\partial z} + \frac{\sigma_{\mathrm{zr}}}{r}\right)\underline{e}_{\mathrm{z}}$$

Change of coordinates for the base vectors
$$\underline{e}_{\mathrm{r}} = \cos\theta\,\underline{e}_1 + \sin\theta\,\underline{e}_2$$
$$\underline{e}_{\theta} = -\sin\theta\,\underline{e}_1 + \cos\theta\,\underline{e}_2$$
$$\underline{e}_{\mathrm{z}} = \underline{e}_3$$

(continued)

$$\underline{e}_1 = \cos\theta\,\underline{e}_r - \sin\theta\,\underline{e}_\theta$$
$$\underline{e}_2 = \sin\theta\,\underline{e}_r + \cos\theta\,\underline{e}_\theta$$
$$\underline{e}_3 = \underline{e}_z$$

Change of coordinates for a vector $\underline{u}$

$$u_1 = u_r\cos\theta - u_\theta\sin\theta$$
$$u_2 = u_r\sin\theta + u_\theta\cos\theta$$
$$u_3 = u_z$$

$$u_r = u_1\cos\theta + u_2\sin\theta$$
$$u_\theta = -u_1\sin\theta + u_2\cos\theta$$
$$u_z = u_3$$

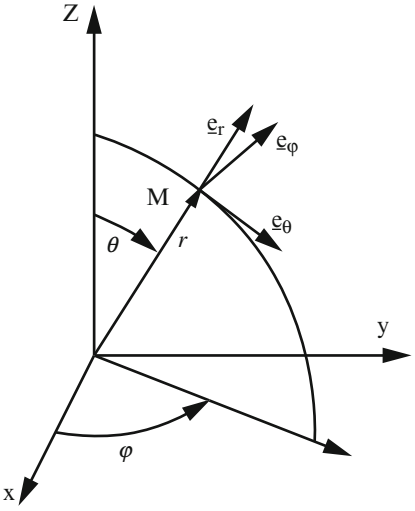
Change of coordinates for a symmetric second-order tensor $\underline{\underline{\varepsilon}}$

$$\varepsilon_{11} = \varepsilon_{rr}\cos^2\theta + \varepsilon_{\theta\theta}\sin^2\theta - 2\varepsilon_{r\theta}\sin\theta\cos\theta$$
$$\varepsilon_{22} = \varepsilon_{rr}\sin^2\theta + \varepsilon_{\theta\theta}\cos^2\theta + 2\varepsilon_{r\theta}\sin\theta\cos\theta$$
$$\varepsilon_{33} = \varepsilon_{zz}$$
$$\varepsilon_{12} = (\varepsilon_{rr} - \varepsilon_{\theta\theta})\sin\theta\cos\theta + \varepsilon_{r\theta}(\cos^2\theta - \sin^2\theta)$$
$$\varepsilon_{13} = \varepsilon_{rz}\cos\theta - \varepsilon_{\theta z}\sin\theta$$
$$\varepsilon_{23} = \varepsilon_{rz}\sin\theta + \varepsilon_{\theta z}\cos\theta$$

$$\varepsilon_{rr} = \varepsilon_{11}\cos^2\theta + \varepsilon_{22}\sin^2\theta + 2\varepsilon_{12}\sin\theta\cos\theta$$
$$\varepsilon_{\theta\theta} = \varepsilon_{11}\sin^2\theta + \varepsilon_{22}\cos^2\theta - 2\varepsilon_{12}\sin\theta\cos\theta$$
$$\varepsilon_{zz} = \varepsilon_{33}$$
$$\varepsilon_{r\theta} = (\varepsilon_{22} - \varepsilon_{11})\sin\theta\cos\theta + \varepsilon_{12}(\cos^2\theta - \sin^2\theta)$$
$$\varepsilon_{r\theta} = \varepsilon_{13}\cos\theta + \varepsilon_{23}\sin\theta$$
$$\varepsilon_{\theta z} = -\varepsilon_{13}\sin\theta + \varepsilon_{23}\cos\theta$$

F.3 Spherical Polars

Definitions



$$\underline{e}_r = \frac{\partial \underline{\underline{M}}}{\partial r} \quad \underline{e}_\theta = \frac{1}{r}\frac{\partial \underline{\underline{M}}}{\partial \theta} \quad \underline{e}_\varphi = \frac{1}{r\sin\theta}\frac{\partial \underline{\underline{M}}}{\partial \varphi}$$

(continued)

Coordinates

$$\underline{OM} = r\underline{e}_{\mathrm{r}}$$

Vector field $\underline{u}(M)$

$$\underline{u}(M) = u_{\mathrm{r}}\underline{e}_{\mathrm{r}} + u_{\theta}\underline{e}_{\theta} + u_{\varphi}\underline{e}_{\varphi}$$

Unit base vectors

$$\begin{aligned}
\mathrm{d}\underline{e}_{\mathrm{r}} &= \mathrm{d}\theta\ \underline{e}_{\theta} + \sin\theta\ \mathrm{d}\varphi\ \underline{e}_{\varphi}\\
\mathrm{d}\underline{e}_{\theta} &= -\mathrm{d}\theta\ \underline{e}_{\mathrm{r}} + \cos\theta\ \mathrm{d}\varphi\ \underline{e}_{\varphi}\\
\mathrm{d}\underline{e}_{\varphi} &= -\sin\theta\ \mathrm{d}\varphi\ \underline{e}_{\mathrm{r}} - \cos\theta\ \mathrm{d}\varphi\ \underline{e}_{\theta}\\
\mathrm{d}\underline{M} &= \mathrm{d}r\ \underline{e}_{\mathrm{r}} + r\mathrm{d}\theta\ \underline{e}_{\theta} + r\sin\theta\ \mathrm{d}\varphi\ \underline{e}_{\varphi}
\end{aligned}$$

Gradient of a scalar f $\mathrm{d}f = \underline{\mathrm{grad}}\, f \cdot \mathrm{d}\underline{M}$

$$\underline{\mathrm{grad}}\, f = \frac{\partial f}{\partial r}\underline{e}_{\mathrm{r}} + \frac{1}{r}\frac{\partial f}{\partial \theta}\underline{e}_{\theta} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \varphi}\underline{e}_{\varphi}$$

Gradient of a vector field $\underline{u}$ $\mathrm{d}\underline{u} = (\underline{\underline{\mathrm{grad}}}\ \underline{u}).\mathrm{d}\underline{M}$

$$\underline{\underline{\mathrm{grad}}}\,\underline{u} = \begin{bmatrix}
\dfrac{\partial u_{\mathrm{r}}}{\partial r} & \dfrac{1}{r}\left(\dfrac{\partial u_{\mathrm{r}}}{\partial \theta} - u_{\theta}\right) & \dfrac{1}{r}\left(\dfrac{1}{\sin\theta}\dfrac{\partial u_{\mathrm{r}}}{\partial \varphi} - u_{\varphi}\right)\\
\dfrac{\partial u_{\theta}}{\partial r} & \dfrac{1}{r}\left(\dfrac{\partial u_{\theta}}{\partial \theta} + u_{\mathrm{r}}\right) & \dfrac{1}{r}\left(\dfrac{1}{\sin\theta}\dfrac{\partial u_{\theta}}{\partial \varphi} - \dfrac{u_{\varphi}}{\tan\theta}\right)\\
\dfrac{\partial u_{\varphi}}{\partial r} & \dfrac{1}{r}\dfrac{\partial u_{\varphi}}{\partial \theta} & \dfrac{1}{r}\left(\dfrac{1}{\sin\theta}\dfrac{\partial u_{\varphi}}{\partial \varphi} + \dfrac{u_{\theta}}{\tan\theta} + u_{\mathrm{r}}\right)
\end{bmatrix}$$

Deformation field $\underline{\underline{\varepsilon}} = (1/2)$ $(\underline{\underline{\mathrm{grad}}}\ \underline{u} + \underline{\underline{\mathrm{grad}}}^{\mathrm{T}}\underline{u})$

$$\underline{\underline{\varepsilon}} = \begin{bmatrix}
\dfrac{\partial u_{\mathrm{r}}}{\partial r} & \dfrac{1}{2}\dfrac{\partial u_{\theta}}{\partial r} + \dfrac{1}{2r}\left(\dfrac{\partial u_{\mathrm{r}}}{\partial \theta} - u_{\theta}\right) & \dfrac{1}{2}\dfrac{\partial u_{\varphi}}{\partial r} + \dfrac{1}{2r}\left(\dfrac{1}{\sin\theta}\dfrac{\partial u_{\mathrm{r}}}{\partial \varphi} - u_{\varphi}\right)\\
 & \dfrac{1}{r}\left(\dfrac{\partial u_{\theta}}{\partial \theta} + u_{\mathrm{r}}\right) & \dfrac{1}{2r}\dfrac{\partial u_{\varphi}}{\partial \theta} + \dfrac{1}{2r}\left(\dfrac{1}{\sin\theta}\dfrac{\partial u_{\theta}}{\partial \varphi} - \dfrac{u_{\varphi}}{\tan\theta}\right)\\
\mathrm{sym} & & \dfrac{1}{r}\left(\dfrac{1}{\sin\theta}\dfrac{\partial u_{\varphi}}{\partial \varphi} + \dfrac{u_{\theta}}{\tan\theta} + u_{\mathrm{r}}\right)
\end{bmatrix}$$

Divergence of a vector field $\mathrm{div}\,\underline{u} = \mathrm{Tr}\left(\underline{\underline{\mathrm{grad}}}\ \underline{u}\right)$

$$\mathrm{div}\underline{u} = \frac{\partial u_{\mathrm{r}}}{\partial r} + \frac{1}{r}\frac{\partial u_{\theta}}{\partial \theta} + 2\frac{u_{\mathrm{r}}}{r} + \frac{1}{r\sin\theta}\frac{\partial u_{\varphi}}{\partial \varphi} + \frac{u_{\theta}}{r\tan\theta}$$

Laplacian of a scalar $\Delta f = \mathrm{div}\left(\underline{\mathrm{grad}}\, f\right)$

$$\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{1}{r^2\tan\theta}\frac{\partial f}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \varphi^2}$$

Laplacian of a vector field $\underline{\Delta u} = \underline{\mathrm{div}}(\underline{\underline{\mathrm{grad}}}\ \underline{u})$

$$\begin{aligned}
\underline{\Delta u} = &\left(\Delta u_{\mathrm{r}} - \frac{2u_{\mathrm{r}}}{r^2} - \frac{2}{r^2\sin\theta}\frac{\partial(u_{\theta}\sin\theta)}{\partial \theta} - \frac{2}{r^2\sin\theta}\frac{\partial u_{\varphi}}{\partial \varphi}\right)\underline{e}_{\mathrm{r}}\\
&+ \left(\Delta u_{\theta} + \frac{2}{r^2}\frac{\partial u_{\mathrm{r}}}{\partial \theta} - \frac{u_{\theta}}{r^2\sin^2\theta} - \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial u_{\varphi}}{\partial \varphi}\right)\underline{e}_{\theta}\\
&+ \left(\Delta u_{\varphi} + \frac{2}{r^2\sin^2\theta}\frac{\partial u_{\mathrm{r}}}{\partial \varphi} - \frac{u_{\varphi}}{r^2\sin^2\theta} + \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial u_{\theta}}{\partial \varphi}\right)\underline{e}_{\varphi}
\end{aligned}$$

Second-order tensor field $\underline{\underline{T}}$

$$\begin{aligned}
\underline{\underline{T}} = &T_{\mathrm{rr}}\underline{e}_{\mathrm{r}} \otimes \underline{e}_{\mathrm{r}} + T_{\mathrm{r}\theta}\underline{e}_{\mathrm{r}} \otimes \underline{e}_{\theta} + T_{\mathrm{r}\varphi}\underline{e}_{\mathrm{r}} \otimes \underline{e}_{\varphi}\\
&+ T_{\theta\mathrm{r}}\underline{e}_{\theta} \otimes \underline{e}_{\mathrm{r}} + T_{\theta\theta}\underline{e}_{\theta} \otimes \underline{e}_{\theta} + T_{\theta\varphi}\underline{e}_{\theta} \otimes \underline{e}_{\varphi}\\
&+ T_{\varphi\mathrm{r}}\underline{e}_{\varphi} \otimes \underline{e}_{\mathrm{r}} + T_{\varphi\theta}\underline{e}_{\varphi} \otimes \underline{e}_{\theta} + T_{\varphi\varphi}\underline{e}_{\varphi} \otimes \underline{e}_{\varphi}
\end{aligned}$$

(continued)

Divergence of a symmetric second-order tensor field $\underline{\underline{\mathrm{div}}}\,\underline{\underline{\sigma}}$

$$\begin{aligned}\underline{\underline{\mathrm{div}}}\,\underline{\underline{\sigma}} = &\left(\frac{\partial\sigma_{rr}}{\partial r}+\frac{1}{r}\frac{\partial\sigma_{r\theta}}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial\sigma_{r\varphi}}{\partial\varphi}+\frac{1}{r}(2\sigma_{rr}-\sigma_{\theta\theta}-\sigma_{\varphi\varphi}+\sigma_{r\theta}\cot\theta)\right)\underline{e}_r\\ &+\left(\frac{\partial\sigma_{\theta r}}{\partial r}+\frac{1}{r}\frac{\partial\sigma_{\theta\theta}}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial\sigma_{\theta\varphi}}{\partial\varphi}+\frac{1}{r}(\sigma_{\theta\theta}\cot\theta-\sigma_{\varphi\varphi}\cot\theta+3\sigma_{r\theta})\right)\underline{e}_\theta\\ &+\left(\frac{\partial\sigma_{\varphi r}}{\partial r}+\frac{1}{r}\frac{\partial\sigma_{\varphi\theta}}{\partial\theta}+\frac{1}{r\sin\theta}\frac{\partial\sigma_{\varphi\varphi}}{\partial\varphi}+\frac{1}{r}(3\sigma_{r\varphi}+2\sigma_{\theta\varphi}\cot\theta)\right)\underline{e}_\varphi\end{aligned}$$

Change of coordinates for the base vectors

$$\begin{cases}\underline{e}_r=\sin\theta\cos\varphi\,\underline{e}_1+\sin\theta\sin\varphi\,\underline{e}_2+\cos\theta\,\underline{e}_3\\ \underline{e}_\theta=\cos\theta\cos\varphi\,\underline{e}_1+\cos\theta\sin\varphi\,\underline{e}_2-\sin\theta\,\underline{e}_3\\ \underline{e}_\varphi=-\sin\varphi\,\underline{e}_1+\cos\varphi\,\underline{e}_2\end{cases}$$

$$\begin{cases}\underline{e}_1=\sin\theta\cos\varphi\,\underline{e}_r+\cos\theta\cos\varphi\,\underline{e}_\theta-\sin\varphi\,\underline{e}_\varphi\\ \underline{e}_2=\sin\theta\sin\varphi\,\underline{e}_r+\cos\theta\sin\varphi\,\underline{e}_\theta+\cos\varphi\,\underline{e}_\varphi\\ \underline{e}_3=\cos\theta\,\underline{e}_r-\sin\theta\,\underline{e}_\theta\end{cases}$$

Change of coordinates for a vector $\underline{u}$

$$\begin{cases}u_1=u_r\sin\theta\cos\varphi+u_\theta\cos\theta\cos\varphi-u_\varphi\sin\varphi\\ u_2=u_r\sin\theta\sin\varphi+u_\theta\cos\theta\sin\varphi+u_\varphi\cos\varphi\\ u_3=u_r\cos\theta-u_\theta\sin\theta\end{cases}$$

$$\begin{cases}u_r=u_1\sin\theta\cos\varphi+u_2\sin\theta\sin\varphi+u_3\cos\theta\\ u_\theta=u_1\cos\theta\cos\varphi+u_2\cos\theta\sin\varphi-u_3\sin\theta\\ u_\varphi=-u_1\sin\varphi+u_2\cos\varphi\end{cases}$$

Change of coordinates for a symmetric second-order tensor $\underline{\underline{\varepsilon}}$

$$\begin{cases}\varepsilon_{11}=\varepsilon_{rr}\sin^2\theta\cos^2\varphi+\varepsilon_{\theta\theta}\cos^2\theta\cos^2\varphi+\varepsilon_{\varphi\varphi}\sin^2\varphi\\ \qquad+2\varepsilon_{r\theta}\sin\theta\cos\theta\cos^2\varphi-2\varepsilon_{r\varphi}\sin\theta\sin\varphi\cos\varphi-2\varepsilon_{\theta\varphi}\cos\theta\sin\varphi\cos\varphi\\ \varepsilon_{22}=\varepsilon_{rr}\sin^2\theta\sin^2\varphi+\varepsilon_{\theta\theta}\cos^2\theta\sin^2\varphi+\varepsilon_{\varphi\varphi}\cos^2\varphi\\ \qquad+2\varepsilon_{r\theta}\sin\theta\cos\theta\sin^2\varphi+2\varepsilon_{r\varphi}\sin\theta\sin\varphi\cos\varphi+2\varepsilon_{\theta\varphi}\cos\theta\sin\varphi\cos\varphi\\ \varepsilon_{33}=\varepsilon_{rr}\cos^2\theta+\varepsilon_{\theta\theta}\sin^2\theta-2\varepsilon_{r\theta}\sin\theta\cos\theta\\ \varepsilon_{12}=\varepsilon_{rr}\sin^2\theta\sin\varphi\cos\varphi+\varepsilon_{\theta\theta}\cos^2\theta\sin\varphi\cos\varphi-\varepsilon_{\varphi\varphi}\sin\varphi\cos\varphi\\ \qquad+2\varepsilon_{r\theta}\sin\theta\cos\theta\sin\varphi\cos\varphi+(\varepsilon_{r\varphi}\sin\theta+\varepsilon_{\theta\varphi}\cos\theta)(\cos^2\varphi-\sin^2\varphi)\\ \varepsilon_{13}=(\varepsilon_{rr}-\varepsilon_{\theta\theta})\sin\theta\cos\theta\cos\varphi+\varepsilon_{r\theta}(\cos^2\theta-\sin^2\theta)\cos\varphi\\ \qquad-\varepsilon_{r\varphi}\cos\theta\sin\varphi+\varepsilon_{\theta\varphi}\sin\theta\sin\varphi\\ \varepsilon_{23}=(\varepsilon_{rr}-\varepsilon_{\theta\theta})\sin\theta\cos\theta\sin\varphi+\varepsilon_{r\theta}(\cos^2\theta-\sin^2\theta)\sin\varphi\\ \qquad+\varepsilon_{r\varphi}\cos\theta\cos\varphi-\varepsilon_{\theta\varphi}\sin\theta\cos\varphi\end{cases}$$

$$\begin{cases}\varepsilon_{rr}=\varepsilon_{11}\sin^2\theta\cos^2\varphi+\varepsilon_{22}\sin^2\theta\sin^2\varphi+\varepsilon_{33}\cos^2\theta\\ \qquad+2\varepsilon_{12}\sin^2\theta\sin\varphi\cos\varphi+2\varepsilon_{13}\sin\theta\cos\theta\cos\varphi+2\varepsilon_{23}\sin\theta\cos\theta\sin\varphi\\ \varepsilon_{\theta\theta}=\varepsilon_{11}\cos^2\theta\cos^2\varphi+\varepsilon_{22}\cos^2\theta\sin^2\varphi+\varepsilon_{33}\sin^2\theta\\ \qquad+2\varepsilon_{12}\cos^2\theta\sin\varphi\cos\varphi-2\varepsilon_{13}\sin\theta\cos\theta\cos\varphi-2\varepsilon_{23}\sin\theta\cos\theta\sin\varphi\\ \varepsilon_{\varphi\varphi}=\varepsilon_{11}\sin^2\varphi+\varepsilon_{22}\cos^2\varphi-2\varepsilon_{12}\sin\varphi\cos\varphi\\ \varepsilon_{r\theta}=\varepsilon_{11}\sin\theta\cos\theta\cos^2\varphi+\varepsilon_{22}\sin\theta\cos\theta\sin^2\varphi-\varepsilon_{33}\sin\theta\cos\theta\\ \qquad+2\varepsilon_{12}\sin\theta\cos\theta\sin\varphi\cos\varphi+(\varepsilon_{13}\cos\varphi+\varepsilon_{23}\sin\varphi)(\cos^2\theta-\sin^2\theta)\\ \varepsilon_{r\varphi}=(\varepsilon_{22}-\varepsilon_{11})\sin\theta\sin\varphi\cos\varphi+\varepsilon_{12}\sin\theta(\cos^2\varphi-\sin^2\varphi)\\ \qquad-\varepsilon_{13}\cos\theta\sin\varphi+\varepsilon_{23}\cos\theta\cos\varphi\\ \varepsilon_{\theta\varphi}=(\varepsilon_{22}-\varepsilon_{11})\cos\theta\sin\varphi\cos\varphi+\varepsilon_{12}\cos\theta(\cos^2\varphi-\sin^2\varphi)\\ \qquad+\varepsilon_{13}\sin\theta\sin\varphi-\varepsilon_{23}\sin\theta\cos\varphi\end{cases}$$

Appendix G
Notations

Scalar Quantities

∂V	surface of a volume V
a	distance between crystallographic planes
	unit cell parameter
	radius of the neck cross-section of a tensile specimen
a_{T}	shift factor (relaxation of polymers)
a_{V}	shift factor (creep of polymers)
a_0	atomic radius
$a_{0\mathrm{S}}$	solute atomic radius
a_1, a_2, a_3	initial coordinates of a point
b	interatomic distance
	Burgers vector amplitude
c	atomic concentration
	volume fraction
c_1	longitudinal wave speed
c_2	shear wave speed
d	distance between dislocations
	distance between slip lines
	distance of a dislocation to an interface
	grain size
	standard deviation
e	density of internal energy (per unit mass)
	extension
f	amplitude of a force acting on a dislocation
	density of free energy (per unit mass)
	volume fraction
$\mathrm{f}(t)$	creep function
f_{t}	total Helmoltz free energy
f_{v}	volume fraction
	Helmoltz free energy per unit volume

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

g	slip amplitude
h, k, l	Miller indexes
h	distance between dislocations in a wall
	distance between slip planes
	work-hardening modulus
k	Boltzmann constant
k	yield strength in simple shear
	stiffness coefficient
	coefficient of compressibility, elastic bulk modulus
l	dislocation line
l	length
	distance between trees
l_{D}	mean distance travelled by a dislocation
l_{G}	.... by a geometrically necessary dislocation
m	fraction transformed
	strain rate sensitivity exponent
m_{S}	Schmid factor
m_{T}	Taylor factor
n	number of dislocations; in particular in a pile-up
	number of jogs
	number of phantom molecular chains of an elastomer
	strain-hardening exponent
	atomic fraction of vacancies
n_i	components of the normal to the slip plane
p	cumulative plastic deformation
	pressure
q	deformation parameter
	displacement
$\dot{q}$	deformation rate parameter
r	polar coordinate (radial)
	rate of heat flow into unit volume
	radius
	inter-atomic distance
r^*	complex modulus
$\mathrm{r}(t)$	relaxation function
$<r>$	mean end to end distances of molecular chains
r_0	core radius of dislocations
	inter-atomic distance at equilibrium
s	curvilinear coordinate
	displacement in twinning
	entropy density (per unit mass)
t	line tension
	wall thickness
	time
t_i	stress components

t_R	time to fracture
u	displacement
u_{GB}	intergranular displacement
v	dislocation velocity
w	width of a precipitate or of an inclusion
	width of a ribbon of stacking fault
	length
w^*	critical length of a dislocation loop
x_c	degree of crystallinity
x_1, x_2, x_3	present coordinates of a point
x, y, z	cartesian coordinates
z	cylindrical coordinate (axial)
A	area swept by a dislocation
	percentage elongation after fracture
A_g	percentage elongation non-proportional at maximum force
C	torque
	concentration of vacancies
C_p	specific heat (per unit mass)
C_∞	polymer chain characteristic ratio
D	diffusion coefficient
	diameter
	dimension
	relative cross-section ratio
D_{GB}	intergranular diffusion coefficient of vacancies
D_L	diffusion coefficient of vacancies
D_v	self-diffusion coefficient
E	Young modulus
	elastic energy
E_d	elastic energy of a dislocation
E_k	energy of a kink
E_{PN}	energy for overcoming Peierls-Nabarro hills
E_T	elastic energy due to external forces
E^*	complex modulus
E'	storage modulus (real part of E^*)
E''	loss modulus (imaginary part of E^*)
F	load
	Helmoltz free energy
	inertia parameter of polymer chain units (molar mass of chains / number of rotatable bonds)
G	Gibbs free energy (free enthalpy)
G'	storage modulus (real part of r^*)
G''	loss modulus (imaginary part of r^*)
ΔG	variation of Gibbs free energy (free enthalpy)
ΔG_f	Gibbs free energy for the creation of a vacancy
$H(t\text{-}\tau)$	Heaviside unit step function

HDT	heat deflection temperature (for polymers)
J	Jacobian
	flux of vacancies
	creep compliance
I_1, I_2, I_3	invariants of the stress tensor
J_1, J_2, J_3	invariants of the deviatoric stress tensor
K	force exerted on a dislocation by an obstacle
	constant in Norton's law
K_{Ic}	fracture toughness
L	distance between points of anchorage of a dislocation
	mean distance between dislocations
	size of a pile-up – length of the slip lines
	dislocation length
	length between marks for a tensile specimen
L_0	initial length between marks for a tensile specimen
L_e	gauge length of extensometer
L_F	size of the Frank's net (mean distance between dislocations)
M	strain rate exponent
	rigidity coefficient ($= \mu/2\pi$ for a screw dislocation $=\mu/2\pi(1\text{-}\nu)$ for an edge dislocation)
	molar mass
M_c	molecular weight
N	number of dislocations
	number of molecular chains per unit volume
	number of loops
	strain-hardening exponent
Q	load parameter
	load
	activation energy
Q'	heat rate
Q_0	activation energy of self-diffusion
P_D	dynamic stiffness of molecules
P_S	static stiffness of molecules
R	Perfect gas constant
R	radius of curvature
	radius of action of a dislocation
	radius of Mohr circle
	radius of a test piece
	radius of a cylindrical vessel
	Lankford coefficient
$\bar{R}$	mean value of the Lankford coefficient
R_c	radius of a dislocation loop
R_e	limit of proportionality
R_e'	yield strength in compression
R_{eH}	upper yield strength

R_{el}	lower yield strength
R_f	average plastic flow stress
R_m	tensile strength
R_p	proof strength
$R_{p0.2}$	conventional proof strength at 0.2% extension
R_t	proof strength total extension
R_r	permanent set strength
R_u	true tensile strength
S_0	initial cross-sectional area of a tensile specimen
S_u	surface over which displacements or velocities are imposed
S_T	surface over which tractions are imposed
ΔS	variation of entropy
T	absolute temperature
T_g	glass transition temperature
T_{gl}	glass transition temperature of a hypothetical linear copolymer
$T_{g\infty}$	characteristic temperature of a polymer (linked to the chain dynamic stiffness)
T_m	melting point
T_β	β transition temperature
U	interaction energy cohesion energy
U_0	cohesion energy at equilibrium
U_L	energy of a dislocation loop
U_L*	activation energy for the escape of a dislocation loop
V	volume striker velocity flow variables associated with the work-hardening parameter
V_L	longitudinal sound wave velocity
V_T	transverse sound wave velocity
$\Delta V/V$	relative volume change in a transformation relative variation of the atomic volume
W	work
W'_{def}	virtual deformation work
W_k	kinetic energy
W_p	work done in plastic deformation
W_S	stored energy
X	atomic fraction
Y	parameter for isotropic work-hardening
Y_m	work-hardening parameter
Z	percentage reduction of area
α	coefficient of thermal expansion portion of the Burgers vector b by which the centre of a dislocation is displaced
α, β	slip lines
α, β	factors

δ	relative cross-section increment ratio
	loss angle (tanδ is the damping factor)
δ_{GB}	thickness of grain boundary
γ	shear strain, slip ($\dot{\gamma}$ slip rate)
γ_A	antiphase energy
γ_f	stacking fault energy
γ_i	interface energy
γ_s	surface energy
γ'	precipitate of Ni_3Al type
ε	strain
ε^v	viscous strain
$\dot{\varepsilon}^v$	viscous strain rate
$\dot{\varepsilon}^{sv}$	secondary creep strain rate
ε_{GB}	intergranular deformation
ε_{hg}	homogeneous strain
ε_p	plastic strain
ε_T	total deformation
η	size factor
	viscosity
λ	angle between the slip direction and the tensile axis
	distance between hard points in work-hardening: cell size
	extension ratio
	Lamé coefficient
μ	shear modulus (or Coulomb modulus)
μ_d	depressed shear modulus
ν	Poisson ratio
	cross-link density in a polymer
ω	angular frequency
π	density of elastic potential
φ	angle of rotation
ϖ	density of complementary elastic potential
ρ	mass density (per unit volume)
	dislocation density
	internal friction coefficient
ρ_a	density of amorphous polymer
ρ_c	density of crystalline polymer
ρ_D	dislocation density
ρ_G	density of geometrically necessary dislocations
ρ_M	density of mobile dislocations
ρ_S	density of dislocation sources
σ	tensile stress
σ_v	"viscous stress"
σ_{ij}	stress tensor components
σ_i, σ_I, σ_{II}, σ_{III}	principal stresses
σ_h	stress parameter in the constitutive equation

σ_m	mean stress (= 1/3 Tr$\underline{\underline{\sigma}}$)
σ_p	tensile yield strength of a single crystal
σ_s	threshold stress
$\bar{\sigma}$	equivalent stress
τ	shear stress
	relaxation time
τ_c	critical resolved shear stress
τ_{chem}	chemical stress (martensitic transformation)
τ_i	internal frictional shear stress on dislocations
τ_M	maximum shear stress
τ_p	flow shear stress
τ_{PN}	Peierls-Nabarro critical shear stress
τ_R	maximum shear stress at the surface of a torsion specimen
τ_w	stress needed to nucleate a twin
τ_0	theoretical yield strength in shear
θ	volume dilatation (= Tr$\underline{\underline{\varepsilon}}$)
	angle, polar coordinate
	cooling rate
ζ	$a/2(1\text{-}\nu)$, a being the distance between slip planes; 2ζ is the half-width of a dislocation
Λ	distance between precipitates
	length of dislocation
ϕ	diameter
	angle between the normal to the slip plane and the tensile axis
	flux of vacancies
Φ	volume dissipation of energy
Ω	volume of a solid
	atomic volume
$\partial\Omega$	boundary of a solid
Δ_{ij}	Green-Lagrange strain tensor
$\dot{\Delta}_{ij}$	Green-Lagrange strain rate tensor
∇^2	Laplacian operator, often written Δ

Vectors

$\underline{b}$	Burgers vector
$\underline{f}$	force per unit length exerted on a dislocation body-force field
$\underline{g}$	vector of reciprocal lattice
$\underline{l}$	unit vector along a dislocation line
$\underline{n}$	unit normal vector
$\underline{q}$	heat flow

$\underline{u}(\underline{x})$	displacement field
$\underline{u}_D$	displacement of the surface of a solid in creating a dislocation
$\underline{u}^{\mathrm{g}}$	imposed displacement
$\underline{u}'$	virtual displacement vector
$\underline{T}$	finite transformation
$\underline{v}(\underline{x})$	velocity field
$\underline{v}^{\mathrm{g}}$	velocity imposed on a boundary
$\underline{v}'(\underline{x})$	kinematically admissible velocity field
$\underline{T}$	external applied forces
$\underline{T}_{\mathrm{D}}$	external forces to create a dislocation
$\underline{\gamma}(\underline{x})$	acceleration field

G.1 Tensors

Second-order tensors are underlined twice; higher order tensors are written with bold roman letters. Matrixes are denoted by square brackets: $[\alpha]$.

$\mathbf{A}$	strain concentration tensor
$\mathbf{B}$	stress concentration tensor
$\underline{\underline{\varepsilon}}$	Eulerian strain tensor
$\underline{\underline{d}}$	Eulerian strain rate tensor
$\underline{\underline{\Delta}}$	Green-Lagrange strain tensor
$\underline{\underline{e}}$	deviatoric strain tensor ($\underline{\underline{\varepsilon}} = \frac{1}{3}\theta\underline{\underline{\delta}} + \underline{\underline{e}}$)
$\underline{\underline{\varepsilon}}'(\underline{x})$	admissible strain field
$\underline{\underline{\dot{\varepsilon}}}'(\underline{x})$	admissible strain rate field
$\underline{\underline{E}}$	macroscopic strain
$\underline{\underline{\varepsilon}}^{\mathrm{PT}}$	transformation strain
$\underline{\underline{F}}$	transformation gradient
$\mathbf{c}$	elastic moduli
$\mathbf{C}$	elastic moduli (C_{ijkl}, C_{IJ})
$\mathbf{S}$	elastic compliances ($S_{ijkl,}S_{IJ}$)
$\mathbf{S}^{\mathrm{Esh}}$	Eshelby tensor
$\underline{\underline{X}}$	kinematic work-hardening parameter
$\underline{\underline{\sigma}}$	Cauchy stress tensor
$\underline{\underline{s}}$	deviatoric stress tensor ($\underline{\underline{\sigma}} = \sigma_m\underline{\underline{\delta}} + \underline{\underline{s}}$)
$\underline{\underline{\sigma}}^*(\underline{x})$	statically admissible stress field
$\underline{\underline{\sigma}}'(\underline{x})$	stress associated with $\underline{\underline{\dot{\varepsilon}}}'(\underline{x})$
$\underline{\underline{\Pi}}$	Piola-Kirchoff stress tensor
$\underline{\underline{\Sigma}}$	macroscopic stress tensor

$\underline{\underline{\delta}}$ second-order unit tensor (or Kronecker tensor) ($\delta_{ij} = 1$ if $i = j$ and 0 otherwise)

$\mathbf{I}$ fourth-order symmetric unit tensor ($I_{ijkl} = 1/2\left(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl}\right)$)

$\underline{\underline{\sigma}} = \sigma_m\underline{\underline{\delta}} + \underline{\underline{s}} = -p\underline{\underline{\delta}} + \underline{\underline{s}}$, *i.e.* $\sigma_{ij} = \sigma_m\delta_{ij} + s_{ij} = -p\delta_{ij} + s_{ij}$, p is the *hydrostatic* stress (pressure).

$\underline{\underline{\sigma}}^*$ statically admissible stress field

G.2 Vector and Tensor Operations

The summation convention is mostly used, repetition of a suffix implying summation over all values for that suffix, *e.g.* $a_{ij}b_j = \Sigma_j a_{ij}b_j$

$\underline{u}.\underline{v} = \sum_j u_j v_j (= u_j v_j)$(inner product)

$\underline{w} = \underline{u} \times \underline{v}$ ($w_i = \in_{ijk} u_j v_k$) (cross product) $\in_{ijk} = 0$ if any two suffixes are equal, $+1$ or -1 according as *ijk* is an even or an odd permutation of 123.

$\underline{\underline{w}} = \underline{u} \otimes \underline{v}$ ($w_{ij} = u_i v_j$) (dyadic or tensorial product)

$\underline{t} = \underline{\underline{\sigma}}.\underline{n}$ ($t_i = \sigma_{ij} n_j$)

$\underline{\underline{\sigma}} : \underline{\underline{\varepsilon}} = \sigma_{ij}\varepsilon_{ij}$

$\mathbf{C} = \underline{\underline{a}} \otimes \underline{\underline{b}}$ ($C_{ijkl} = a_{ij}b_{kl}$)

$\mathbf{B} = \mathbf{A}.\underline{u}$ ($B_{ijk} = A_{ijkl}u_l$)

$\underline{\underline{b}} = \mathbf{A} : \underline{\underline{a}}$ ($b_{ij} = A_{ijkl}a_{kl}$)

$\mathbf{C} = \mathbf{A} \otimes \underline{\underline{a}}$ ($C_{ijklmn} = A_{ijkl}a_{mn}$)

$\underline{f} = \text{div}\ \underline{\underline{\sigma}}$ ($f_i = \sigma_{ij,j}$ for rectangular cartesian coordinates)

$\underline{\underline{\eta}} = \mathbf{Inc}\underline{\underline{\varepsilon}}$ ($\eta_{ij} = \in_{ikl} \in_{jmn} \varepsilon_{km,ln}$ for rectangular cartesian coordinates)

Author Index

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012

Subject Index

D. François et al., *Mechanical Behaviour of Materials: Volume 1: Micro- and Macroscopic Constitutive Behaviour*, Solid Mechanics and Its Applications 180,
DOI 10.1007/978-94-007-2546-1, © Springer Science+Business Media B.V. 2012