G.G.N. Angilella · Antonino La Magna

Editors

Correlations in Condensed Matter under Extreme Conditions

A tribute to Renato Pucci on the occasion of his 70th birthday

Springer

Correlations in Condensed Matter under Extreme Conditions

G.G.N. Angilella · Antonino La Magna
Editors

Correlations in Condensed Matter under Extreme Conditions

A tribute to Renato Pucci on the occasion of his 70th birthday

Springer

Editors
G.G.N. Angilella
Dipartimento di Fisica e Astronomia
University of Catania
Catania
Italy

Antonino La Magna
CNR-IMM
Catania
Italy

Preface

Correlations among several many-body systems manifest themselves at various scales especially under extreme conditions. In collecting invited contributions for this special volume, we especially had in mind many-electron systems both in solids and in molecules or clusters, under high pressure or low temperature, and especially in solids or molecular systems with reduced dimensionality.

These in fact embrace the main research interests of Prof. Renato Pucci, whose 70th birthday we would hereby like to celebrate. Professor Pucci started his career immediately after he graduated in physics from the University of Catania (Italy) in 1968, where he eventually became full professor of structure of matter (condensed matter theory) in 1990. During his long and fruitful career, he enjoyed a constant collaboration with Prof. Norman H. March (Oxford University, UK), with regular visits of Prof. Pucci to Oxford first and of Prof. March to Catania later. Professor Pucci also collaborated with several other distinguished scientists, including Profs. R.G. Parr (North Carolina University, USA), F. Flores (Universidad Autónoma de Madrid, Spain), and L.S. Brizhik (Ukrainian Academy of Science, Kiev, Ukraine). He is the author of more than 150 papers on several peer-reviewed international journals and serves as referee for numerous national and international journals, such as the Physical Review and the Physical Review Letters, and the Journal of Chemical Physics.

Professor Pucci is quite well known within the high-pressure scientific community. After organizing two Archimedes Workshops on high-pressure science, he chaired the European High Pressure Research Group (EHPRG) meeting in 1998 and the Joint International Conference of the International Association for the Advancement of High Pressure Science and Technology (AIRAPT) and of the EHPRG in 2007. He was also scientific secretary of the national congress of the Italian Physical Society (SIF) in the International Year of Physics 2005. For both the EHPRG and the AIRAPT, he has been the member of the respective scientific committees for several terms.

At the University of Catania, Prof. Pucci served as head of the undergraduate program in physics, head of the Department of Physics (later: and Astronomy), dean of the Faculty of Sciences (for two terms), and member of the Board of Administration of the university.

His research interests are quite broad and span from chemistry to physics, especially in the solid state. These are reflected, to some extent, in the topics covered by the present volume, which includes invited contributions from collaborators, colleagues, and friends of Renato's.[1] In particular, Prof. Pucci has given contributions to the fields of positron annihilation, to the many-body theory of surfaces (see Chap. 2 by Flores and Goldberg), to nonlinear effects in electronic transport in low-dimensional systems (see Chap. 15 by Brizhik and Chap. 10 by Deretzis et al.), including polyacetylene (see Chap. 11 by Baldo).

Professor Pucci and his group devoted much attention to superconductivity in novel compounds, such as the high-T_c superconductors (see Chap. 3 by Angilella), especially under high pressure (see Chap. 4 by Schilling), or the ruthenocuprates (see also Chap. 5 by Citro).

Within the field of high-pressure physics, Prof. Pucci and his group predicted that the light alkali metals should become insulating (and then possibly reenter a metallic phase) with increasing pressure (see the reprinted article by Siringo et al. on Chap. 18). This has stimulated much research, both experimental and theoretical, with the eventual identification of insulating phases occurring indeed in lithium and sodium under pressure (see Chap. 3 by Angilella for a review). These findings may also be relevant for the pressurized electronic phases of several electrides (see Chap. 6 by Dong and Oganov).

In chemical and molecular physics, Prof. Pucci is probably best known for the so-called Yang–Parr–Pucci formulae (see Chap. 19 by Echegaray et al.), which are the key working equations for most practical applications of the Fukui functions—a reliable descriptor of molecular electronic structure and chemical reactivity. Together with N.H. March, Prof. Pucci has given contributions in the general area of density functional theory (see the reprint in Chap. 21 by Pucci, for a review).

More recently, Prof. Pucci's group has given contributions to the study of the effect of strain on several transport (both electrical and optical) properties of graphene (see Chap. 3, and especially Sect. 3.4, by Angilella, for a review, Chap. 12 by Pellegrino, Chap. 13 by Deretzis and La Magna, and Chap. 14 by Pidatella and Mazzarello).

Part IV contains some reprints or translations (and one original contribution by Ruggieri, Chap. 24) of several papers by Prof. Pucci et al. on the philosophy and the history of science and on the 'correlations' thereof with another deeply rooted aspect of man's endeavor toward knowledge, with its own peculiar 'language,' *viz.* theology—a subject under 'extreme conditions' in its own respect!

---

[1]Some of these were delivered at a workshop which took place in Catania, on September 23, 2016.

This volume is meant to be a testimony to Prof. Pucci's scientific achievements during his career thus far, a small token of our gratitude toward his efforts as a group leader, and of course a hopeful encouragement for future successes!

Catania, Italy   G.G.N. Angilella
September 2016   Antonino La Magna

Contents

Contributors

J.S.M. Anderson Computational Materials Science Research Team, Minatojima-minami-machi, Chuo-ku, Kobe, Hyogo, Japan

G.G.N. Angilella Dipartimento di Fisica e Astronomia, Università di Catania, Catania, Italy; CNISM, UdR Catania, Catania, Italy; IMM-CNR, UdR Catania, Catania, Italy; INFN, Sez. Catania, Catania, Italy; Scuola Superiore di Catania, Università di Catania, Catania, Italy

P.W. Ayers Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario, Canada

M. Baldo Istituto Nazionale di Fisica Nucleare, Sezione di Catania, Catania, Italy

L.S. Brizhik Bogolyubov Institute for Theoretical Physics, Nonlinear Condensed Matter Physics Department, National Academy of Sciences of Ukraine, Kiev, Ukraine

G. Cassone Institute of Biophysics, Czech Academy of Sciences, Brno, Czech Republic

R. Citro Dipartimento di Fisica "E.R. Caianiello" and Spin-CNR, Università degli Studi di Salerno, Fisciano (SA), Italy

I. Deretzis Istituto per la Microelettronica e Microsistemi (CNR-IMM), Catania, Italy

C.D. Dewhurst Institut Laue-Langevin, Grenoble, France

K. Dikmenli Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario, Canada

Xiao Dong Center for High Pressure Science and Technology Advanced Research, Beijing, China

E. Echegaray Facultad de Química, Nucleus Millennium Chemical Processes and Catalysis (CPC), Pontificia Universidad Católica de Chile, Santiago, Chile

Fernando Flores Departamento de Física Teórica de la Materia Condensada and IFIMAC, Universidad Autónoma de Madrid, Madrid, Spain

G. Forte Dipartimento di Scienze del Farmaco, Università di Catania, Catania, Italy

P.V. Giaquinta Dipartimento di Scienze Matematiche e Informatiche, Scienze Fisiche e Scienze della Terra, Università degli Studi di Messina, Messina, Italy

G. Giriat Laboratory of Nanostructures and Novel Electronic Materials, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

E.C. Goldberg Instituto de Física del Litoral (CONICET-UNL), Santa Fe, Argentina; Departamento Ing. Materiales, Facultad de Ing. Química, Universidad Nacional del Litoral, Santa Fe, Argentina

A. Grassi Dipartimento di Scienze del Farmaco, Università di Catania, Catania, Italy

F. Heidar-Zadeh Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario, Canada

Wilfried B. Holzapfel Department-Physik, University Paderborn, Paderborn, Germany

Russell J. Hemley Department of Civil and Environmental Engineering, The George Washington University, Washington, DC, USA; Lawrence Livermore National Laboratory, Livermore, CA, USA

K.V. Kamenev Centre for Science at Extreme Conditions and School of Engineering, University of Edinburgh, Edinburgh, UK

A. La Magna Istituto per la Microelettronica e Microsistemi (CNR-IMM), Catania, Italy

Yunfeng Liang Department of Physics and Engineering Physics, University of Saskatchewan, Saskatoon, SK, Canada; Research into Artifacts, Center for Engineering, University of Tokyo, Kashiwa Campus, Kashiwa, Japan

G.M. Lombardo Dipartimento di Scienze del Farmaco, Università di Catania, Catania, Italy

S.F. Lombardo Istituto per la Microelettronica e Microsistemi (CNR-IMM), Catania, Italy; Dipartimento di Fisica e Astronomia, Università di Catania, Catania, Italy

N.H. March EMAT, Department of Physics, University of Antwerp, Antwerp, Belgium; Abdus Salam International Centre for Theoretical Physics, Trieste, Italy; Oxford University, Oxford, UK

C.C. Matthai Department of Physics and Astronomy, Cardiff University, Cardiff, UK

R. Mazzarello Institut für Theoretische Festkörperphysik, RWTH Aachen University, Aachen, Germany

Artem R. Oganov Skolkovo Innovation Center, Skolkovo Institute of Science and Technology, Moscow, Russia; State University of New York, Stony Brook, NY, USA; Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia; International Center for Materials Design, Northwestern Polytechnical University, Xi'an, China

Robert G. Parr Department of Chemistry, University of North Carolina, Chapel Hill, NC, USA

N.A. Parzyk Department of Physics, University of Warwick, Coventry, UK

D.M. Paul Department of Physics, University of Warwick, Coventry, UK

F.M.D. Pellegrino NEST and Scuola Normale Superiore and Istituto Nanoscienze-CNR, Pisa, Italy

A. Pidatella Institut für Festkörperphysik, Technische Universität Dresden, Dresden, Germany

R. Pucci Dipartimento di Fisica e Astronomia, Università di Catania, Catania, Italy; CNISM, UdR Catania, Catania, Italy; IMM-CNR, UdR Catania, Catania, Italy

N. Rabi Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario, Canada

S. Rabi Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario, Canada

Fernando Rodríguez MALTA TEAM, DCITIMAC, Facultad de Ciencias, Universidad de Cantabria, Santander, Spain

G. Ruggieri Istituto per le Scienze Religiose Bologna, Bologna, Italy

F. Saija CNR-IPCF, Messina, Italy

A.M. Saitta Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC), Sorbonne Universités, Université Pierre et Marie Curie Paris 06, Paris, France; CNRS, Paris, France

James S. Schilling Department of Physics, Washington University, St. Louis, MO, USA

Hongliang Shi Key Laboratory of Micro-Nano Measurement-Manipulation and Physics, Department of Physics, Beihang University, Beijing, China

Fabio Siringo Dipartimento di Fisica e Astronomia, Università di Catania, and INFN Sezione di Catania, Catania, Italy

A. Toro-Labbe Facultad de Química, Nucleus Millennium Chemical Processes and Catalysis (CPC), Pontificia Universidad Católica de Chile, Santiago, Chile

John S. Tse Department of Physics and Engineering Physics, University of Saskatchewan, Saskatoon, SK, Canada

X. Wang Centre for Science at Extreme Conditions and School of Engineering, University of Edinburgh, Edinburgh, UK

Weitao Yang Department of Chemistry, University of North Carolina, Chapel Hill, NC, USA

Wai-Leung Yim Institute of High Performance Computing, Agency for Science, Technology and Research, Connexis, Singapore

Part I
Condensed Matter Theory

Chapter 1
Can the *d*-Orbital Splitting Unveil the Local Structure of Cu²⁺ Ions? Study on the K₂ZnF₄:Cu²⁺ Archetype

Fernando Rodríguez

Abstract Jahn-Teller (JT) transition-metal ions like Cu^{2+} (d^9) or Mn^{3+} (d^4) in octahedral coordination exhibit larger distortions than non-JT ions like Mn^{2+} or Fe^{3+} (both d^5) in oxides and halides with perovskite type structure. Their mutual interactions eventually determine the type of distortion and the way they couple each other, being at the core of some relevant physical properties. When a JT ion is introduced as an impurity in an octahedral or nearly octahedral site, it provokes a low-symmetry lattice distortion as a consequence of the instability associated with the electronic ground state degeneracy, $e_g(x^2 - y^2, 3z^2 - r^2)$. The distortion degree, ρ, depends mainly on the electron-ion coupling interaction related to the $E \otimes e$ JT effect, and it is modulated by the host crystal structure. This scenario explains for example why Cu^{2+} induces a large distortion of the CuF_6 octahedron, when Cu^{2+} replaces Zn^{2+} either in the perovskite $KZnF_3$ or in the layered perovskite K_2ZnF_4. However, there is a long debate about whether the splitting, Δ_e of the O_h $e_g(x^2 - y^2, 3z^2 - r^2)$ orbitals into a_{1g} and b_{1g} is proportional to ρ or Δ_e contains additional contributions from the rest-of-the-lattice (crystal anisotropy) aside ρ. Elucidation of this controversy is important in order to establish structural correlations between Δ_e and ρ, for an eventual local structure determination of JT impurities from optical spectroscopy. Recent studies on $K_2ZnF_4:Cu^{2+}$ report different views of this problem. Here we show that Δ_e scales linearly with ρ. High pressure experiments and JT-ion compound series of different dimensionality give support for the proposed scenario and provide structural correlations relating ρ and Δ_e in Cu^{2+} and Mn^{3+} systems.

Introduction

Transition-metal (M) inorganic compounds with perovskite type structures $A_n M X_m$ ($n = 1 - 3, m = 3 - 6$, A: alkali or alkaline, X: O or halogen) display an ample variety of structures with relevant physical phenomena such as high T_c superconductivity,

F. Rodríguez (✉)
MALTA TEAM, DCITIMAC, Facultad de Ciencias, Universidad de Cantabria,
E-39005 Santander, Spain
e-mail: fernando.rodriguez@unican.es

3

colossal magnetoresistance, ferro-antiferro-non-magnetic behavior associated with exotic high-spin to low-spin magnetism, metal-insulating transition, multifunctionality, piezochromism and piezomagnetism, or extreme polymorphism at high pressure, which are induced by subtle structural changes around the M ion [1–6]. Correlations between electronic structure and coordination geometry of M are, in most cases, essential to interpret these phenomena. These correlations are particularly important in compounds involving Jahn-Teller (JT) M ions [6–11] since additional structural complexities arise. This situation is even more complicated if we deal with JT M ions introduced as impurities in these materials [12, 13]. Compounds $A_n M X_m$ form lattices of interconnected polyhedra in which the M ion is octahedrally coordinated to X anions as a MX_6 cluster [14, 15]. The actual M coordination symmetry corresponds to a perfect octahedron in case of a pure perovskite structure AMX_3 ($Pm\bar{3}m$). But slight or large deviations of the octahedral symmetry can be attained depending on the crystal structure (composition and dimensionality): structures associated with MX_6 rotations (i.e. as a result of softening of R_{25} and M_3 phonon modes) [16], or displacements of the anion lattice with respect to cation lattices (ferroelectric type transitions) or those attained in low-dimensional systems: layered AMX_4 or $A_2 M X_4$ (with compressed or elongated MX_6 octahedra); linear chains of elongated MX_6 octahedra in $A_2 M X_5$ or AMX_3, or independent MX_6 octahedra in $A_3 M X_6$ or face-sharing structures in AMX_3 are some examples of how the X-sharing and packing of MX_6 octahedra influence their final symmetry [14, 15]. M ions with a $3d^5$ electronic configuration like Fe^{3+} or Mn^{2+}, which have a singlet orbital multielectronic ground state ($^6A_{1g}$), are archetypal examples of this behavior. Nevertheless, octahedral distortions for these M are small and thus their corresponding coordination polyhedra are usually considered as regular octahedra. In case of Mn^{2+} or Zn^{2+}, they can pass from a perfect octahedron in perovskites $TlMnCl_3$ or $RbMnF_3$ [17, 18] to a D_{3d} distorted octahedron in the one dimensional face-sharing $CsMnCl_3$ or $CsMnF_3$ (with three short bonds $R_s = 2.54$ and 2.12 Å, and three long bonds $R_l = 2.58$ and 2.16 Å, respectively.) [19, 20] or to a slightly tetragonal distortion in layered perovskites with either compressed (Rb_2MnCl_4-$4xR_{eq} = 2.525$ Å, and $2xR_{ax} = 2.495$ Å [21]) or elongated (Rb_2MnF_4-$4xR_{eq} = 2.115$ Å, and $2xR_{ax} = 2.126$ Å [22]) octahedron. The effects of crystal anisotropy on the octahedral distortion on passing from 0D to 1D and 2D, can be measured through parameter ρ defined in terms of the M–X bond distances, R_i, as $\rho = \sqrt{\sum_{i=1}^{6}(R_i - R_0)^2}$, with R_0 being the average M–X distance. In case of tetragonal distortions, this parameter coincides with the normal coordinate, Q_θ, which is given in terms of the equatorial and axial bond distances as $Q_\theta = 2/\sqrt{3}(R_{ax} - R_{eq})$. This parameter varies with the dimensionality in the above compounds by 0, 0.05, and -0.035 (and 0.013) Å, respectively.

Results and Discussion

Structural Correlations in Octahedral Jahn-Teller Transition-Metal Ions

Low-symmetry distortions in octahedral M ions with a singlet electronic ground state are significantly smaller than those attained in isostructural compounds in which M has an orbitally degenerate ground state (JT ions). Mn^{3+} ($3d^4$) in fluorides [8, 23, 24] or Cu^{2+} ($3d^9$) in chlorides [25, 26] and fluorides [8, 12] provide examples of isostructural compounds where the MX_6 polyhedra display larger distortions than those involving M ions with nondegenerate ground state, even dealing with compounds of cubic symmetry. This well-known behavior is a consequence of the JT effect exhibited by Mn^{3+} and Cu^{2+}. More precisely, these ions have a doubly degenerate ground state in O_h symmetry (^5E_g and ^2E_g, respectively) both showing strong electron-lattice couplings with lattice (or ligand) distortions of either tetragonal or
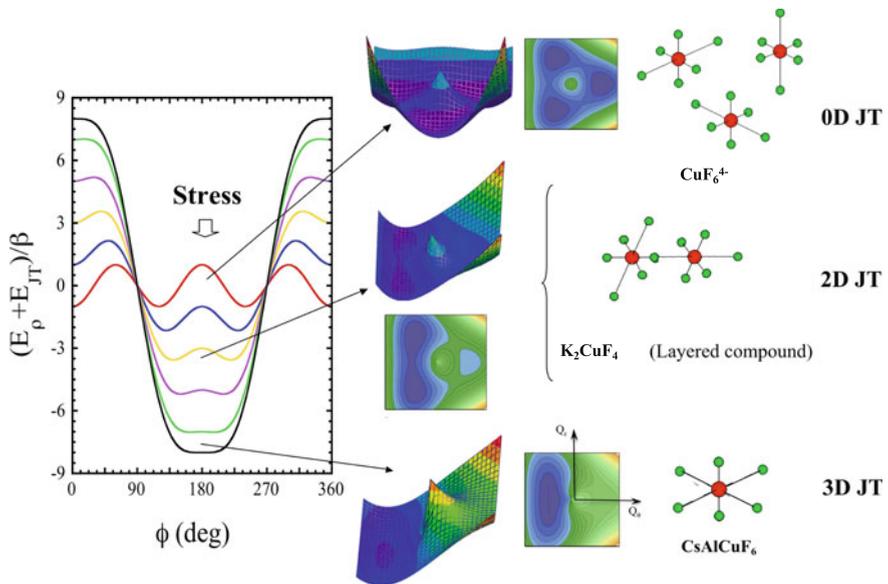
Fig. 1.1 Effect of crystal anisotropy represented through an effective axial stress in the ground-state energy in (Q_θ, Q_ε)-space given by the $E \otimes e$ Jahn-Teller theory [6–11]. The (Q_θ, Q_ε)-space ground-state energy surface is shown on the *left*. Note the evolution from elongated-to-compressed coordination geometry upon increasing the axial stress. The collapse into the compressed geometry takes place at $S_{crit} = 9\beta$. The parameter β contains anharmonic and second-order JT interactions yielding warping of the Mexican-hat-type energy surface in (Q_θ, Q_ε)-space. 2β corresponds to the energy barrier for jumping among energy minima at $S = 0$. Structures shown on the *right-hand side* illustrate CuF_6^{4-} equillibrium geometries predicted by the perturbed JT model

rhombic symmetry, described by local O_h modes (Q_θ, Q_ε) of E symmetry: $E \otimes e$ JT effect [6–12].

Such electron-ion couplings are eventually responsible for the large distortion of the MX_6 octahedra observed experimentally in Mn^{3+} and Cu^{2+} by x-ray diffraction [8, 24, 26]. Phenomenological models based on $E \otimes e$ JT theory [9–12, 27, 28] properly explain the wide variety of distortions of JT M ions in perovskite type crystals. Starting from an octahedral symmetry and considering the linear electron-ion coupling related to $E(Q_\theta, Q_\varepsilon)$ distortions, A, the force constant associated with the elastic distortion, k', the magnitude of the local distortion of the M, ρ, is basically determined by

$$\rho = -\frac{A}{k'}\sqrt{Q_\theta^2 + Q_\varepsilon^2}. \tag{1.1}$$

Anharmonic effects and JT second-order terms of the electron-ion interaction, as well as external stresses (this is introduced to simulate crystal anisotropy effects), aside its magnitude, determine the symmetry of the MX_6 distortion as tetragonal elongated, tetragonal compressed, or any other rhombic intermediate (Fig. 1.1). All possible distortions arose from the warping of the Mexican hat energy surface describing the first order $E \otimes e$ JT coupling configurations [9–12, 27, 28]. In this way the family of Mn^{3+} and Cu^{2+} fluorides illustrate how JT ions exhibit larger octahedral distortions than the singlet-orbital M counterparts. So that the ZnF_6 octahedral distortion in the cubic perovskite $KZnF_3$ and tetragonal layered perovskite K_2ZnF_4, which are $Q_\theta = 0$ ($R_{ax} = R_{eq} = 2.080$ Å) [29] and $Q_\theta = -0.002$ Å ($2xR_{ax} = 2.027$ Å; $4xR_{eq} = 2.029$ Å) [30], respectively, contrasts to distortions in $KCuF_3$ and K_2CuF_4, which are $Q_\theta = 0.38$ Å, $Q_\varepsilon = 0.16$ Å, $\rho = 0.41$ Å, and $Q_\theta = 0.42$ Å, $Q_\varepsilon = 0.09$ Å, $\rho = 0.43$ Å [12], respectively, (Table 1.1). A similar behavior is found for JT Mn^{3+} (^2E ground state) and non-JT Fe^{3+} ($^6A_1$ ground state) fluorides. FeF_6 and MnF_6 distortions are: $Q_\theta = 0$ in Rb_2KFeF_6 [31] and 0.16 Å in Na_3MnF_6 [24] (0D), $Q_\theta = 0.14$ Å in Cs_2FeF_5 [32] and 0.29 Å in Tl_2MnF_5 [24] (1D), and $Q_\theta = 0.09$ Å in $KFeF_4$ [33], and 0.37 Å in $CsMnF_4$ [24] (2D), thus emphasizing the additional distortion displayed by Mn^{3+} due to the $E \otimes e$ JT coupling (Table 1.1). A similar situation is encountered for Cu^{2+} [8, 12, 25, 26, 34].

The different distortion of MnF_6^{3-} depending on the lattice dimensionality, *i.e.*, different F-sharing lattices, is noteworthy. A similar situation is observed in Cu^{2+} fluorides. In particular, the distortion degree increases with the number of Mn^{3+} (or Cu^{2+}) sharing common ligands in the structure: none (0D) in Na_3MnF_6 ($Q_\theta = 0.16$ Å), two (1D) in Tl_2MnF_5 ($Q_\theta = 0.29$ Å), and four (2D) in $CsMnF_4$ ($Q_\theta = 0.37$ Å) [24] or K_2CuF_4 ($Q_\theta = 0.42$ Å) [12]. Such a dimensionality-induced distortion can be easily explained on the basis of the JT models if we take into account that, in an isolated (not sharing; 0D) MnF_6^{3-} or CuF_6^{4-} system, tetragonal distortions are mainly given by $\rho_0 \approx Q_\theta = -A/k'$, which corresponds to the first-order energy-minimum condition of $E = A\rho + \frac{1}{2}k'\rho^2$, yielding $E_{JT} = -\frac{1}{2}A\rho_0 \approx -\frac{1}{2}AQ_\theta$. However, $E(\rho)$ is modified by JT-ion ligand sharing because the lattice distortion energy is shared by neighboring MnF_6^{3-} (or CuF_6^{4-}), and therefore the corresponding distortion increases significantly with respect to isolated systems. In ligand-sharing structures, the

Table 1.1 Structural and spectroscopic parameters for different JT Mn^{3+} and Cu^{2+} fluorides together and other transition metal fluorides for structural comparison. E_1 represents the crystal-filed energy associated with the $b_{1g}(x^2 - y^2) \rightarrow a_{1g}(3z^2 - r^2)$, Δ_e, R_{ax}, R_{eq1}, and R_{eq2} are the M–F distances of the MF_6 complex. The normal coordinates, Q_θ and Q_ε are related to these distances as $Q_\theta = 2/\sqrt{3}(R_{ax} - R_{eq})$ and $Q_\varepsilon = R_{eq1} - R_{eq2}$. The geometric average of both distortions is represented by $\rho = \sqrt{Q_\theta^2 + Q_\varepsilon^2}$. Lattice parameters and space group of some compounds are also included (See text for references of structural and spectroscopic data)

	Compound	$E_1 = \Delta_e$ (eV)	R_{ax} (Å)	E_{eq1} (Å)	E_{eq2} (Å)	E_{eq} (Å)	Q_θ (Å)	Q_ε (Å)	ρ (Å)	Space group	a (Å)	b (Å)	c (Å)
Mn^{3+} (d^4)	K_2NaMnF_6	1.116	2.060	–	–	1.860	0.231	–	0.231	–	–	–	–
	Na_3MnF_6 [b]	2.380	2.018	1.862	1.897	1.880	0.160	0.035	0.164	–	–	–	–
Fe^{3+} (d^5)	Rb_2KFeF_6	0	1.915	1.915	1.915	1.976	0	0	0	Fm-$3m$	8.865	8.865	8.865
Mn^{3+} (d^4)	Na_2MnF_5 [d]	1.550	2.103	1.829	1.872	1.851	0.291	0.043	0.288	$P2_1/c$	7.719	5.236	10.862
Fe^{3+} (d^5)	Cs_2FeF_5 [d]	–	2.07	1.97	1.93	1.95	0.14	0.04	0.14	$Pnmac$	7.84	5.15	12.567
	$NaMnF_4$ [h]	1.915	2.167	1.869	1.808	1.839	0.379	0.061	0.383	$P2_1/c$	5.736	4.892	5.748
Mn^{3+} (d^4)	$CsMnF_4$ [i]	1.922	2.162	1.860	1.821	1.841	0.371	0.039	0.373	$P4/nmm$	7.944	7.944	6.338
Fe^{3+} (d^5)	$KFeF_4$ [i]	–	1.965	1.95	1.88	1.915	0.06	0.07	0.09	$Pmn2_1$	7.608	7.767	12.272
	$KCuF_3$	1.00	2.26	1.98	1.86	1.92	0.39	0.12	0.41				
	CuF_2	0.93	2.32	1.93	1.90	1.915	0.47	0.03	0.47				
Cu^{2+} (d^9)	$CsAlCuF_6$	0.72	1.915	2.06	2.06	2.06	−0.17	0	0.17				
	Ba_2CuF_6	1.20	2.32	1.93	1.85	1.89	−0.50	0.08	0.50				
	K_2CuF_4	1.03	2.27	1.95	1.86	1.91	0.42	0.07	0.43				
Co^{2+} (d^7)	K_2CoF_4	0.05	2.005	2.035	2.035	2.035	−0.034	0	0.034	$I4/mmm$	4.073	4.073	13.087
Ni^{2+} (d^8)	K_2NiF_4	0.00	2.001	2.006	2.006	2.006	−0.006	0	0.006	$I4/mmm$	4.012	4.012	13.076
Zn^{2+} (d^{10})	K_2ZnF_4		2.026	2.029	2.029	2.029	−0.0035	0	0.0035	$I4/mmm$	4.058	4.058	13.019

distortion increases due to JT cooperative effect: $E \approx (1 + n/6)A\rho + \frac{1}{2}k'\rho^2$, where n is the number of shared ligands in MnF_6^{3-} (or CuF_6^{4-}). So that $n = 0$ in isolated systems (0D), 2 in corner-sharing linear chains (1D), 4 in layered perovskites (2D), and 6 in perovskite structure (3D). According to this model, the JT distortion of MX_6 increases with the dimensionality as $\rho_0 \approx -(1 + n/6)A/k'$:

$$\rho_{0D} \approx -\frac{A}{k'}, \tag{1.2a}$$

$$\rho_{1D} \approx -\frac{4}{3}\frac{A}{k'}, \tag{1.2b}$$

$$\rho_{2D} \approx -\frac{5}{3}\frac{A}{k'}, \tag{1.2c}$$

$$\rho_{3D} \approx -2\frac{A}{k'}. \tag{1.2d}$$

This simple model explains why JT ions in ligand-sharing structures like layer perovskites display octahedral distortions which are about twice as large as those observed in structures with isolated JT ions (ligand sharing with non-JT ions).

In general, XRD-determined local distortions in MnF_6^{3-} or CuF_6^{4-} (and $CuCl_6^{4-}$) follow approximately this trend. Values of $\rho = 0.2\,\text{Å}$ (0D), $0.29\,\text{Å}$ (1D), and $0.38\,\text{Å}$ (2D) in Mn^{3+}, and $\rho = 0.2\,\text{Å}$ (0D) and $0.43\,\text{Å}$ (2D), $0.47\,\text{Å}$ (3D) in Cu^{2+} fluorides (Table 1.1) confirm it. This result clearly indicates that the M local distortion in JT systems depends mainly on the electron-ion coupling interaction (A), the elastic force constant associated with the local lattice distortion (k'), and the number of shared ligands with neighboring JT ions (n), aside other minor contributions due to rest-of-the-lattice effects.

Electronic Structure in JT Octahedral-Distorted Transition-Metal Ions

A salient feature concerning the JT distortion in MnF_6^{3-} and CuF_6^{4-} is the additional splitting shown by the parent octahedral e_g and t_{2g} d-orbitals of the M ion. In particular, the splitting of e_g into $a_{1g} + b_{1g}$ can be measured by optical spectroscopy (Fig. 1.2). The corresponding $a_{1g} \rightarrow b_{1g}$ transition energy gives Δ_e ($\approx 4E_{JT}$) directly. According to JT theory, there is a direct first-order relationship between Δ_e and ρ, which is given by $\Delta_e = 2A\rho$. Experimentally, the dependence of Δ_e on ρ has been measured for compound series involving both MnF_6^{3-} [24] and CuF_6^{4-} [12] (and $CuCl_6^{4-}$ [25, 26]). In these systems, the measured Δ_e and ρ show a linear dependence (Figs. 1.3 and 1.4), from which we can extract suitable values of the electron-ion JT coupling: $A = 2.55\,\text{eV/Å}$ for MnF_6^{3-} [24], $A = 1.25\,\text{eV/Å}$ for CuF_6^{4-} (and $A = 1.15\,\text{eV/Å}$ for $CuCl_6^{4-}$) [26], and $A = 1.15\,\text{eV/Å}$ for CuO_6^{10-} (Fig. 1.4), the latter correlation being derived from pressure experiments in $CuWO_4$
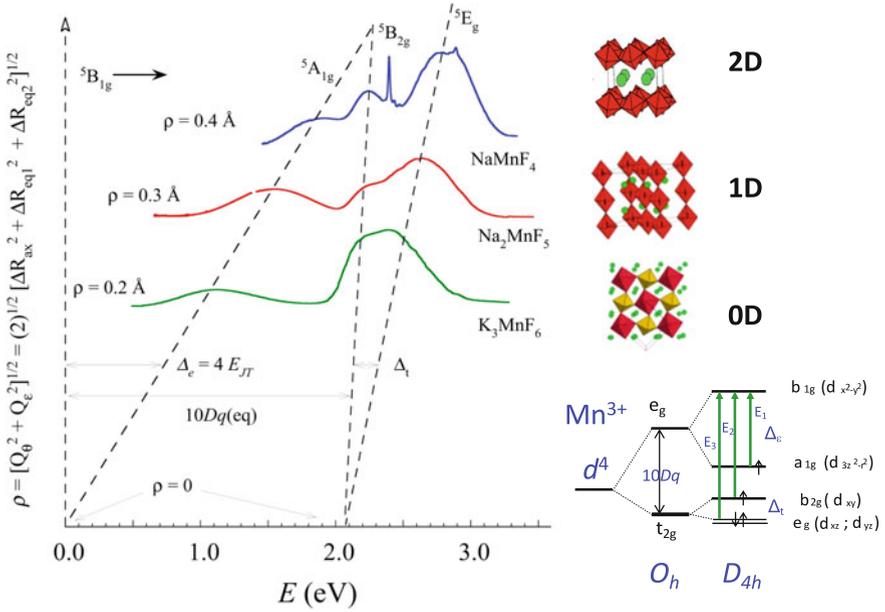
Fig. 1.2 Optical absorption spectra of K_3MnF_6 (0D), Na_2MnF_5 (1D), and $NaMnF_4$ (2D) Mn^{3+} compounds. The three bands shift to higher energies as dimensionality increases. Note that the tetragonal splitting, Δ_e and Δ_t, increase with the distortion parameter ρ defined as a function of the normal coordinates, Q_θ and Q_ε, representing the tetragonal and rhombic distortions, respectively, and the three variations of Mn–F distance with respect to the average Mn–F distance, ΔR_{ax}, ΔR_{eq1}, and ΔR_{eq2}. An O_h symmetry yielding splitting closure is produced for $\rho = 0$. An energy level diagram of the Mn^{3+} *d*-orbitals are shown for octahedral and tetragoanl symmetries on the *right-hand side*. The electronic transition energies are related to the tetragonal splitting of the O_h e_g and t_{2g} orbitals as $\Delta_e = E_1$ and $\Delta_t = E_3 - E_2$. Figure adapted from Fig. 4 of Ref. [28]

[35]. This result is remarkable since it establishes that the low-symmetry local distortion of JT M ions in octahedral coordination can be derived from optical spectroscopy provided that we know the relationship between Δ_e and ρ. Reciprocally, the splitting Δ_e, and hence the JT energy as $E_{JT} = \Delta_e/4$, can be estimated from structural data (ρ) if we know $\Delta_e(\rho)$. The experimentally observed linear relationship between Δ_e and ρ along different crystal structures is significant. It means that Δ_e mainly depends on the low-symmetry distortion of MX_6 irrespective of the particular crystal structure. This implies that the local structure of MX_6 essentially depends on the JT coupling modulated by the type of crystal structure. In any case, the measured Δ_e reflects the actual low-symmetry distortion ρ of MX_6.

An important application of such structural correlations is the local structure determination of JT M ions, which are placed as diluted impurities at octahedral coordination sites through optical spectroscopy measurements. Interestingly, this procedure is simpler and more direct than those based on x-ray absorption
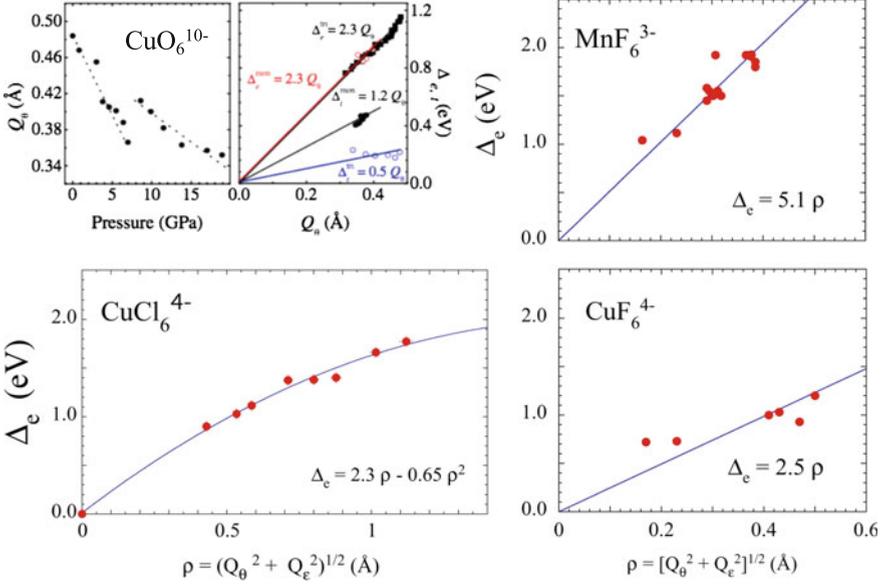
Fig. 1.3 Variation of the parent octahedral e_g splitting, Δ_e, obtained by optical absortion, with the low-symmetry distortion, $\rho = \sqrt{\sum_{i=1}^{6}(R_i - R_0)2} = \sqrt{Q_\theta^2 + Q_\varepsilon^2}$, derived from XRD, in compounds series invloving CuF_6^{4-} [8, 12, 27], $CuCl_6^{4-}$ [25, 26], and MnF_6^{3-} [8, 23, 24]. For Cu^{2+} in six coordination oxides: CuO_6, the corresponding variations of the e_g- and t_{2g}-orbital splittings Δ_e and Δ_t as a function of Q_θ were obtained from optical absorption at high pressure in $CuWO_4$ for low-pressure triclinic and high-pressure monoclinic phases. Results are shown in the *top-left* figure (adapted from Ref. [35]). The coupling coefficient of the observed linear dependence between Δ_e and ρ is included in each plot

spectroscopy (XAS). Besides, XAS has serious limitations if the impurity concentration is below 0.1 mol%.

According to this structural model, the local structure of Cu^{2+} in K_2ZnF_4—ZnF_6 distortion: $\rho = |Q_\theta| = 0.003$ Å [30]—can be derived from the absorption band associated with the $a_{1g} \rightarrow b_{1g}$ transition located at 0.64 eV [12, 34]. Therefore, the local structure of the compressed CuF_6^{4-} is then $\rho = |Q_\theta| = 0.64/2.5 = 0.26$ Å. This distortion fairly agrees to Reinen's estimates ($\rho = |Q_\theta| = 0.27$ Å) obtained from optical spectroscopy and electronic spin resonance [12].

Recently, there has been a controversy for determining the JT distortion of Cu^{2+} in K_2ZnF_4 on the basis of ab initio DFT calculations [13]. Following DFT estimates, the idea of deriving ρ from Δ_e was ruled out in Ref. [13] as Δ_e in Cu^{2+}-doped K_2ZnF_4 contains two main contributions: one is due to the JT effect by Cu^{2+}, and the other one comes from additional contributions to the tetragonal e_g splitting due to rest-of-the-lattice crystal field (beyond CuF_6 coordination shell) due to crystal anisotropy. In fact, the calculated tetragonal crystal-field contribution to the e_g splitting, Δ_e(*lattice*), in K_2ZnF_4 at the Zn^{2+}-substituted Cu^{2+} site is about 50% of Δ_e, the remainder 50% is
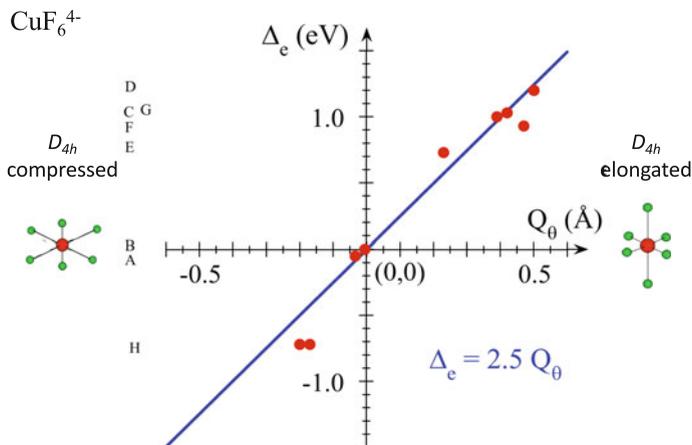
Fig. 1.4 Variation of the parent octahedral e_g splitting, Δ_e, obtained by polarized optical absorption, with the low-symmetry distortion, Q_θ, derived from XRD, in the layered perovskites K_2CoF_4 (*A*), K_2NiF_4 (*B*), K_2CuF_4 (*C*). Structural and spectroscopic data from Ba_2CuF_6 (*D*), $CuZrF_6$ (*E*), CuF_2 (*F*), $KCuF_3$ (*G*), and $CsAlCuF_6$ (*H*), are also included (Table 1.1). Capital letters identify each compound in the $\Delta_e(Q_\theta)$ plot. Note the linear dependence between Δ_e and Q_θ for all transition-metal ions yielding electron-ion coupling coefficient: $A = 1.25\,\text{eV·Å}^{-1}$

due to the JT effect by Cu^{2+} [13]. However, this conclusion is somewhat controversial in view of structural correlations established in compound series. Anyway, DFT simulations rule out the use of such correlations to derive ρ in a JT M ion from optical spectroscopy since Δ_e depends not only on ρ but also on the peculiarities of crystal structure (anisotropy). On the basis of DFT estimates the local distortion as a consequence of the JT coupling is $Q_\theta = -0.13\,\text{Å}$ [13], while a value $Q_\theta = -0.26\,\text{Å}$ is obtained using the experimental structural correlations through ρ —i.e., none direct contribution from rest-of-the-lattice to Δ_e.

In order to ascertain which method provides suitable values of the Cu^{2+} local structure, it must be noted that DFT-based methodology in [13] involves two contributions to Δ_e —JT effect and rest-of-the-lattice crystal field— as: $\Delta_e = \Delta_e(lattice) + \Delta_e(JT)$. The Cu^{2+} local structure in K_2ZnF_4 derived from DFT is $Q_\theta = -0.13\,\text{Å}$. Such a distortion yields $\Delta_e(JT) \approx 0.3\,\text{eV}$, the remainder $0.34\,\text{eV}$ coming from $\Delta_e(lattice)$. In order to estimate $\Delta_e(lattice)$, the d-orbital splitting of Cu^{2+} was calculated by DFT fixing the CuF_6^{4-} structure to a perfect octahedron in the K_2ZnF_4 lattice. With this constraint, the DFT splitting of the parent O_h e_g orbitals of Cu^{2+} is $0.34\,\text{eV}$ [13]. This together with the JT contribution accounts for the experimental Δ_e. Nevertheless, such model estimates lack from experimental verification due to the absence of real systems having Cu^{2+} in a perfect octahedral environment. Actually, sixfold coordinated Cu^{2+} fluorides show large O_h-distorted equilibrium geometries, irrespective of the host crystal structure.

A chief methodologic issue arises at this stage: whether the calculated $\Delta_e(lattice)$ for CuF_6^{4-} in a regular octahedral coordination does really come from the K_2ZnF_4

crystal field at Cu^{2+} beyond the $6F^-$ coordination sphere, or it is caused by a non-symmetric redistribution of the d-orbital electronic density due to instabilities associated with the twofold degeneracy ground state (^2E) in O_h Cu^{2+} (JT-like electronic effect). In fact, such an electronic density redistribution of 'tetragonal' symmetry was claimed elsewhere [13] but keeping artificially the CuF_6 O_h symmetry (identical Cu–F distances). This latter possibility could be named as the static version of the 'forced' JT effect —the instability response to freezing ions in O_h symmetry is to modify the electronic density associated with the e_g orbitals yielding $\Delta_e(lattice)$. However, it must be pointed out that such a situation does not occur in real systems, given that a nonsymmetric electronic density redistribution gives rise necessary to a tetragonal CuF_6 distortion, the degree of which depends mainly on the $E \otimes e$ JT coupling and its type as tetragonal elongated or compressed, or any rhombic intermediate, can be triggered by the crystal anisotropy. The stronger the JT coupling, the larger instability yielding CuF_6 tetragonal distortion in K_2ZnF_4.

In order to shed light on this controversy, proposed models deserve experimental verification. Given the impossibility to find real systems involving CuF_6 with O_h symmetry, we focus on exploring $\Delta_e(lattice)$ in isostructural compounds involving other M ions with nearly O_h symmetry. The tetragonal crystal-field strength at Zn^{2+} by the K_2ZnF_4 lattice beyond the $6F^-$ coordination sphere must be ascertained in isomorphous compounds having M ions in a regular octahedral symmetry, for which Δ_e is known by optical spectroscopy. Layered perovskites K_2CoF_4 [36], K_2NiF_4 [37], and K_2ZnF_4 [30], provide a nearly O_h symmetry for Co^{2+}, Ni^{2+} and Zn^{2+}, with a c/a lattice parameter ratio (crystal anisotropy) of 3.21, 3.26, and 3.23, respectively, (Table 1.1). Interestingly, the crystal anisotropy in K_2ZnF_4 is in between K_2CoF_4 and K_2NiF_4, and thus the tetragonal crystal-field splitting due to the rest-of-the-lattice, $\Delta_e(lattice)$, can be firmly ascertained for this type of structures since Δ_e for Co^{2+} and Ni^{2+} is known from polarized optical absorption [38, 39], and their local geometry from XRD [30, 36, 37] (Table 1.1).

The low temperature polarized optical absorption spectra of K_2CoF_4 and K_2NiF_4 [38, 39] provide information on the tetragonal crystal-field D_s and D_t parameters ($\Delta_e = 4D_s + 5D_t$) [25, 40] by comparing the peak positions measured in the spectra taken with **E** parallel to the principal axis of the refractive-index ellipsoid (uniaxial system)—*i.e.* tetragonal c axis— and perpendicular to it. Polarized absorption measurements indicate that $\Delta_e = -0.05 \pm 0.03$ eV in K_2CoF_4 [38], and $\Delta_e = 0.00 \pm 0.03$ eV in K_2NiF_4 [39]. No difference between the two polarized spectra was observed in the latter system (perfect O_h system). In fact, the measured Δ_e is consistent with the slight O_h distortion of CoF_6^{4-} ($Q_\theta = -0.013$ Å) and NiF_6^{4-} ($Q_\theta = -0.003$ Å) in their respective lattices (Table 1.1), and point out that the rest-of-the-lattice contribution to Δ_e is negligible: $\Delta_e(lattice) \approx 0$. Figure 1.4 shows the variation of the measured Δ_e splitting for Ni^{2+}, Co^{2+}, and Cu^{2+} along the K_2MF_4 series as a function of the local tetragonal distortion at M site, Q_θ, measured by XRD. (Δ_e, Q_θ) values for several Cu^{2+} fluorides are also included in Fig. 1.4. Indeed we observe that Δ_e depends linearly on $Q_\theta (= \rho)$ as $\Delta_e = K Q_\theta$ with $K = 2A = 2.5$ eV/Å. Curiously, present results confirm that Δ_e scales with $\rho (\approx Q_\theta)$ irrespective of the M ion and crystal anisotropy, *i.e.*, $\Delta_e = \Delta_e(JT)$ with $\Delta_e(lattice) \approx 0$.

According to this, the spectroscopically measured $\Delta_e = 0.64\,\text{eV}$ for Cu^{2+} in K_2ZnF_4 is thus consistent with a local distortion $Q_\theta = -0.64/2.5 = -0.26\,\text{Å}$ (compressed octahedron). It is worth noting that the local distortion of Cu^{2+} in K_2ZnF_4 is shorter than the distortion attained in K_2CuF_4 (elongated antiferrodistortive structure, $Q_\theta = 0.47\,\text{Å}$) due to JT sharing. In fact, this difference is consistent with the estimates based on number of neighboring JT ions sharing common ions as described above.

Conclusions

We showed that JT Cu^{2+} ions in sixfold octahedral coordination exhibit large low-symmetry distortions due mostly to the JT effect. The distortion is modulated by the lattice structure depending on the number of neighboring JT ions sharing common ligands. We have established empirical correlations between the *d*-orbital splitting of the parent octahedral $e_g(x^2 - y^2, 3z^2 - r^2)$ orbitals into a_{1g} and b_{1g} levels, Δ_e, and the local structure, ρ, of JT ions Cu^{2+} and Mn^{3+}, through crystal series of different dimensionality and high pressure experiments. We demonstrated that Δ_e exhibits a linear dependence with ρ as $\Delta_e = 2A\rho$. The experimental correlations provide the $E \otimes e$ JT electron-ion coupling coefficient, A, for Mn^{3+} fluorides and Cu^{2+} in chlorides, fluorides, and oxides, as $A = 2.5\,\text{eV·Å}^{-1}$ (Mn^{3+}) and $1.2\,\text{eV·Å}^{-1}$ (Cu^{2+}). We have applied these structural correlations to determine the local structure in the impurity system Cu^{2+}-doped K_2ZnF_4 from the first absorption band associated to the $a_{1g} \rightarrow b_{1g}$ transition, $\Delta_e = 0.64\,\text{eV}$, yielding $Q_\theta = -0.26\,\text{Å}$. We demonstrated that the contribution of the rest-of-the-lattice —beyond CuF_6— to Δ_e is negligible. Correlations in isostructural crystals K_2CoF_4 and K_2NiF_4, involving nearly octahedral Co^{2+} and Ni^{2+}, confirm it. The larger low-symmetry distortion displayed by K_2CuF_4 $(\rho = 0.43\,\text{Å})$ in comparison to the spectroscopically determined distortion in Cu^{2+}in K_2ZnF_4 $(\rho = |Q_\theta| = 0.26\,\text{Å})$ is due to the lack of neighboring JT Cu^{2+} ions in the latter system as Cu^{2+} is introduced as impurity in the Zn^{2+} lattice.

References

1. F. Aguado, F. Rodríguez, R. Valiente, J.P. Itiè, M. Hanfland, Phys. Rev. B **85**, 100101 (2012). DOI 10.1103/PhysRevB.85.100101
2. F. Siringo, R. Pucci, N.H. March, High Pressure Res. **2**(2), 109 (1990). DOI 10.1080/08957959008201442

3. G.G.N. Angilella, R. Pucci, F. Siringo, Phys. Rev. B **54**, 15471 (1996). DOI 10.1103/PhysRevB.54.15471

4. G.G.N. Angilella, A. Sudbø, R. Pucci, Int. J. Mod. Phys. B **14**, 3306 (2000). DOI 10.1142/S0217979200003782

5. R. Citro, G.G.N. Angilella, M. Marinaro, R. Pucci, Phys. Rev. B **71**, 134525 (2005). DOI 10.1103/PhysRevB.71.134525

6. H.A. Jahn, E. Teller, Proc. Roy. Soc. (London) A **161**(905), 220 (1937). DOI 10.1098/rspa.1937.0142

7. R.G. Burns, *Mineralogical applications of crystal field theory* (Cambridge University Press, Cambridge, 1993)

8. D. Oelkrug, Berichte Bunsengesellschaft für physikalische Chemie **70**(7), 736 (1966). DOI 10.1002/bbpc.19660700710

9. D. Reinen, M. Atanasov, Mag. Res. Rev. **15**, 167 (1991)

10. M.A. Hitchman, Comments Inorg. Chem. **15**(3–4), 197 (1994). DOI 10.1080/02603599408035843

11. J.S. Griffith, *The Theory of Transition Metal Ions* (Cambridge University Press, Cambridge, 1980)

12. D. Reinen, Inorg. Chem. **51**(8), 4458 (2012). DOI 10.1021/ic202209c

13. J.A. Aramburu, J.M. García-Lastra, P. García-Fernández, M.T. Barriuso, M. Moreno, Inorg. Chem. **52**(12), 6923 (2013). DOI 10.1021/ic400105z

14. A. Tressaud, J.M. Dance, *Relationships between structure and low-dimensional magnetism in fluorides* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1982), vol. 52, pp. 87–146. DOI 10.1007/BFb0111297

15. E.G. Tulsky, J.R. Long, Chem. Mater. **13**(4), 1149 (2001). DOI 10.1021/cm0007858

16. M.C. Marco de Lucas, F. Rodriguez, M. Moreno, A. Tressaud, J. Phys.: Condens. Matter **6**(31), 6353 (1994). DOI 10.1088/0953-8984/6/31/034

17. M.C. Marco De Lucas, F. Rodríguez, C. Prieto, M. Verdaguer, H.U. Güdel, J. Phys. Chem. Solids **56**(7), 995 (1995). DOI 10.1016/0022-3697(95)00041-0

18. L.J. de Jongh, D.J. Breed, Solid State Commun. **15**(6), 1061 (1974). DOI 10.1016/0038-1098(74)90531-6

19. J. Goodyear, D.J. Kennedy, Acta Cryst. B **29**(4), 744 (1973). DOI 10.1107/S0567740873003286

20. R.E. Schmidt, M. Welsch, S. Kummer-Dörner, D. Babel, Z. Anorg. Allg. Chemie **625**(4), 637 (1999). DOI 10.1002/(SICI)1521-3749(199904)625:4<637::AID-ZAAC637>3.0.CO;2-N

21. J. Goodyear, E.M. Ali, G.A. Steigmann, Acta Cryst. B **33**(9), 2932 (1977). DOI 10.1107/S0567740877009868

22. K. Iwasa, M. Nishi, H. Ikeda, J. Suzuki, J. Phys. Soc. Jpn. **63**(5), 1900 (1994). DOI 10.1143/JPSJ.63.1900

23. F. Rodríguez, P.N. nez, M.C. Marco de Lucas, J. Solid State Chem. **110**(2), 370 (1994). DOI 10.1006/jssc.1994.1182

24. F. Rodriguez, F. Aguado, J. Chem. Phys. **118**(24), 10867 (2003). DOI 10.1063/1.1569847

25. R. Valiente, F. Rodriguez, J. Phys. Chem. Solids **57**(5), 571 (1996). DOI 10.1016/0022-3697(96)00293-4

26. R. Valiente, F. Rodríguez, Phys. Rev. B **60**, 9423 (1999). DOI 10.1103/PhysRevB.60.9423

27. M.J. Riley, L. Dubicki, G. Moran, E.R. Krausz, I. Yamada, Chem. Phys. **145**(3), 363 (1990). DOI 10.1016/0301-0104(90)87045-D

28. F. Aguado, F. Rodríguez, R. Valiente, J.P. Itié, P. Munsch, Phys. Rev. B **70**, 214104 (2004). DOI 10.1103/PhysRevB.70.214104

29. M.C. Marco de Lucas, F. Rodríguez, M. Moreno, Phys. Rev. B **50**, 2760 (1994). DOI 10.1103/PhysRevB.50.2760

30. E. Herdtweck, D. Babel, Z. Kristallogr. **153**(1–4), 189 (1980). DOI 10.1524/zkri.1980.153.14.189

31. A. Tressaud, S. Khaïroun, J.P. Chaminade, M. Couzi, Physica Status Solidi A **98**(2), 417 (1986). DOI 10.1002/pssa.2210980212

32. J.M. Dance, J.L. Soubeyroux, R. Sabatier, L. Fournes, A. Tressaud, P. Hagenmuller, J. Magn. Mag. Mat. **15**, 534 (1980). DOI 10.1016/0304-8853(80)91164-6
33. M. Hidaka, Z.Y. Zhou, B.M. Wanklyn, Physica Status Solidi A **115**(1), 149 (1989). DOI 10.1002/pssa.2211150114
34. D. Reinen, G. Steffen, M.A. Hitchman, H. Stratemeier, L. Dubicki, E.R. Krausz, M.J. Riley, H.E. Mathies, K. Recker, F. Wallrafen, Chem. Phys. **155**(1), 117 (1991). DOI 10.1016/0301-0104(91)87012-K
35. J. Ruiz-Fuertes, A. Segura, F. Rodríguez, D. Errandonea, M.N. Sanz-Ortiz, Phys. Rev. Lett. **108**, 166402 (2012). DOI 10.1103/PhysRevLett.108.166402
36. H. Ikeda, J. Phys. Soc. Jpn. **37**(3), 660 (1974). DOI 10.1143/JPSJ.37.660
37. D. Babel, E. Herdtweck, Z. Anorg. Allg. Chem. **487**(1), 75 (1982). DOI 10.1002/zaac.19824870108
38. J. Ferguson, T.E. Wood, H.J. Guggenheim, Inorg. Chem. **14**(1), 177 (1975). DOI 10.1021/ic50143a038
39. P. Day, C.D. Flint, in *Electronic Structure and Magnetism of Inorganic Compounds*, vol. 6, ed. by P. Day (The Royal Society of Chemistry, 1980), pp. 148–209. DOI 10.1039/9781847555960-00148
40. A.B.P. Lever, *Inorganic Electronic Spectroscopy* (Elsevier, Amsterdam, 1984)

Chapter 2
Introducing the Inner Structure of the Magnetic Atom in the Interaction Between a Transition Metal Atom Impurity and a Metal Surface

Fernando Flores and E.C. Goldberg

Abstract This paper presents a review of the work performed by the authors to incorporate the inner structure of a d-shell magnetic atom in the description of the many-body interaction between a transition metal atom impurity and a metal host. Two main assumptions are made in this approach: (i) the magnetic atom has an orbital singlet with a total electronic spin S (as corresponds to a case for which the angular moment is quenched); (ii) the first Hund's rule determines the inner electronic structure of the magnetic atom. Using these two assumptions and the rotational symmetry of the electronic spin, an ionic Hamiltonian is introduced and, in a further step, the effective exchange metal atom coupling and the impurity Kondo temperature are analyzed.

Introduction

Kondo related problems are still an important issue in many-body condensed matter [1]. In the earliest approaches, very idealistic models [2–5] with a very poor description of the inner structure of the magnetic atom involved in the problem were introduced. This was a very convenient way for understanding the general properties of that basic problem. In particular, in the Anderson-like Hamiltonian a spin

F. Flores (✉)
Departamento de Física Teórica de la Materia Condensada and IFIMAC,
Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain
e-mail: fernando.flores@uam.es

E.C. Goldberg
Instituto de Física del Litoral (CONICET-UNL), Güemes 3450,
S3000GLN Santa Fe, Argentina

E.C. Goldberg
Departamento Ing. Materiales, Facultad de Ing. Química, Universidad Nacional
del Litoral, Santiago del Estero 2829, S3000AOM Santa Fe, Argentina

degenerate level, with an intra-site Coulomb interaction U, interacting with a metal band is introduced. Kondo-like properties of this Hamiltonian are crucially dependent on the correlation effects associated with its twofold degenerate level and the U-value. In section "The Anderson Model Revisited", we present a summary of the well-known main properties of that Hamiltonian, including the effective exchange coupling between the atom and the metal, and its Kondo temperature.

Although it has been a great success of nanoscience the realization of different many-body systems in the quantum wells fabricated using different techniques, Kondo resonances and their electron correlation properties have been found to play also an important role in the interaction between d orbitals (f orbitals) of transition (rare-earth) metal impurities and the electrons of a metal band [6]. However, the analysis of the properties of these impurities has not yet been fully developed except for the so-called N-fold degenerate case [2].

Recent data in the experiments of inelastic tunneling spectroscopy of transition metal magnetic atoms [7–9] or single atom transistors [10] also points out to the need to understand the properties of those systems including the inner structure of the atom with its Hund's rule. A first attempt in this direction was taken by Hirst [11] who introduced the following Hamiltonian:

$$\hat{H}^{\text{Hirst}} = \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}\sigma}^{\dagger} \hat{c}_{\mathbf{k}\sigma} + \sum_{m} \left[E_{n-1} |n-1, m\rangle\langle n-1, m| + E_n |n, m\rangle\langle n, m| \right]$$
$$+ \sum_{\mathbf{k}\sigma m} \left[V_{\mathbf{k}\sigma}^{*}(m) \hat{c}_{\mathbf{k}\sigma}^{\dagger} |n-1, m-\sigma\rangle\langle n, m| \right.$$
$$\left. + V_{\mathbf{k}\sigma}(m) |n, m\rangle\langle n-1, m-\sigma| \hat{c}_{\mathbf{k}\sigma} \right], \tag{2.1}$$

where, for simplicity, only fluctuations of one electron in the atom are considered. In the Hamiltonian, Eq. (2.1), one electron from state $|n, m\rangle$, with n electrons and $S_z = m$, jumps to a $\mathbf{k}$, σ state leaving the atom in a $|n-1, m-\sigma\rangle$ state. Apparently, the Hamiltonian, Eq. (2.1), is rather complicated because it requires many parameters to specify it; however, using Hund's rule and the spin symmetry of the states associated with the d-shell we have introduced an ionic Hamiltonian that only depends on one parameter [12].

In section "The Anderson Model Revisited", the Anderson Hamiltonian is revisited and used to introduce, in section "An Ionic Hamiltonian for d Electrons", the main ideas leading to our generalization of the Hirst Hamiltonian. Then, in section "Effective Exchange Coupling and Kondo Temperature" we review the main properties of that model, including its effective exchange coupling with the metal and its Kondo temperature [13]. Finally, in section "Discussion and Conclusions" we present our main conclusions.

The Anderson Model Revisited

In the Anderson model [3], the interaction between an atom with a nondegenerate d-level and a metal is described by the following Hamiltonian:

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}, \tag{2.2}$$

where

$$\hat{H}_0 = \varepsilon_0 \left[| \uparrow 0\rangle\langle\uparrow 0| + |0 \downarrow\rangle\langle 0 \downarrow |\right] + (2\varepsilon_0 + U)| \uparrow\downarrow\rangle\langle\uparrow\downarrow | + \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} \hat{c}_{\mathbf{k}\sigma}^{\dagger} \hat{c}_{\mathbf{k}\sigma}, \tag{2.3}$$

$|\mathbf{k}\sigma\rangle$ being the metal states, and $| \uparrow 0\rangle, 0 \downarrow\rangle, | \uparrow\downarrow\rangle$ the single and configurations. The atom-metal interaction is described by

$$\hat{H}_{\text{int}} = \sum_{\mathbf{k}} \left\{ V_{\mathbf{k}} \left[| \uparrow 0\rangle\langle 00|\hat{c}_{\mathbf{k}\uparrow} + |0 \downarrow\rangle\langle 00|\hat{c}_{\mathbf{k}\downarrow} + | \uparrow\downarrow\rangle\langle 0 \downarrow |\hat{c}_{\mathbf{k}\uparrow} - | \uparrow\downarrow\rangle\langle\uparrow 0|\hat{c}_{\mathbf{k}\downarrow}\right] + \text{c.c.}\right\} \tag{2.4}$$

with electrons being transferred between the atom and the metal (the metal Fermi energy, E_{F}, is taken as the origin of energies). We stress that, for the convenience of our discussion below, we are using the projector operators for the atom instead of the more conventional creation and annihilation operators [14].

This Hamiltonian has been analyzed using many different (exact and approximate) solutions [2]; we mention here the solution obtained using a many-body technique that uses a Green function method combined with an appropriate self-energy [15, 16]. Figure 2.1 shows the DOS calculated with that approach taking $U = 0.2$, $U/\Gamma = 10$ (Γ is the usual linewidth of the one-electron problem: $\Gamma = \pi |V|^2 \rho_0$, ρ_0 being the metal density of states and $V = V_{\mathbf{k}}$ assuming that $V_{\mathbf{k}}$ is $\mathbf{k}$-independent), and different values of ε (-0.1, -0.15, -0.2, and -0.25). For $\varepsilon_0 = -0.1$, the solution has an electron–hole symmetry with a narrow Kondo peak at the Fermi level; for $\varepsilon_0 = -0.15$ the solution is asymmetric, but still there is a Kondo peak at E_{F}; for $\varepsilon_0 = -0.20$, $\varepsilon_0 + U = 0$, the Kondo peak disappears and the second electron level, $\varepsilon_0 + U$, is already the only one appearing practically in the DOS; finally for $\varepsilon_0 = -0.25$ this regime is completely developed and the one-electron solution coincides practically with the many-body one.

The Hamiltonian, Eq. (2.2) can be transformed into an effective spin scattering Hamiltonian by means of a Schrieffer-Wolff transformation [17], assuming that the atom is mostly in a spin state with $S = 1/2$. We illustrate here briefly how this transformation can be obtained using the projector operators just discussed; this will also be useful for our discussion below.

The main idea is to eliminate the atomic excited states, $|00\rangle, | \uparrow\downarrow\rangle$, using a second-order perturbation theory in $\hat{H}_{\text{int}}$:
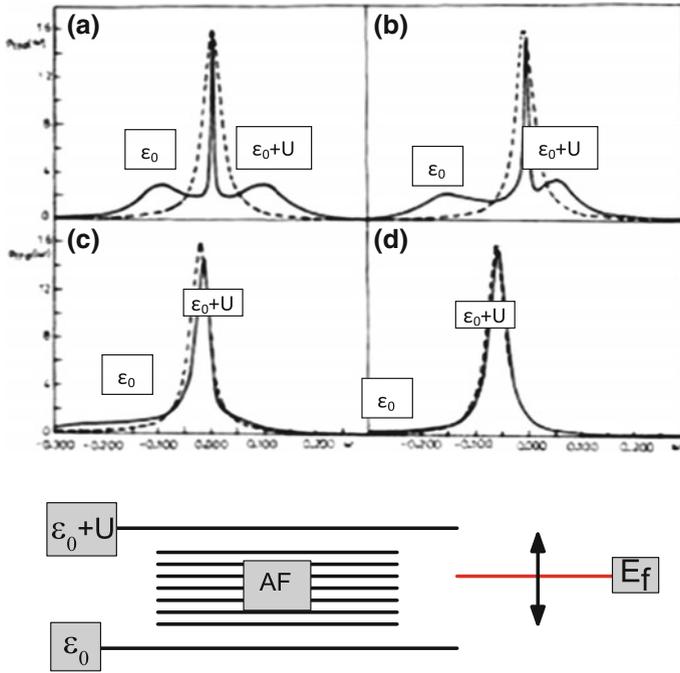
Fig. 2.1 *Upper panel* (from Ref. [14]): DOS for $U/\Gamma = 10$, $U = 0.2$ and **a** $\varepsilon_0 = -0.1$; **b** $\varepsilon_0 = -0.15$; **c** $\varepsilon_0 = -0.2$, and **d** $\varepsilon_0 = -0.25$. *Dashed line* one-electron solution without correlation effects. *Full line* many-body solution following Ref. [15]. *Lower panel* indicates the energy window, $\varepsilon_0 < E_F < \varepsilon_0 + U$, for which the effective atom-metal exchange coupling is antiferromagnetic

$$\hat{H}_{\text{eff}} = -\sum_n \frac{\hat{H}_{\text{int}}|n\rangle\langle n|\hat{H}_{\text{int}}}{E_n - E}. \tag{2.5}$$

Taking $|n\rangle = |00\rangle$ and $|n\rangle = |\uparrow\downarrow\rangle$ in Eq. (2.5), and assuming the energy levels, ε_0 ($\varepsilon_0 + U$), much lower (higher) than the Fermi level, so that $E_0 - E \approx |\varepsilon_0|$ and $E_{\uparrow\downarrow} - E \approx |\varepsilon_0 + U|$, Eq. (2.5) yields:

$$
\begin{aligned}
\hat{H}_{\text{eff}} &= -\sum_{\mathbf{kk'}\sigma\sigma'}\left[\frac{V_{\mathbf{k}}V_{\mathbf{k'}}^*}{|\varepsilon_0|} + \frac{V_{\mathbf{k}}V_{\mathbf{k'}}^*}{|\varepsilon_0 + U|}\right]|\sigma\rangle\langle\sigma'|\hat{c}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k'}\sigma'}^{\dagger}\\
&= -\sum_{\mathbf{kk'}\sigma\sigma'}J_{\mathbf{kk'}}|\sigma\rangle\langle\sigma'|\hat{c}_{\mathbf{k}\sigma}\hat{c}_{\mathbf{k'}\sigma'}^{\dagger},
\end{aligned}
\tag{2.6}
$$

where $|\sigma\rangle = |\uparrow 0\rangle$ for $\sigma = \frac{1}{2}$, and $|\sigma\rangle = |0\downarrow\rangle$ for $\sigma = -\frac{1}{2}$, and we have introduced

$$J_{\mathbf{kk'}} = V_{\mathbf{k}}V_{\mathbf{k'}}^*\left[\frac{1}{|\varepsilon_0|} + \frac{1}{|\varepsilon_0 + U|}\right]. \tag{2.7}$$

Now, introducing the spin operators for the atom and the **k** states, respectively,

$$\hat{S}^+ = |\tfrac{1}{2}\rangle\langle-\tfrac{1}{2}|; \quad \hat{S}^- = |-\tfrac{1}{2}\rangle\langle\tfrac{1}{2}|; \quad \hat{S}_z = \frac{1}{2}\left[|\tfrac{1}{2}\rangle\langle\tfrac{1}{2}| - |-\tfrac{1}{2}\rangle\langle-\tfrac{1}{2}|\right], \quad (2.8a)$$

$$\hat{s}^+_{\mathbf{kk'}} = \hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k'}\downarrow}; \quad \hat{s}^-_{\mathbf{kk'}} = \hat{c}^\dagger_{\mathbf{k}\downarrow}\hat{c}_{\mathbf{k'}\uparrow}; \quad \hat{s}_{z,\mathbf{kk'}} = \frac{1}{2}\left[\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k'}\uparrow} - \hat{c}^\dagger_{\mathbf{k}\downarrow}\hat{c}_{\mathbf{k'}\downarrow}\right], \quad (2.8b)$$

we can write Eq. (2.6) in the following way:

$$\hat{H}_{\text{eff}} = \sum_{\mathbf{kk'}} J_{\mathbf{kk'}} \hat{S}\cdot\hat{s}_{\mathbf{kk'}} + \frac{1}{2}\sum_{\mathbf{kk'}} J_{\mathbf{kk'}}\left[\hat{c}^\dagger_{\mathbf{k}\uparrow}\hat{c}_{\mathbf{k'}\uparrow} + \hat{c}^\dagger_{\mathbf{k}\downarrow}\hat{c}_{\mathbf{k'}\downarrow}\right] - \sum_{\mathbf{k}} J_{\mathbf{kk}}. \quad (2.9)$$

This equation can be written in a more transparent way if we assume a localized description of the **k**-band states, $\phi_{\mathbf{k}} = \sum_\alpha a^{\mathbf{k}}_\alpha \varphi_\alpha$, with the index α denoting the different metal local (Wannier) orbitals. If the **k**-band and the atom interact through a single channel state α [18], the atomic spin part of Hamiltonian, Eq. (2.9), can be written as follows:

$$\hat{H}^{(\text{spin})}_{\text{eff}} = 2J_\alpha \hat{S}\cdot\hat{s}_\alpha, \quad (2.10)$$

a $\hat{S}$-$\hat{s}_\alpha$ antiferromagnetic interaction, with $J_\alpha = |V_\alpha|^2[1/|\varepsilon_0| + 1/|\varepsilon_0 + U|]$ (remember that we have assumed $\varepsilon_0 < E_{\text{F}} < \varepsilon_0 + U$). Equation (2.10) represents the *s-d* model of the Anderson Hamiltonian [19]. Notice that in the limit of $U \to \infty$, $J_\alpha \approx |V_\alpha|^2/|\varepsilon_0|$; however, when $E_{\text{F}} < \varepsilon_0$, $J_\alpha = -|V_\alpha|^2/|\varepsilon_0|$, so that the $\hat{S}$-$\hat{s}_\alpha$ interaction becomes ferromagnetic. Likewise, if the $(\varepsilon_0 + U)$ level is much closer to E_{F} than ε_0, $J_\alpha \approx |V_\alpha|^2/|\varepsilon_0 + U|$; but if $E_{\text{F}} < \varepsilon_0 + U$, $J_\alpha \approx -|V_\alpha|^2/|\varepsilon_0 + U|$, showing that there is a $\hat{S}$-$\hat{s}_\alpha$ ferromagnetic interaction. This is shown schematically in the lower panel of Fig. 2.1, where the window of energy is indicated, $\varepsilon_0 + U > E_{\text{F}} > \varepsilon_0$, for which there appears an antiferromagnetic interaction and, consequently, a Kondo resonance, as it can be seen in the upper panel of Fig. 2.1.

This Kondo-temperature, T_{K}, can be calculated by applying the poor man's scaling method to the antiferromagnetic interaction given by Hamiltonian, Eq. (2.6) [2, 3, 5], whereby metal **k**-states in the energy intervals $(D, D - \delta D)$ and $(-D, -D + \delta D)$ are removed from that Hamiltonian, Eq. (2.6); the metal states are characterized by a constant density of states per spin, ρ_0, and a bandwidth extending from $-D$ to D. This is achieved by using second-order perturbation theory and summing upon the **K**-states in the following equation:

$$\delta H_{\text{eff}} = -\sum_{\mathbf{K}} \frac{\hat{H}_{\text{eff}}|\mathbf{K}\rangle\langle\mathbf{K}|\hat{H}_{\text{eff}}}{E_{\mathbf{K}} - E_0}, \quad (2.11)$$

where $|\mathbf{K}\rangle$ represents excited states like $\hat{c}^\dagger_{\mathbf{k}\sigma}|\sigma'\rangle$ or $\hat{c}_{\mathbf{k}\sigma}|\sigma'\rangle$. This analysis yields the following T_{K}:

$$k_{\text{B}}T_{\text{K}} \approx D\exp\left(-\tfrac{1}{2}\rho_0 J_0\right). \quad (2.12)$$

An Ionic Hamiltonian for *d* Electrons

We describe the interaction between a metal and a *d*-shell atom by the following Hamiltonian:

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}. \tag{2.13}$$

In Eq. (2.13) $\hat{H}_0 = \sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}} \hat{n}_{\mathbf{k}\sigma} + \hat{H}_{\text{atom}}$ includes the energy terms of both, the solid and the atom. The solid is described by the conduction band energies $\varepsilon_{\mathbf{k}}$ with an occupation number given by $\hat{n}_{\mathbf{k}\sigma} = \hat{c}_{\mathbf{k}\sigma}^{\dagger} \hat{c}_{\mathbf{k}\sigma}$. The atomic part, $\hat{H}_{\text{atom}}$, in the extended version appropriate for treating any multi-electron atom [20], takes the form:

$$\hat{H}_{\text{atom}} = \sum_{m\sigma} \varepsilon_m \hat{n}_{m\sigma} + \sum_m U_d \hat{n}_{m\uparrow} \hat{n}_{m\downarrow} + \frac{1}{2} \sum_{m \neq m', \sigma} J_d \hat{n}_{m\sigma} \hat{n}_{m'-\sigma}$$
$$+ \frac{1}{2} \sum_{m \neq m', \sigma} (J_d - J_d^x) \hat{n}_{m\sigma} \hat{n}_{m'\sigma} - \frac{1}{2} \sum_{m \neq m', \sigma} J_d^x \hat{c}_{m\sigma}^{\dagger} \hat{c}_{m-\sigma} \hat{c}_{m'-\sigma}^{\dagger} \hat{c}_{m'\sigma}. \tag{2.14}$$

Here, $\hat{c}_{m\sigma}^{\dagger}$ ($\hat{c}_{m\sigma}$) are the fermionic operators creating (annihilating) an electron with spin projection σ in the orbital m and $\hat{n}_{m\sigma} = \hat{c}_{m\sigma}^{\dagger} \hat{c}_{m\sigma}$; the intra-atomic Coulomb interactions U_d and J_d, as well as the intra-atomic exchange interaction J_d^x, are assumed to be constants independent of the orbital index m. The last term, related to spin-flip processes, restores the invariance under rotation in spin space.

The interaction term, $\hat{H}_{\text{int}}$, contemplates the charge exchange between the atom and the solid through a one-electron tunneling mechanism described by the following expression:

$$\hat{H}_{\text{int}} = \sum_{\mathbf{k}m\sigma} \left[V_{\mathbf{k}m} \hat{c}_{\mathbf{k}\sigma}^{\dagger} \hat{c}_{m\sigma} + V_{\mathbf{k}m}^* \hat{c}_{m\sigma}^{\dagger} \hat{c}_{\mathbf{k}\sigma} \right]. \tag{2.15}$$

In transition metal atoms where the orbital contribution to the angular moment is often quenched due to crystal field effects associated with the low symmetry of the environment (this implies that the ground state of the atom is an angular singlet), the maximum spin associated with a given number of electrons, say N, determines the atomic ground state (first Hund rule). Figure 2.2 shows the atomic states and their corresponding energies as defined by the atomic Hamiltonian, Eq. (2.14), for $N \leq 5$; for $N > 5$, we find a hole–electron symmetry with respect to the previous case (their energies are included in the caption of Fig. 2.2).

Then, we introduce the electronic states of total spin S and spin projection M, $|S, M\rangle_e$ and $|S, M\rangle_h$ for $N \leq 5$ and $N > 5$ respectively, to rewrite the atomic Hamiltonian, Eq. (2.14), as follows:

$$\hat{H}_{\text{atom}} = \sum_{SM} \left[E_{S,e} |S, M\rangle_e \langle S, M|_e + E_{S,h} |S, M\rangle_h \langle S, M|_h \right], \tag{2.16}$$

$$E_s = 5\varepsilon_d + 10J_d - 10J_d^x$$
$$S=5/2$$

$$E_s = 4\varepsilon_d + 6J_d - 6J_d^x$$
$$S=2$$

$$E_s = 3\varepsilon_d + 3J_d - 3J_d^x$$
$$S=3/2$$

$$E_s = 2\varepsilon_d + J_d - J_d^x$$
$$S=1$$

$$E_s = \varepsilon_d$$
$$S=1/2$$

$$E_s = 0$$
$$S=0$$
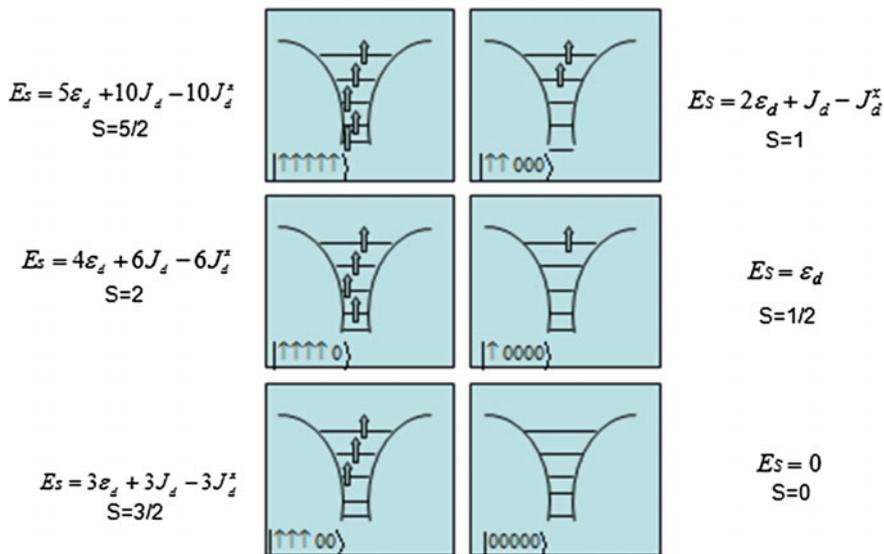
Fig. 2.2 It is shown the atomic d^N states and their energies for different number of electrons, N; $S = N/2$ and M is taken equal to S ($N \leq 5$). For $N > 5$ we find: $E_S = 6\varepsilon_d + U_d + 14J_d - 14J_d^x$ for $N = 6$; $E_S = 7\varepsilon_d + 2U_d + 19J_d - 19J_d^x$ for $N = 7$; $E_S = 8\varepsilon_d + 3U_d + 25J_d - 25J_d^x$ for $N = 8$; $E_S = 9\varepsilon_d + 4U_d + 32J_d - 32J_d^x$ for $N = 9$ and $E_S = 10\varepsilon_d + 5U_d + 40J_d - 40J_d^x$ for $N = 10$

where the energies E_S are given in Fig. 2.2. In this way we reduce the configuration space of Hamiltonian, Eq. (2.14), to the one spanned by those eigenvalues, so that we have $\sum_{SM} [|S, M\rangle_e \langle S, M|_e + |S, M\rangle_h \langle S, M|_h] = 1$.

Notice that the one-electron levels, $E(d^N)$, associated with the many-body Hamiltonian, Eq. (2.16), are defined by the equation: $E(d^N) = E_S(N) - E_{S-\frac{1}{2}}(N - 1)$. This yields:

$$E(d^1) = \varepsilon_d, \tag{2.17a}$$

$$E(d^2) = \varepsilon_d + J_d - J_d^x, \tag{2.17b}$$

$$E(d^3) = \varepsilon_d + 2J_d - 2J_d^x, \tag{2.17c}$$

$$\cdots$$

$$E(d^5) = \varepsilon_d + 4J_d - 4J_d^x, \tag{2.17d}$$

$$E(d^6) = \varepsilon_d + U_d + 4J_d - 4J_d^x, \tag{2.17e}$$

$$E(d^7) = \varepsilon_d + U_d + 5J_d - 5J_d^x, \tag{2.17f}$$

$$\cdots$$

$$E(d^{10}) = \varepsilon_d + U_d + 8J_d - 8J_d^x. \tag{2.17g}$$

For $E(d^N) < E_F < E(d^{N+1})$ we can expect the atom to have N electrons. However, when E_F approaches the levels $E(d^N)$ or $E(d^{N+1})$, the atom should start to have fluctuations to either $N - 1$ or $N + 1$ electrons.

We describe those metal atom processes by assuming that an atomic ground state with N electrons, $|S, M\rangle$, has fluctuations to the states with $N - 1$ or $N + 1$ electrons, which for $N < 5$, correspond to the states $|S - \frac{1}{2}, M\rangle_e, |S + \frac{1}{2}, M\rangle_e$ respectively. For the sake of simplicity, we restrict our discussion to $N < 5$, and neglect the sub-index e from the atomic states. Then, the corresponding interaction Hamiltonian has the form:

$$
\hat{H}_{\text{int}} = \sum_{kM\sigma} \left[V^{S*}_{kM\sigma} \hat{c}^\dagger_{k\sigma} |S - \tfrac{1}{2}, M - \sigma\rangle\langle S, M| + V^S_{kM\sigma} |S, M\rangle\langle S - \tfrac{1}{2}, M - \sigma| \hat{c}_{k\sigma} \right]
$$
$$
+ \sum_{kM\sigma} \left[V^{S+\frac{1}{2}*}_{kM\sigma} \hat{c}^\dagger_{k\sigma} |S, M - \sigma\rangle\langle S + \tfrac{1}{2}, M| + V^{S+\frac{1}{2}}_{kM\sigma} |S + \tfrac{1}{2}, M\rangle\langle S, M - \sigma| \hat{c}_{k\sigma} \right],
$$

$$(2.18)$$

where the different interaction elements, $V^S_{kM\sigma}$ and $V^{S+\frac{1}{2}}_{kM\sigma}$, are calculated by ensuring to have spin rotation invariance. This is illustrated in Fig. 2.3, where we
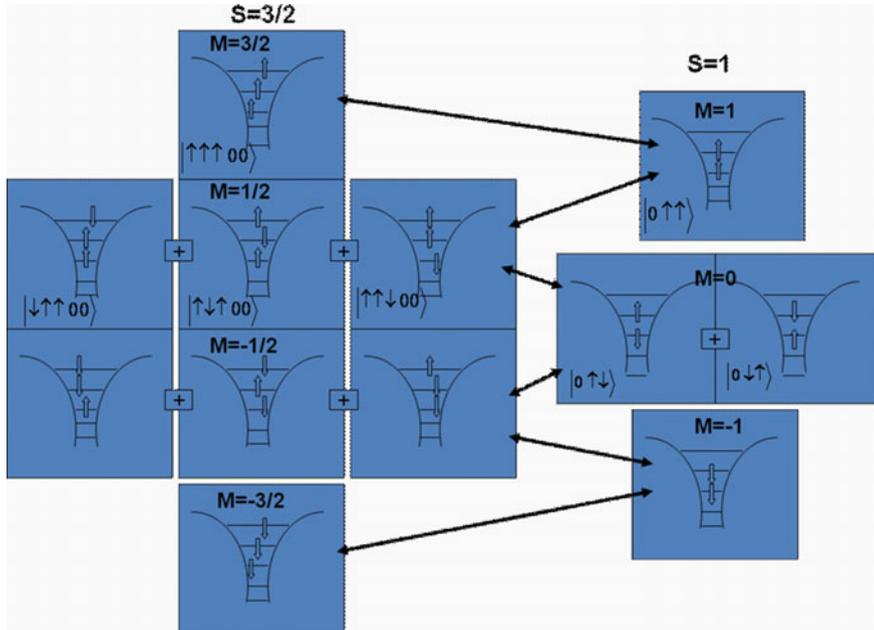
Fig. 2.3 Illustrates how to calculate the matrix elements between the states $|S, M\rangle$ and $|S - \frac{1}{2}, M'\rangle$ due to the interaction: $\sum_{k\sigma} \left[V_k \hat{c}^\dagger_{k\sigma} \hat{c}_{1\sigma} + V^*_k \hat{c}^\dagger_{1\sigma} \hat{c}_{k\sigma} \right]$

consider states with $S = \frac{3}{2}$ and $S = 1$, and assume to have the following interaction: $\sum_{\mathbf{k}\sigma} \left[V_{\mathbf{k}} \hat{c}_{\mathbf{k}\sigma}^{\dagger} \hat{c}_{1\sigma} + V_{\mathbf{k}}^{*} \hat{c}_{1\sigma}^{\dagger} \hat{c}_{\mathbf{k}\sigma} \right]$ between them, where 1 refers to the first d state in the sequence of the 5 d-states shown in that figure. In Fig. 2.3, the different $|S, M\rangle$ or $|S - \frac{1}{2}, M'\rangle$ states are generated by successive application of the operator $\hat{S}^{-}$. In this way the following expression for the coupling terms, $V_{\mathbf{k}M\sigma}^{S}$ in Eq. (2.18), with $N \to N - 1$, is calculated for the case of a half-filled or less than half-filled shell ($N \leq 5$) [21]:

$$V_{\mathbf{k}M\sigma}^{S} = \sqrt{\frac{S + (-1)^{p} M}{2S}} V_{\mathbf{k}d}. \tag{2.19}$$

For the sake of completeness, we also mention that for $N \geq 5$ and fluctuation from N to $N + 1$ (S to $S - \frac{1}{2}$ in the hole picture), we obtain:

$$V_{\mathbf{k}M\sigma}^{S} = (-1)^{p} \sqrt{\frac{S - (-1)^{p} M}{2S}} V_{\mathbf{k}d}. \tag{2.20}$$

Effective Exchange Coupling and Kondo Temperature

Equations (2.16), (2.18), and (2.19) define our ionic Hamiltonian. In this section, we are going to analyze the effective exchange coupling and the Kondo temperature associated with it [13]. We calculate the effective exchange coupling by using the Schrieffer–Wolff transformation as performed above for the case $S = \frac{1}{2}$. We also start by considering only the atomic fluctuations, $|S, M\rangle \to |S - \frac{1}{2}, M'\rangle$ and $N \leq 5$. Then, using Eq. (2.5), and $|n\rangle = |S - \frac{1}{2}, M\rangle$ we find the following effective Hamiltonian:

$$\hat{H}_{\mathrm{eff}} = \sum_{\substack{\mathbf{k}M\sigma \\ \mathbf{k}'\sigma'}} \frac{V_{\mathbf{k}M\sigma}^{S} \hat{c}_{\mathbf{k}\sigma} |S, M\rangle \langle S, M - \sigma + \sigma'| V_{\mathbf{k}'M-\sigma+\sigma'\sigma'}^{*} \hat{c}_{\mathbf{k}'\sigma'}^{\dagger}}{E^{S} - E^{S-\frac{1}{2}} - (\varepsilon_{\mathbf{k}'} - \varepsilon_{\mathbf{k}})}. \tag{2.21}$$

Next, approximate $E^{S} - E^{S-\frac{1}{2}} - (\varepsilon_{\mathbf{k}'} - \varepsilon_{\mathbf{k}})$ by $E^{S} - E^{S-\frac{1}{2}} = E(d^{N}) \equiv -\Delta$, use the expression, Eq. (2.19) for $V_{\mathbf{k}M\sigma}^{S}$, the following equations for $\hat{S}$,

$$\hat{S}^{+} = \sum_{M} \sqrt{S(S + 1) - M(M + 1)} |S, M + 1\rangle \langle S, M|, \tag{2.22a}$$

$$\hat{S}^{-} = \sum_{M} \sqrt{S(S + 1) - M(M - 1)} |S, M - 1\rangle \langle S, M|, \tag{2.22b}$$

$$\hat{S}_{z} = \sum_{M} M |S, M\rangle \langle S, M|, \tag{2.22c}$$

and the expressions of $\hat{s}^+_{\mathbf{kk'}}$, $\hat{s}^-_{\mathbf{kk'}}$, and $\hat{s}_{z,\mathbf{kk'}}$ given by Eq. (2.8b). All this leads to

$$\hat{H}_{\text{eff}} = \sum_{\mathbf{kk'}} \frac{J_{\mathbf{kk'}}}{S} \hat{S} \cdot \hat{s}_{\mathbf{kk'}} + \sum_{\mathbf{kk'}} \frac{J_{\mathbf{kk'}}}{2} \left(\hat{c}^\dagger_{\mathbf{k'}\uparrow} \hat{c}_{\mathbf{k}\uparrow} + \hat{c}^\dagger_{\mathbf{k'}\downarrow} \hat{c}_{\mathbf{k}\downarrow} - 2\delta_{\mathbf{kk'}} \right) \qquad (2.23)$$

which defines a Heisenberg exchange interaction between a local moment and the conduction electrons with a coupling constant $J_{\mathbf{kk'}} = V_{\mathbf{k}} V^*_{\mathbf{k'}}/\Delta$.

The spin part of this effective Hamiltonian in the case where the $\mathbf{k}$-band and the atom interact through a single channel is given by:

$$\hat{H}^{(\text{spin})}_{\text{eff}} = \frac{J_\alpha}{S} \hat{S} \cdot \hat{s}_\alpha, \qquad (2.24)$$

with $J_\alpha = |V_\alpha|^2/\Delta$. This is an antiferromagnetic exchange interaction between $\hat{S}$ and $\hat{s}_\alpha$ with a coupling constant J_α/S when $\Delta > 0$, or equivalently when $E(S) - E(S - \frac{1}{2}) = E(d^N) < 0$.

Up to this point, we have assumed the atom to fluctuate from the spin S (the normal state of the atom) to spin $S - \frac{1}{2}$. If the atom fluctuates to $S + \frac{1}{2}$, our analysis yields that the effective spin Hamiltonian is (written in the channel representation):

$$\hat{H}^{(\text{spin})}_{\text{eff}} = -\frac{J_\alpha}{S + \frac{1}{2}} \hat{S} \cdot \hat{s}_\alpha, \quad \text{for } S \to S + \frac{1}{2}, \qquad (2.25)$$

where $J_\alpha = |V_\alpha|^2/\Delta'$, with $\Delta' = E(d^{N+1})$. This is a ferromagnetic exchange interaction if $E(d^{N+1}) > 0$.

Notice that in all these cases, there appear renormalization factors, S or $S + \frac{1}{2}$, changing the J-coupling, defined in the conventional way $J_\alpha = |V_\alpha|^2/\Delta$, to either J/S or $J/(S + \frac{1}{2})$. Then, we have that a large spin would imply a reduction in the effective $\hat{S} \cdot \hat{s}_\alpha$ interaction and in the corresponding Kondo resonance for the antiferromagnetic case. This is illustrated in the results shown in Fig. 2.4. In this figure, we show the impurity atom spectral densities for $0 \le N \le 5$ electrons in the d-shell and $S \to S - \frac{1}{2}$ fluctuations. These ones have been obtained from the imaginary part of the Green function defined in the projection operator language,

$$G^{SM}_{S-\frac{1}{2}\, M-\frac{1}{2}}(t, t') = i\theta(t - t') \left\langle \left\{ |S, M\rangle\langle S - \tfrac{1}{2}, M - \tfrac{1}{2}|_{t'}, \ |S - \tfrac{1}{2}, M - \tfrac{1}{2}\rangle\langle S, M|_t \right\} \right\rangle \qquad (2.26)$$

and solved by using the equation of motion technique up to second order in the coupling term $V_{\mathbf{k}}$ [12].

From this figure a decreasing antiferromagnetic exchange interaction between $\hat{S}$ and $\hat{s}_\alpha$ accordingly with an effective coupling constant J/S results evident. Consistently, the corresponding Kondo temperature should decrease with S (see below). Experimental evidences of these results exist from long time ago [22] being recently confirmed by using scanning tunneling spectroscopy [23]. Jamneala et al. have found

Fig. 2.4 The impurity density of states $\rho(\omega)$ around the Fermi level for the antiferromagnetic case with $\Delta = 0.1D$ and $\Gamma = 0.01D$. From up to down: $d^1 \leftrightarrow d^0$, $d^2 \leftrightarrow d^1$, $d^3 \leftrightarrow d^2$, $d^4 \leftrightarrow d^3$, $d^5 \leftrightarrow d^4$. The energies ω are measured in units of D. From Ref. [21]
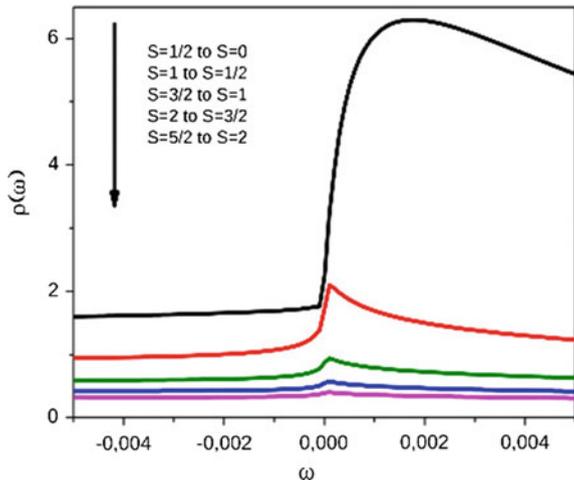
that atoms near the middle of the $3d$ row, such as V, Cr, Mn and Fe, show no discernible features of the local density of states at low energy, while atoms near the ends of the row, such as Ti, Co, and Ni, show narrow resonances near the Fermi energy [23].

Although the cases $S \rightarrow S - \frac{1}{2}$ and $S \rightarrow S + \frac{1}{2}$ have been presented independently, they can be combined simultaneously in one equation because, in the second-order perturbation theory used to calculate the exchange interaction, both terms appear as contributions that add to each other in the effective Hamiltonian. This means that we can combine Eqs. (2.24) and (2.25) into the equation

$$\hat{H}_{\text{eff}}^{(\text{spin})} = \frac{J_\alpha}{S} \hat{S} \cdot \hat{s}_\alpha - \frac{J_\alpha'}{S + \frac{1}{2}} \hat{S} \cdot \hat{s}_\alpha, \quad \text{for } S \rightarrow S - \frac{1}{2} \text{ and } S \rightarrow S + \frac{1}{2}(N < 5).$$
(2.27)

A similar argument can be applied to the case $N \geq 5$, with the magnetic atom of spin S fluctuating to spin $S - \frac{1}{2}$ or $S + \frac{1}{2}$, S defining the normal state of the atom. In general, our analysis yields the same result, Eq. (2.27), for the corresponding spin effective Hamiltonian. It should be stressed, however, that there is an exception to Eq. (2.27), when the normal state of the atom corresponds to the half-filled shell ($N = 5$). In such a case the N to $N - 1$ fluctuation is related to the spin fluctuation $S \rightarrow S - \frac{1}{2}$ within the electron picture, and the N to $N + 1$, in the hole picture, is related to the same spin fluctuation $S \rightarrow S - \frac{1}{2}$. Therefore, due to the hole–electron symmetry of the problem we find that both fluctuations have an antiferromagnetic character if $E(d^N) < E_F < E(d^{N+1})$. Then, the effective spin Hamiltonian reads as:

$$\hat{H}_{\text{eff}}^{(\text{spin})} = \frac{J_\alpha}{S} \hat{S} \cdot \hat{s}_\alpha + \frac{J_\alpha'}{S} \hat{S} \cdot \hat{s}_\alpha \quad (S \text{ corresponds to the half-filled shell}).$$
(2.28)

On the other hand, it is worth mentioning that our results for the effective Hamiltonian keep the rotational symmetry of the problem in all the cases, giving an independent confirmation to the validity of the ionic Hamiltonian introduced for d-transition metal atoms.

Regarding the Kondo temperature associated with our ionic Hamiltonian, we can start our analysis with the spin effective Hamiltonians, Eqs. (2.27) and (2.28), and write

$$\hat{H}_{\text{eff}}^{(\text{spin})} = J' \hat{S} \cdot \hat{s}_\alpha, \tag{2.29}$$

where $J' = (J_\alpha/S) - (J'_\alpha + \frac{1}{2})$ for $S \neq \frac{5}{2}$, or $J' = (J_\alpha/S) + (J'_\alpha/S)$ for $S = \frac{5}{2}$. Equation (2.29) allows us to calculate the Kondo temperature for the antiferromagnetic case by using results obtained by other researchers for the same Hamiltonian, Eq. (2.29) [1, 4]. In this way, we can write:

$$k_B T_K \approx D \exp(-1/\rho_0 J'). \tag{2.30}$$

Assuming that the term J_α/S dominates the J-value for the case $S \neq \frac{5}{2}$, we obtain the following Kondo temperature

$$k_B T_K \approx D \exp(-S/\rho_0 J_\alpha) \quad (S \neq \tfrac{5}{2}). \tag{2.31a}$$

This equation indicates how the Kondo temperature decreases with increasing values of S. On the other hand, for $S = \frac{5}{2}$, if we assume that $E_F = [E(d^6) + E(d^5)]/2$ and $J_\alpha = J'_\alpha$, we can write

$$k_B T_K \approx D \exp(-S/2\rho_0 J_\alpha) \quad (S = \tfrac{5}{2}). \tag{2.31b}$$

Discussion and Conclusions

In this paper, we have discussed the properties of an ionic Hamiltonian introduced to describe the interaction between a metal band and a d-shell magnetic atom with an orbital singlet. As an introduction to that general problem, we have reviewed in section "The Anderson Model Revisited" the properties of the Anderson Hamiltonian, where only one-electron d-level is introduced; in particular we have discussed for this Hamiltonian the exchange coupling between the metal and the atomic spin as well as the corresponding Kondo temperature.

In section "An Ionic Hamiltonian for d Electrons", we have presented our ionic Hamiltonian [12] and have shown how the different hopping parameters between the metal and the atomic wave functions, $|S, M\rangle$, defined by the first Hund's rule can be calculated, up to a factor, using the spin rotational symmetry.

In section "Effective Exchange Coupling and Kondo Temperature", we have discussed [13] how to apply the Schrieffer–Wolff transformation to our ionic

Hamiltonian to obtain the effective exchange coupling interaction between the metal and the atom. Our results show that this effective interaction has the form: $\Gamma \hat{S} \cdot \hat{s} + \gamma \hat{I}$, confirming that our Hamiltonian has the appropriate rotational spin symmetry. Consider the case $N < 5$; S is related to the number of electrons in the d-shell by the equation $2S = N$. The atomic occupancy is determined by the position of the Fermi level; for $E(d^{N+1}) > E_F > E(d^N)$, we can expect the atom to be in the d^N-state (with N electrons and $S = N/2$), and to develop charge fluctuations to states with either $N + 1$ (d^{N+1}) or $N - 1$ (d^{N-1}) electrons. Depending on those fluctuations we find different effective spin interactions with the metal. If E_F is closer to the $E(d^N)$ level, the system develops an antiferromagnetic (AF) interaction, with $N \rightarrow N - 1$ fluctuations; when E_F approaches the $E(d^{N+1})$ level, the effective spin interaction is ferromagnetic (FM) and $N \rightarrow N + 1$. Things are different when $E(d^6) > E_F > E(d^5)$, which shows in both cases, for E_F close to either $E(d^6)$ or $E(d^5)$ an AF interaction; this is due to the electron–hole symmetry of the system between the cases $N < 5$ and $N > 5$. For example, the case $E(d^2) > E_F > E(d^1)$ is the symmetric of $E(d^8) < E_F < E(d^9)$, the system being in both cases AF if E_F is closer to either $E(d^8)$ or $E(d^2)$; for that AF interaction, we find the following fluctuations $d^2(S = 1) \rightarrow d^1(S = \frac{1}{2})$ or $d^8(S = 1) \rightarrow d^9(S = \frac{1}{2})$, and a similar effective spin interaction, $(J_\alpha/S)\hat{S} \cdot \hat{s}_\alpha$, Eq. (2.24), with $S = 1$.

Notice an interesting similarity between the conventional d-shell atom, with a fivefold orbital degeneracy, and the Anderson model. For the Anderson model we have two levels, $E(d^1)$ and $E(d^2)$, and the system shows an AF behavior for $E(d^1) < E_F < E(d^2)$ (see Fig. 2.1), while for the conventional d-shell atom we find the same AF behavior for $E(d^6) > E_F > E(d^5)$. In this last model we find, however, that for any other Fermi energy interval $E(d^{N+1}) > E_F > E(d^N)$, $N \neq 5$, the system shows either an AF or a ferromagnetic behavior depending on the position of E_F with respect to $E(d^N)$ or $E(d^{N+1})$.

Finally, for an AF exchange interaction we expect to have a Kondo resonance in the electronic density of states as shown in Fig. 2.4 [12]; we have also analyzed the Kondo temperature associated with that resonance and have found that it decreases exponentially with the total spin of the normal atomic state [13, 24].

References

1. K.G. Wilson, Rev. Mod. Phys. **47**(4), 773 (1975). DOI 10.1103/RevModPhys.47.773
2. A.C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, Cambridge, 1993). DOI 10.1017/CBO9780511470752
3. P.W. Anderson, Phys. Rev. **124**, 41 (1961). DOI 10.1103/PhysRev.124.41

4. J. Kondo, Prog. Theor. Phys. **32**(1), 37 (1964). DOI 10.1143/PTP.32.37
5. P. Nozières, A. Blandin, J. Phys. (France) **41**(3), 193 (1980). DOI 10.1051/jphys:01980004103019300
6. A. Georges, L. de' Medici, J. Mravlje, Annual Rev. Cond. Matter Phys. **4**(1), 137 (2013). DOI 10.1146/annurev-conmatphys-020911-125045
7. C.F. Hirjibehedin, C.Y. Lin, A.F. Otte, M. Ternes, C.P. Lutz, B.A. Jones, A.J. Heinrich, Science **317**(5842), 1199 (2007). DOI 10.1126/science.1146110
8. A.F. Otte, M. Ternes, S. Loth, C.P. Lutz, C.F. Hirjibehedin, A.J. Heinrich, Phys. Rev. Lett. **103**, 107203 (2009). DOI 10.1103/PhysRevLett.103.107203
9. A. Spinelli, M.P. Rebergen, A.F. Otte, J. Phys.: Cond. Matter **27**(24), 243203 (2015). DOI 10.1088/0953-8984/27/24/243203
10. J. Park, A.N. Pasupathy, J.I. Goldsmith, C. Chang, Y. Yaish, J.R. Petta, M. Rinkoski, J.P. Sethna, H.D. Abruna, P.L. McEuen, D.C. Ralph, Nature **417**(6890), 722 (2002). DOI 10.1038/nature00791
11. L.L. Hirst, Adv. Phys. **27**(2), 231 (1978). DOI 10.1080/00018737800101374
12. E.C. Goldberg, F. Flores, Phys. Rev. B **77**, 125121 (2008). DOI 10.1103/PhysRevB.77.125121
13. F. Flores, E.C. Goldberg, (2016). Submitted for publication
14. J. Hubbard, Proc. R. Soc. Lond. A **285**(1403), 542 (1965). DOI 10.1098/rspa.1965.0124
15. A. Martín-Rodero, F. Flores, M. Baldo, R. Pucci, Solid State Commun. **44**, 911 (1982). DOI 10.1016/0038-1098(82)90303-9
16. J. Ferrer, A. Martín-Rodero, F. Flores, Phys. Rev. B **36**, 6149 (1987). DOI 10.1103/PhysRevB.36.6149
17. J.R. Schrieffer, P.A. Wolff, Phys. Rev. **149**(2), 491 (1966). DOI 10.1103/PhysRev.149.491
18. J.C. Cuevas, E. Scheer, *Molecular electronics* (World Scientific, Singapore, 2010)
19. C. Zener, Phys. Rev. **81**(3), 440 (1951). DOI 10.1103/PhysRev.81.440
20. A. Yoshimori, Prog. Theor. Phys. **55**(1), 67 (1976). DOI 10.1143/PTP.55.67
21. E.C. Goldberg, F. Flores, J. Phys.: Cond. Matter **25**(22), 225001 (2013). DOI 10.1088/0953-8984/25/22/225001
22. M.D. Daybell, W.A. Steyert, Rev. Mod. Phys. **40**, 380 (1968). DOI 10.1103/RevModPhys.40.380
23. T. Jamneala, V. Madhavan, W. Chen, M.F. Crommie, Phys. Rev. B **61**, 9990 (2000). DOI 10.1103/PhysRevB.61.9990
24. J.R. Schrieffer, J. Appl. Phys. **38**(3), 1143 (1967). DOI 10.1063/1.1709517

Chapter 3
Structural Effects on Electronic Properties of Selected Materials

G.G.N. Angilella

Abstract Tuning structural properties, e.g., by means of external pressure or applied uniaxial strain, can significantly modify the electronic properties of several solid systems, thereby influencing, or even inducing, several electronic instabilities. We specifically analyze the effect of pressure or strain on several superconducting materials, including the high-T_c cuprates, some organic salts, and magnesium diboride, on the light alkalis, where pressure can even destroy metallicity, and graphene, whose remarkable electronic and transport properties can be tailored by applied strain.

Introduction

Most electronic transitions in the solid state clearly are directly influenced, when not induced, by specific structural properties of the material under consideration. Notable examples are superconductivity, which in BCS materials is made possible by the formation and condensation of Cooper pairs (electrons pairs bound by exchanged phonons), and some instances of insulator-to-metal transitions, and vice versa. Even in the high-T_c cuprate superconductors, and some related materials, where more

G.G.N. Angilella (✉)
Dipartimento di Fisica e Astronomia, Università di Catania, Via S. Sofia, 64,
95123 Catania, Italy
e-mail: giuseppe.angilella@ct.infn.it

G.G.N. Angilella
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
IMM-CNR, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
INFN, Sez. Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
Scuola Superiore di Catania, Università di Catania, Via Valdisavoia, 9,
95123 Catania, Italy

'exotic', purely electronic or nonexclusively phononic, pairing mechanisms have been proposed, structural effects are nonetheless important in determining the superconducting properties, such as the critical temperature T_c itself. This is particularly true in reduced-dimensional materials, e.g., the same cuprate superconductors, where the main electronic phenomena are believed to reside in the CuO_2 layers. Of course, this applies a fortiori to the electronic (e.g., band or transport) properties of graphene, an atomically thin layer of carbon atoms in the honeycomb lattice, when subjected to external strain or deformation.

In this chapter, a review will be presented of the main contributions by the group led by Professor Renato Pucci in Catania, and collaborators also elsewhere, to the understanding of the effects of structural deformation, as induced by external pressure or strain, on the electronic properties of several superconducting materials, of the 'simple' alkalis, and of graphene.

High-T_c Superconductors

High-T_c cuprate superconductors often crystallize in perovskite structures, and are characterized by CuO_2 planes alternated by charge 'reservoirs' of different elements, often containing rare earths. On doping, these provide the CuO_2 planes with the charge carriers (usually, holes) which eventually take part in superconductivity below the critical temperature T_c. That T_c is a characteristic, dome-shaped function of the doping concentration or, rather, of the hole content n, say, was established by several workers, including Zhang and Sato [1]. Further investigation envisaged a universal correlation between the critical temperature T_c and the superfluid density ρ_s, as proposed by Uemura et al. [2] (see also Refs. [3, 4]). Contributions along these directions within the group led by Pucci concerned the dependence of T_c on (i) pressure, (ii) the number of cuprate layers, and (iii) the symmetry of the order parameter in different material classes.

The critical temperature of both elemental as well as compound superconductors is a function of external pressure, which can be exploited as a tool to boost T_c [5, 6] (see also Chap. 4 by Schilling in this volume [7]). Fruitful discussions with Professor Jim Schilling at an international conference in Warsaw, Poland, in 1995, stimulated Pucci and collaborators to investigate the interplay among T_c, n, and pressure P [8]. Assuming $T_c = T_c(n, P)$, the effect of pressure in the cuprate superconductors was recognized to be possibly twofold: either direct or intrinsic, or indirect, through an implicit dependence of the hole content on pressure, $n = n(P)$. This led Pucci et al. [8] to extract phenomenologically the different contributions to

$$\frac{dT_c}{dP} = \frac{\partial T_c}{\partial P} + \frac{dn}{dP}\frac{\partial T_c}{\partial n}. \tag{3.1}$$

This study was later generalized by Pucci et al. to cuprates with several layers, by explicitly considering a possibly nonuniform distribution of hole content

among inequivalent layers [9]. There, the interlayer tunneling mechanism of high-T_c superconductivity was employed [10], which predicted a sharp anisotropy of the superconducting order parameter, regardless of the pairing symmetry thereof, which remained dependent on the in-plane pairing mechanism only [11].

More generally, Pucci et al. adopted the view proposed by Uemura et al. [2] that T_c should correlate universally with the superconducting coherence length ξ, thus embracing also other material classes, including the heavy Fermion superconductors, characterized by a different symmetry of the order parameter [12, 13]. Remarkably, Pucci et al. were able to derive an analytical relation between superconducting and normal state properties in non-s-wave superconductors, albeit for the simpler Cooper pairing problem, in the form of the surprisingly compact formula [14]

$$\xi^2 = \frac{\hbar^2}{2m^*\varepsilon_{\mathrm{F}}} \left[\frac{4}{3}\frac{1}{x^2} + \frac{\ell(\ell+1)}{1+x}\left(1 - \frac{x\ln x}{1+x}\right) \right]. \tag{3.2}$$

Here, m^* is the effective mass of the individual constituents of a Cooper pair in the normal state, which could be as high as $\approx 200 m_e$ for the heavy Fermion superconductors, ε_{F} is their Fermi energy, implicitly related to their density (hole content n, in the cuprates), $x = |\varepsilon|/2\varepsilon_{\mathrm{F}}$ is a dimensionless measure of the binding energy of a Cooper pair, $|\varepsilon|$, in units of the energy of its unbound constituents, $2\varepsilon_{\mathrm{F}}$, and is thereby a measure of the superconducting strength, as it merges into the superconducting gap in the fully many-body problem. Finally, $\ell = 0, 1, 2$ is an orbital quantum number, parametrizing the orbital symmetry of the superconducting gap as s-, p-, or d-wave, respectively. Equation (3.2) is in good qualitative accord with the experimental data for $k_{\mathrm{B}}T_c$ correlated with the characteristic energy $\varepsilon_c = \hbar^2/2m^*\xi^2$ within various material classes, with different symmetry of the order parameter [12–14].

Effect of Proximity to an Electronic Topological Transition (ETT)

Applied pressure or uniaxial strain can also serve as a tool to evidence the dependence of T_c or other superconducting characteristics on otherwise elusive details of the electronic structure of materials, especially in the case of low-dimensional ones. Again, this is the case of the high-T_c cuprates, these being characterized by a nearly two-dimensional electronic band structure, due to the presence of CuO_2 layers. As is generic of two-dimensional systems, the electronic density of states (DOS) is expected to exhibit logarithmic divergences (Van Hove singularities) as a function of energy [15]. This is a consequence of Morse theory [16], applied to the band dispersion relation $\varepsilon_{\mathbf{k}}$ over the first Brillouin zone (1BZ). Since this is a periodic, smooth function of quasimomentum $\mathbf{k}$, it possesses critical points, and these correspond to Electronic Topological Transitions (ETT), i.e., possible values of the Fermi energy ε_{F} at which the 'Fermi line', implicitly defined by $\varepsilon_{\mathbf{k}} = \varepsilon_{\mathrm{F}}$, changes

topology in momentum space. These are exactly the energy values at which the two-dimensional DOS presents logarithmic singularities. Any proximity to these values is therefore expected to induce remarkable behaviors either in the normal or in the superconducting state of these materials [17].

ETTs were originally studied by I.M. Lifshitz [18], in connection with the electronic properties of three-dimensional (3D) metals (see Refs. [19–21] for reviews). In contrast with the 3D case, Pucci and co-workers [22] found that T_c depends non-monotonically on the Fermi energy, with a maximum T_c^{max} close to an ETT. Furthermore, this maximum monotonically increases as a function of the ratio $r = t'/t$ between the next-nearest to nearest neighbor hopping parameters in a tight-binding dispersion relation [22], in a characteristic way, also depending on the symmetry of the order parameter (either s- or d-wave). Such a behavior was in good qualitative agreement with the phenomenological trend examined by Pavarini et al. [23], by collecting both experimental and theoretical (DFT) results for several cuprates. It was also established that the proximity to an ETT influenced the superconducting fluctuations of these layered superconductors [22, 24], again in agreement with available data for the excess Hall conductivity for several cuprates and cuprate superlattices [25–28].

One effective way to control the proximity to the ETT in the cuprates was recognized as epitaxial strain, as is, e.g., induced by epitaxially growing a thin film of $La_{2-x}Sr_xCuO_4$ in slightly mismatched substrates, thereby realizing either compressive or tensile strains [29]. The strain effect is twofold, as it directly modifies the in-plane electronic band structure, and thereby the ratio $r = t'/t$, by means of a modification of the in-plane crystal structure, and as it indirectly modifies the hole content, by means of a deformation of the crystal structure in the direction perpendicular to applied strain. Another achievement of the group led by Professor Pucci was to show that this again correlates qualitatively well both with the non-monotonic, dome-shaped dependence of T_c on hole content, and with the monotonic dependence of its maximum, T_c^{max} say, on the ratio $r = t'/t$ [30].

Organic Superconductors

Another class of quasi-two-dimensional materials characterized by the proximity to ETTs are the BEDT-TTF–based organic salts, where BEDT-TTF is the acronym for bis(ethylenedithia)tetrathiafulvalene. The κ-phase of the $(BEDT-TTF)_2$-X organic salts differ from the Bechgaard salts in that the arrangement of the constituent organic molecules is planar, rather than linear, with a number of planes stacked on top of each other, and the anions X, where $X = Cu(NCS)_2$, $Cu[N(CN)_2]Br$, or $Cu[N(CN)_2]Cl$, acting as fillers and charge reservoirs between the planes. Such a structure gives rise to a markedly anisotropic conductivity, similar to what happens in the high-T_c cuprates [31]. However, due to the inequivalent orientation of the BEDT-TTF dimers within the planes, next-nearest hopping is here restricted along the (110) direction only. Thus, it turns out that a triangular lattice model is more appropriate to describe such

compounds. Both in the cuprates and in the κ-BEDT-TTF organic salts, however, a nonzero value of the hopping ratio $r = t'/t$ slightly modulates the shape of the Fermi line. Changes in the Fermi energy, as well as changes in the band parameters and thus in r, can be induced by hydrostatic pressure, as well as by anisotropic strain, or by changing the doping level.

In the case of BEDT-TTF-based salts, the Fermi line changes topology twice: once from electron-like to hole-like contours, and the second time from a singly-connected hole-like contour to a doubly connected one, still retaining its hole-like character. However, evidence for a non-monotonic dependence of T_c on uniaxial strain in κ-(BEDT-TTF)$_2$-X is still lacking, below $\sim$10 kbar, thus suggesting that higher pressures or anisotropic strains are required, in order to demonstrate the relevance of ETTs in such quasi-2D organic salts [32] (see also [33]).

Magnesium Diboride (MgB$_2$)

The concept of a proximity to an ETT has also been employed by Pucci and co-workers to investigate the dependence of T_c on doping in a three-dimensional superconductor, MgB$_2$, whose band structure is usually described in terms of the interplay between a σ and π subband. Indeed, in Al$_{1-x}$Mg$_x$B$_2$, at $x = 0.66$, the overall monotonic dependence of T_c on the Mg content x displays a pronounced kink, that has been related to a 3D–2D crossover of the Fermi surface associated with the boron σ subband [34]. This tuning enhances the critical temperature by a shape resonance [35–38] in the boron superlattice [38–41] that is analogous to the Feshbach resonance in ultracold atoms. Within the two-band model of superconductivity, Pucci and co-workers [42] studied the effect of the proximity to a 3D–2D crossover (or ETT) on the doping dependence of the critical temperature and of the isotope effect. The proximity to the ETT correctly takes into account, both quantitatively and qualitatively, for the enhancement of T_c as a consequence of a quantum interference effect between the two electronic bands characterizing the diborides.

Alkali Metals Under Extreme Conditions

Because of their relatively simple electronic structure and relatively high electron density, light alkali metals in normal conditions, including Li, Na, and K, have long been considered to have a 'simple' metallic behavior [43]. The free-electron model was also believed to work even better at higher pressure and electron density. Such an assumption was first questioned by Siringo et al. [44], who, already in the title of their work, explicitly posed the issue whether the light alkali metals are still metals under high pressure. Siringo et al. [44] considered an extended Hubbard model on a bcc lattice, characterized by an on-site electron repulsion U and an intersite electron interaction V. The main result of their work was that such a model electron system

should undergo a metal-to-insulator transition with increasing electron density. Such a transition occurred at a critical value of the adimensional Wigner–Seitz radius of $r_s = 2.6$, corresponding to a pressure $p = 100\,\text{GPa}$ for Na. Later works by the same authors even predicted reentrant metallicity at yet higher pressure, corresponding to oscillations between a symmetric (metallic) phase and a low-symmetry (dimerized and insulating) phase, as a consequence of Friedel oscillations in the electron pair potential [45, 46].

The original result by Siringo et al. [44] was followed soon after by Density Functional Theory (DFT) studies on dense lithium by Neaton and Ashcroft [47], who predicted a structural instability of bcc Li toward a more complex Cmca structure at high pressure. In this structure, each Li ion is coordinated by another ion only, and the pairing thereof can give rise to a metal-to-insulator transition at some $r_s < 2.1$. According to Neaton and Ashcroft [47], the initial distortion is due to the Jahn–Teller mechanism, while further distortion is due to a balance between the Peierls instability and exchange effects.

The above theoretical findings immediately kindled much experimental interest [48–50], which confirmed the tendency of the light alkali metals to lower their symmetry under high pressure. Low-symmetry phases have been indeed observed for Li [49] and Na [51, 52]. In particular for Li, Hanfland et al. [49] established experimentally and theoretically that a cI16 structure stabilizes at $p \simeq 40\,\text{GPa}$ and $T \simeq 180\,\text{K}$. The existence of a broken symmetry phase was confirmed by the observation of Goncharov et al. [53] of a Raman vibrational band in Li above 70–120 GPa, the thus confirming the theoretical prediction by Matsuoka et al. [54] and by Struzhkin et al. [55] of the existence of two further phases of Li, *viz.* Li-VI and Li-VII, which are stable at $T = 25\,\text{K}$ and $p = 69\,\text{GPa}$ and $p = 86\,\text{GPa}$, respectively. Furthermore, Matsuoka and Shimizu [56], by performing electrical measurements up to 105 GPa, have inequivocally found that a metal-to-insulator transition occurs in Li near 80 GPa at $T < 50\,\text{K}$. Preliminary calculations [57] seemed to indicate that both Li and Na show a tendency toward the formation of atomic pairs, and a dimerized oC8 structure was predicted to be the most stable phase above 165 GPa for Li, and above 220 GPa for Na. According to more recent and compelling calculations, several proposals have been made for the identification of the structures of the Li-VI and Li-VII phases, including the oC88 and oC40 structures [58], the c2-24 and Aba2-24 structures [59], and more recently an Aba2-40 structure [60, 61].

Analogies and differences between Li and Na under pressure have been described in detail by Rousseau et al. [62]. Like Li, Na crystallizes in the bcc structure at ambient conditions, and undergoes a structural transition toward the fcc phase at 65 GPa [63]. Increasing pressure up to $p = 103\,\text{GPa}$, Na takes the Na-III phase [52, 64], which has a cI16 symmetry, which is similar to the symmetry adopted by compressed Li. However, above $p = 117\,\text{GPa}$, the Na-III phase transforms into the Na-IV phase [64, 65], which is characterized by an oP8 symmetry. Yet two more phases of Na have been discovered, *viz.* Na-V above $p = 125\,\text{GPa}$ (Refs. [64, 65]), and Na-VI above $p = 200\,\text{GPa}$ (Ref. [66]). This latter phase presents a simple double-hexagonal close-packed (dhcp) structure, with hP4 symmetry, and is typically transparent [66]. The hP4 structure has been confirmed also theoretically by Gatti et al. [67].

The physical reason for dimerization is an open question. Ab initio electronic structure calculations [68] indicated a tendency toward distance alternation, due to a sizeable overlap of $p\pi$ orbitals in the interstitial regions, and to electron localization between ions [69, 70]. On the other hand, it has been proposed [57, 71] that the increase of s-p hybridization could give rise to a low coordination number. However, even the fully dimerized phase is far from any standard covalent solid: the electron density is uniformly spread and almost constant, while the first and second neighbors distances are comparable. Such a dimerized phase is better described as a charge density wave in a high density metal rather than a molecular solid. Such a low-symmetry phase in distorted cubic-based structures, as those observed experimentally in the compressed alkalis, has been connected by Angilella et al. [45, 46] to the presence of Friedel oscillations in the electronic pair potential. These can ultimately be traced back to the existence of a sharp Fermi surface, and should therefore be generic to most metals.

The deceivingly 'simple' light alkali metals are a continuous source of experimental surprises and theoretical challenges. Recently, earlier predictions of superconductivity in compressed lithium [72] have been confirmed experimentally [54, 55, 73, 74]. At variance with Li, the superconducting critical temperature T_c of Na is predicted to be rather small ($T_c < 1$ K) also in fcc phase [75, 76]. If constrained in structures with reduced dimensionality (thin layers or wires), both lithium and sodium have been predicted to exhibit other electronic instabilities, such as a ferromagnetic one [77, 78]. A closer consideration of the phase diagram of both lithium and sodium at high pressure indicates that these light alkali metals should rather prefer the liquid state, with lithium being thus the elemental metal with the lowest melting point [58, 64, 79–82]. Several theoretical works have also been performed both on Li and on Na [83–88]. Further ab initio calculations [61, 89–91] at yet higher pressure indicate more complicated crystal structures, a tendency toward a metal-to-semiconducting transition, and the increasingly relevant role of core orbitals in establishing the electronic behavior of the compressed alkalis. However, both for lithium [56] and sodium [66] a metal-to-insulator transition at high pressure has now been experimentally well established. (See also [92] for a recent review.)

Graphene Under Strain

Graphene is a single, atomically thick, sheet of carbon atoms regularly arranged in the honeycomb lattice, which has been isolated quite recently in the laboratory [93]. Because of its truly remarkable electronic, structural, and transport properties at the nanoscale, and its wide-ranging potential applications, the discovery of graphene has kindled a tremendous outburst of both experimental and theoretical investigation [94]. It was clear at the outset (or even well before its discovery [95]) that the electronic properties of graphene are directly related with its structural properties [96], i.e., its essentially two-dimensional character, and the bipartite nature of the honeycomb lattice. The most prominent feature is, notoriously, the occurrence of two inequivalent

Dirac points, at **K** and **K**′ say, in the band dispersion, where the dispersion relation is linear, thus making elementary excitations massless. It was immediately realized, therefore, that a *change* in the structural properties of graphene would immediately induce potentially relevant modifications in the *electronic* properties thereof.

Pucci and co-workers then started investigating[1] modifications in the local electronic properties of graphene, as those being induced by the insertion of vacancies or impurities, such as hydrogen [97, 98] (see also Chap. 14 in the present volume [99]). In particular, Pucci et al. [98] evaluated the local density of states (LDOS) as a function of energy on an impurity located at high symmetry points in the direct lattice and on its neighboring sites, as well as in reciprocal space, at an energy corresponding to a bound state. The study was relevant to interpret the results of Fourier-transformed scanning tunneling spectroscopy, as they show which states mostly contribute to impurity-induced variations in the LDOS.[2] It was also speculated [101] that an analysis of the LDOS around an isolated impurity in graphene could help discerning among the several symmetries of a superconducting order parameter compatible with the point symmetries of the honeycomb lattice, following the proposal that graphene could sustain intrinsic superconductivity [102].

Pucci's group then undertook an almost systematic study of various electronic, transport, and optical properties of graphene under strain. Indeed, a deformation of the lattice through the application of uniaxial strain or hydrostatic pressure is expected to produce modifications also in the electronic structure of graphene. Recently, it has been proposed that nanodevices based on graphene could be engineered on the basis of the expected strain-induced modifications of the deformed graphene sheet (origami electronics) [103]. This is made possible by the exceptional mechanical properties of graphene, as is the case of other carbon compounds. For instance, despite its reduced dimensionality, graphene is characterized by a sizeable tensile strength and stiffness [104], with graphene sheets being capable to sustain elastic deformations as large as ≈20% [105–109]. Larger strains would then induce a semimetal-to-semiconductor transition, with the opening of an energy gap [110–113].

Uniaxial strain can be included in a standard, noninteracting model Hamiltonian at the tight-binding level, i.e., through the introduction of strain-dependent hopping parameters [111]. In particular, the strain-induced modification of the band structure at a fixed chemical potential may again result in an ETT [18, 20, 114], as seen above for the cuprates and other (especially) low-dimensional material classes. In graphene, one of the main consequences of applied of strain is that of moving the Dirac points, i.e., the points where the band dispersion relations vanish linearly, away from the points of highest symmetry in the first Brillouin zone, and of deforming their low-energy conical approximation. Moreover, for a sufficiently large strain modulus and a for a generic strain direction, two inequivalent Dirac points may merge, thus resulting

---

[1]Well before the Nobel Prize in Physics was awarded to Novoselov and Geim in 2010!

[2]It is amusing to recollect that the authors of Ref. [98] anticipated that the psychedelic LDOS maps in momentum space (in false colors, of course!) might make them useful as tiles for a kitchen wall, say. And indeed three samples of such a tile were crafted, after that anticipation was rewarded by the Physical Review B, which selected that particular figure as the online cover of that issue! [100].

in the opening of a band gap (at strains larger than the critical one) along one specific direction across the degenerate Dirac point. This results in a sublinear density of states exactly at the transition, which in turn gives rise to an unusual magnetic field dependence of the Landau levels [115]. This may also be described as a quantum phase transition, from a semimetal to a semiconductor state, of purely topological origin [116], characterized by low-energy massless quasiparticles developing a finite mass only along a given direction.

Interestingly, besides the cuprates, similar effects have been predicted also for other low-dimensional systems, and that their overall features are generic with respect to their detailed crystal structure. In particular, a similar discussion applies as well to cold atoms in two-dimensional optical lattices [115, 117], which have been proposed to simulate the behavior of Dirac fermions [118].

In Ref. [119], Pellegrino, the present author, and Pucci evaluated the longitudinal optical conductivity of strained graphene as a function of frequency. Their main results were that (a) logarithmic peaks appear in the optical conductivity at sufficiently high frequency, and can be related to the ETTs in the electronic spectrum under strain, and depending on the strain direction; (b) the relative weight of the peaks in general depends on the strain direction and field orientation, and contributes to the generally anisotropic pattern of the optical conductivity as a function of field orientation; (c) the opening of a band gap, where allowed, is signaled by a vanishing optical conductivity. Thus, an experimental study of the optical conductivity in the visible range of frequencies as a function of strain modulus and direction, as well as of field orientation, should enable one to identify the occurrence of the three distinct ETTs predicted for graphene [120].

This study was complemented by an analysis of the strain effect on the Drude weight of graphene by the same authors [121]. They found a nonmonotonic strain dependence of the Drude weight, related to the longitudinal polarization of graphene. This can again be described as an effect of the proximity to several strain-induced ETTs. Given the connection between the Drude weight and the long-wavelength plasmon frequency [122, 123], such a dependence could be evidenced by infrared spectroscopy [124].

More generally, Pellegrino, the present author, and Pucci [123] evaluated the dynamical polarization under strain. This enabled to derive the dependence of the plasmon dispersion relation of graphene on applied uniaxial strain. Besides electron correlation at the level of the random phase approximation (RPA), these authors also included local field effects (LFE) specific for the honeycomb lattice. As a consequence of the two-band character of the electronic band structure, two distinct plasmon branches were found. Besides recovering the square root behavior of the low-energy branch, Pellegrino et al. [123] found a nonmonotonic dependence of the strain-induced modification of its stiffness, as a function of the wavevector orientation with respect to applied strain.

In Ref. [125], Pellegrino, the present author, and Pucci studied a generic linear response function for a deformed graphene sheet, relevant instances being the density–density and current–current correlation functions. This enabled to derive the strain dependence of the plasmon dispersion relation and of the optical conductivity.

Specifically, Pellegrino et al. [125] found that the prefactor in the $\sqrt{q}$-dependence of the plasmon frequency develops an anisotropic character, with maxima occurring when the wavevector is orthogonal to the direction of applied strain. Moreover, they derived a strain-induced anisotropic enhancement of the deviations from linearity of the recently predicted transverse plasmon [126], which should facilitate its experimental detection in suitably strained graphene samples. Also, Ref. [125] by Pellegrino et al. contained results concerning the strain dependences of the magnetic and electric susceptibilities, showing that strain enhances the response of strained graphene to an applied electric field, while suppressing the response to a magnetic field.

Refs. [127, 128] by Pellegrino et al. considered the transport properties across strain-induced nonuniform velocity profiles in graphene, and the possibility that resonant modes may be obtained in strain-induced graphene superlattices. A linear dispersion relation around the Dirac points makes possible the realization of Klein tunneling across potential barriers in graphene [129–133]. Single and multiple barriers can be induced by strain [134, 135], but the main effect of strain has usually been considered to be that of shifting the position of the Dirac points in reciprocal space. However, it has been demonstrated that a nonuniform space variation of the underlying gate potential would result in a modulation of the Fermi velocity [134, 136, 137]. Indeed, Pellegrino et al. [127] showed that both effects are of the same order on the applied strain intensity, and should therefore be considered on the same ground, when studying the transport properties of strained graphene. They therefore explicitly considered not only the strain-induced displacement of the Dirac points in reciprocal space, but also a strain-induced deformation of the Dirac cones, resulting in a strain-dependent anisotropic Fermi velocity. Their results could be generalized to the problem of transport through a tunneling structure, characterized by a *nonuniform* variation of both the Fermi velocity and of the gate potential, as can e.g., be brought about by a continuous deformation or applied uniaxial strain.

As a possible application, Pellegrino et al. [128] then suggested that a strain-induced graphene superlattice may serve as a mode filter for transport in graphene. To this aim, they studied the condition for the occurrence of resonant quasiparticle states for tunneling across such a strain-induced superlattice in graphene. They also derived the band-like spectrum of bound states, as a function of conserved energy and transverse momentum.

More recent developments on graphene can be found in Chap. 12 of the present volume [138].

Conclusions

Summarizing, a brief review has been presented of some of the most recent research interests and results by Professor Renato Pucci and collaborators. These ranged from superconductivity in several material classes, especially the high-T_c cuprates, to the metal-to-insulator transition in the light alkalis under pressure, to the electronic

and transport properties of strained graphene. Some keywords and *leitmotifs* among these wide-ranging subjects are certainly the occurrence of electronic instabilities of various kinds, such as superconducting pairing or the metal-to-insulator transition, the relevance of low dimensionality on the electronic spectrum and density of states of these systems, the proximity to an electronic topological transition being a recurring feature, and of course the possibility of tuning or even inducing any of these effects by changing the structural properties of the systems under study, as can be brought about by external pressure or applied uniaxial strain.

References

1. H. Zhang, H. Sato, Phys. Rev. Lett. **70**, 1697 (1993). DOI 10.1103/PhysRevLett.70.1697
2. Y.J. Uemura, L.P. Le, G.M. Luke, B.J. Sternlieb, W.D. Wu, J.H. Brewer, T.M. Riseman, C.L. Seaman, M.B. Maple, M. Ishikawa, D.G. Hinks, J.D. Jorgensen, G. Saito, H. Yamochi, Phys. Rev. Lett. **66**, 2665 (1991). DOI 10.1103/PhysRevLett.66.2665
3. C.C. Homes, S.V. Dordevic, M. Strongin, D.A. Bonn, R. Liang, W.N. Hardy, S. Komiya, Y. Ando, G. Yu, N. Kaneko, X. Zhao, M. Greven, D.N. Basov, T. Timusk, Nature **430**(6999), 539 (2004). DOI 10.1038/nature02673
4. F. Pistolesi, G.C. Strinati, Phys. Rev. B **49**, 6356 (1994). DOI 10.1103/PhysRevB.49.6356
5. R.J. Wijngaarden, E.N. van Eenige, J.J. Scholtz, H.K. Hemmes, R. Griessen, in *Molecular systems under high pressure: Proceedings of the II Archimedes workshop on Molecular solids under high pressure, Catania, Italy, May 28–31, 1990*, ed. by R. Pucci, G. Piccitto (North-Holland, Amsterdam, 1991), pp. 157–180
6. R.J. Wijngaarden, D.T. Jover, R. Griessen, Physica B **265**(1-4), 128 (1999). DOI 10.1016/S0921-4526(98)01342-8
7. J.S. Schilling, in *Correlations in condensed matter under extreme conditions*, ed. by G.G.N. Angilella, A. La Magna (Springer, Berlin, 2017), chap. 4, p. 47. (This volume.)
8. G.G.N. Angilella, R. Pucci, F. Siringo, Phys. Rev. B **54**, 15471 (1996). DOI 10.1103/PhysRevB.54.15471
9. G.G.N. Angilella, R. Pucci, Physica B **265**(1-4), 136 (1999). DOI 10.1016/S0921-4526(98)01343-X
10. S. Chakravarty, A. Sudbø, P.W. Anderson, S. Strong, Science **261**(5119), 337 (1993). DOI 10.1126/science.261.5119.337
11. G.G.N. Angilella, R. Pucci, F. Siringo, A. Sudbø, Phys. Rev. B **59**, 1339 (1999). DOI 10.1103/PhysRevB.59.1339
12. G.G.N. Angilella, N.H. March, R. Pucci, Phys. Rev. B **62**, 13919 (2000). DOI 10.1103/PhysRevB.62.13919
13. G.G.N. Angilella, N.H. March, R. Pucci, Phys. Rev. B **65**(9), 092509 (2002). DOI 10.1103/PhysRevB.65.092509
14. G.G.N. Angilella, F.E. Leys, N.H. March, R. Pucci, Phys. Lett. A **322**, 375 (2004). DOI 10.1016/j.physleta.2003.12.066
15. W. Jones, N.H. March, *Theoretical Solid-State Physics. Perfect Lattices in Equilibrium*, vol. 1 (Dover, New York, 1986)

16. J. Milnor, *Morse theory*, *Ann. Math. Studies*, vol. 51 (Princeton University Press, Princeton, 1963)
17. R.S. Markiewicz, J. Phys. Chem. Solids **58**(8), 1179 (1997). DOI 10.1016/S0022-3697(97)00025-5
18. I.M. Lifshitz, Sov. Phys. JETP **11**, 1130 (1960). Zh. Eksp. Teor. Fiz. **38**, 1569 (1960)
19. A.A. Varlamov, V.S. Egorov, A.V. Pantsulaya, Adv. Phys. **38**, 469 (1989). DOI 10.1080/00018738900101132
20. Ya. M. Blanter, M.I. Kaganov, A.V. Pantsulaya, A.A. Varlamov, Phys. Rep. **245**(4), 159 (1994). DOI 10.1016/0370-1573(94)90103-1
21. E. Bruno, B. Ginatempo, E.S. Giuliano, A.V. Ruban, Y.K. Velikov, Phys. Rep. **249**(6), 353 (1994). DOI 10.1016/0370-1573(94)90056-6
22. G.G.N. Angilella, E. Piegari, A.A. Varlamov, Phys. Rev. B **66**(1), 014501 (2002). DOI 10.1103/PhysRevB.66.014501
23. E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, O.K. Andersen, Phys. Rev. Lett. **87**(4), 047003 (2001). DOI 10.1103/PhysRevLett.87.047003
24. G.G.N. Angilella, R. Pucci, A.A. Varlamov, F. Onufrieva, Phys. Rev. B **67**, 134525 (2003). DOI 10.1103/PhysRevB.67.134525
25. W. Lang, W. Kula, R. Sobolewski, Physica B **194-196**, 1643 (1994). DOI 10.1016/0921-4526(94)91321-8
26. W. Lang, G. Heine, P. Schwab, X.Z. Wang, D. Bäuerle, Phys. Rev. B **49**(7), 4209 (1994). DOI 10.1103/PhysRevB.49.4209
27. W. Lang, G. Heine, W. Kula, R. Sobolewski, Phys. Rev. B **51**(14), 9180 (1995). DOI 10.1103/PhysRevB.51.9180
28. L.M. Wang, H.C. Yang, H.E. Horng, Phys. Rev. B **59**(21), 14031 (1999). DOI 10.1103/PhysRevB.59.14031
29. J. Locquet, J. Perret, J. Fompeyrine, E. Machler, J.W. Seo, G. Van Tendeloo, Nature (London) **394**, 453 (1998). DOI 10.1038/28810
30. G.G.N. Angilella, G. Balestrino, P. Cermelli, P. Podio-Guidugli, A.A. Varlamov, Eur. Phys. J. B **26**, 67 (2002). DOI 10.1140/epjb/e20020067
31. R.H. McKenzie, Science **278**(5339), 820 (1997). DOI 10.1126/science.278.5339.820
32. G.G.N. Angilella, E. Piegari, R. Pucci, A.A. Varlamov, in *Frontiers of high pressure research II: Application of high pressure to low-dimensional novel electronic materials*, *NATO Science Series*, vol. 48, ed. by H.D. Hochheimer, B. Kuchta, P.K. Dorhout, J.L. Yarger (Kluwer, Dordrecht, 2001), *NATO Science Series*, vol. 48. DOI 10.1007/978-94-010-0520-3_27
33. M.O. Goerbig, J.N. Fuchs, G. Montambaux, F. Piéchon, Phys. Rev. B **78**(4), 045415 (2008). DOI 10.1103/PhysRevB.78.045415
34. A. Bianconi, S. Agrestini, D.D. Castro, G. Campi, G. Zangari, N.L. Saini, A. Saccone, S.D. Negri, M. Giovannini, G. Profeta, A. Continenza, G. Satta, S. Massidda, A. Cassetta, A. Pifferi, M. Colapietro, Phys. Rev. B **65**(17), 174515 (2002). DOI 10.1103/PhysRevB.65.174515
35. A. Bianconi, D.D. Castro, S. Agrestini, G. Campi, N.L. Saini, A. Saccone, S.D. Negri, M. Giovannini, J. Phys.: Condens. Matter **13**(33), 7383 (2001). DOI 10.1088/0953-8984/13/33/318
36. J.M. Blatt, C.J. Thompson, Phys. Rev. Lett. **10**(8), 332 (1963). DOI 10.1103/PhysRevLett.10.332
37. C.J. Thompson, J.M. Blatt, Phys. Lett. **5**(1), 6 (1963). DOI 10.1016/S0375-9601(63)80003-1
38. A. Perali, A. Bianconi, A. Lanzara, N.L. Saini, Solid State Commun. **100**(3), 181 (1996). DOI 10.1016/0038-1098(96)00373-0
39. A. Bianconi, Solid State Commun. **89**(11), 933 (1994). DOI 10.1016/0038-1098(94)90354-9
40. A. Valletta, A. Bianconi, A. Perali, N.L. Saini, Z. Phys. B **104**(4), 707 (1997). DOI 10.1007/s002570050513
41. A. Bianconi, A. Valletta, A. Perali, N.L. Saini, Physica C **296**(3-4), 269 (1998). DOI 10.1016/S0921-4534(97)01825-X

42. G.G.N. Angilella, A. Bianconi, R. Pucci, J. Supercond. **18**(5), 619 (2006). DOI 10.1007/s10948-005-0049-3

43. N.W. Ashcroft, N.D. Mermin, *Solid State Physics* (Saunders College Publ., Fort Worth, 1976)

44. F. Siringo, R. Pucci, G.G.N. Angilella, High Press. Research **15**(4), 255 (1997). DOI 10.1080/08957959708244246. Also available as preprint arXiv:cond-mat/9512011, and reprinted in this volume, p. 261

45. G.G.N. Angilella, F. Siringo, R. Pucci, Eur. Phys. J. B **32**, 323 (2003). DOI 10.1140/epjb/e2003-00105-8

46. G.G.N. Angilella, F. Siringo, R. Pucci, High Press. Res. **23**(3), 243 (2003). DOI 10.1080/0895795032000102414

47. J.B. Neaton, N.W. Ashcroft, Nature (London) **400**, 141 (1999). DOI 10.1038/22067

48. M. Hanfland, I. Loa, K. Syassen, U. Schwarz, K. Takemura, Sol. State Commun. **112**(3), 123 (1999). DOI 10.1016/S0038-1098(99)00322-1

49. M. Hanfland, K. Syassen, N.E. Christensen, D.L. Novikov, Nature (London) **408**, 174 (2000). DOI 10.1038/35041515

50. V.E. Fortov, V.V. Yakushev, K.L. Kagan, I.V. Lomonosov, V.I. Postnov, T.I. Yakusheva, JETP Letters **70**, 628 (1999). Pis'ma Zh. Eksp. Teor. Fiz. **70**, 620 (1999)

51. J.B. Neaton, N.W. Ashcroft, Phys. Rev. Lett. **86**(13), 2830 (2001). DOI 10.1103/PhysRevLett.86.2830

52. M.I. McMahon, E. Gregoryanz, L.F. Lundegaard, I. Loa, C. Guillaume, R.J. Nelmes, A.K. Kleppe, M. Amboage, H. Wilhelm, A.P. Jephcoat, Proc. Nat. Acad. Sci. **104**(44), 17297 (2007). DOI 10.1073/pnas.0709309104

53. A.F. Goncharov, V.V. Struzhkin, H.k. Mao, R.J. Hemley, Phys. Rev. B **71**(18), 184114 (2005). DOI 10.1103/PhysRevB.71.184114

54. T. Matsuoka, S. Onoda, M. Kaneshige, Y. Nakamoto, K. Shimizu, T. Kagayama, Y. Ohishi, J. Phys.: Conf. Series **121**(5), 052003 (2008). DOI 10.1088/1742-6596/121/5/052003

55. V.V. Struzhkin, M.I. Eremets, W. Gan, H. Mao, R.J. Hemley, Science **298**(5596), 1213 (2002). DOI 10.1126/science.1078535

56. T. Matsuoka, K. Shimizu, Nature **458**, 186 (2009). DOI 10.1038/nature07827

57. N.E. Christensen, D.L. Novikov, Sol. State Commun. **119**(8-9), 477 (2001). DOI 10.1016/S0038-1098(01)00277-0

58. C.L. Guillaume, E. Gregoryanz, O. Degtyareva, M.I. McMahon, M. Hanfland, S. Evans, M. Guthrie, S.V. Sinogeikin, H. Mao, Nat. Phys. **7**, 211 (2011). DOI 10.1038/nphys1864

59. Y. Yao, J.S. Tse, D.D. Klug, Phys. Rev. Lett. **102**(11), 115503 (2009). DOI 10.1103/PhysRevLett.102.115503

60. M. Marqués, M.I. McMahon, E. Gregoryanz, M. Hanfland, C.L. Guillaume, C.J. Pickard, G.J. Ackland, R.J. Nelmes, Phys. Rev. Lett. **106**(9), 095502 (2011). DOI 10.1103/PhysRevLett.106.095502

61. J. Lv, Y. Wang, L. Zhu, Y. Ma, Phys. Rev. Lett. **106**(1), 015503 (2011). DOI 10.1103/PhysRevLett.106.015503

62. B.R.Y. Xie, Y. Ma, A. Bergara, Eur. Phys. J. B **81**, 1 (2011). DOI 10.1140/epjb/e2011-10972-9

63. M. Hanfland, I. Loa, K. Syassen, Phys. Rev. B **65**(18), 184109 (2002). DOI 10.1103/PhysRevB.65.184109

64. E. Gregoryanz, L.F. Lundegaard, M.I. McMahon, C. Guillaume, R.J. Nelmes, M. Mezouar, Science **320**, 1054 (2008). DOI 10.1126/science.1155715

65. L.F. Lundegaard, E. Gregoryanz, M.I. McMahon, C. Guillaume, I. Loa, R.J. Nelmes, Phys. Rev. B **79**(6), 064105 (2009). DOI 10.1103/PhysRevB.79.064105

66. Y. Ma, M. Eremets, A.R. Oganov, Y. Xie, I. Trojan, S. Medvedev, A.O. Lyakhov, M. Valle, V. Prakapenka, Nature **458**, 182 (2009). DOI 10.1038/nature07786

67. M. Gatti, I.V. Tokatly, A. Rubio, Phys. Rev. Lett. **104**(21), 216404 (2010). DOI 10.1103/PhysRevLett.104.216404

68. R. Rousseau, D. Marx, Chem. Eur. J. **6**(16), 2982 (2000). DOI 10.1002/1521-3765(20000818)6:16<2982::AID-CHEM2982>3.0.CO;2-S

69. B. Rousseau, N.W. Ashcroft, Phys. Rev. Lett. **101**, 046407 (2008). DOI 10.1103/PhysRevLett. 101.046407
70. A.M. Pendás, M.A. Blanco, A. Costales, P.M. Sánchez, V. Luaña, Phys. Rev. Lett. **83**(10), 1930 (1999). DOI 10.1103/PhysRevLett.83.1930
71. A. Bergara, J.B. Neaton, N.W. Ashcroft, Phys. Rev. B **62**(12), 8494 (2000). DOI 10.1103/PhysRevB.62.8494
72. N.E. Christensen, D.L. Novikov, Phys. Rev. Lett. **86**(9), 1861 (2001). DOI 10.1103/PhysRevLett.86.1861
73. K. Shimizu, H. Ishikawa, D. Takao, T. Yagi, K. Amaya, Nature (London) **419**, 597 (2002). DOI 10.1038/nature01098
74. S. Deemyad, J.S. Schilling, Phys. Rev. Lett. **91**(16), 167001 (2003). DOI 10.1103/PhysRevLett.91.167001
75. N.E. Christensen, D.L. Novikov, Phys. Rev. B **73**(22), 224508 (2006). DOI 10.1103/PhysRevB.73.224508
76. L. Shi, D.A. Papaconstantopoulos, Phys. Rev. B **73**(18), 184516 (2006). DOI 10.1103/PhysRevB.73.184516
77. A. Bergara, J.B. Neaton, N.W. Ashcroft, Int. J. Quantum Chem. **91**(2), 239 (2003). DOI 10.1002/qua.10414
78. N.W. Ashcroft, in *Frontiers of High Pressure Research II: Application of High Pressure to Low-Dimensional Novel Electronic Materials*, ed. by H.D. Hochheimer, B. Kuchta, P.K. Dorhout, J.L. Yarger (Kluwer, Dordrecht, 2001), NATO Science Series, p. 87. DOI 10.1007/978-94-010-0520-3_7
79. E. Gregoryanz, O. Degtyareva, M. Somayazulu, R.J. Hemley, H. Mao, Phys. Rev. Lett. **94**(18), 185502 (2005). DOI 10.1103/PhysRevLett.94.185502
80. R. Boehler, Phys. Rev. B **27**(11), 6754 (1983). DOI 10.1103/PhysRevB.27.6754
81. A. Lazicki, Y. Fei, R.J. Hemley, Solid State Commun. **150**(13Ž01314), 625 (2010). http://dx. doi.org/10.1016/j.ssc.2009.12.029
82. A.M.J. Schaeffer, W.B. Talmadge, S.R. Temple, S. Deemyad, Phys. Rev. Lett. **109**(18), 185702 (2012). DOI 10.1103/PhysRevLett.109.185702
83. I. Tamblyn, J.Y. Raty, S.A. Bonev, Phys. Rev. Lett. **101**(7), 075703 (2008). DOI 10.1103/PhysRevLett.101.075703
84. E.R. Hernández, J. Íñiguez, Phys. Rev. Lett. **98**(5), 055501 (2007). DOI 10.1103/PhysRevLett. 98.055501
85. E.R. Hernández, A. Rodriguez-Prieto, A. Bergara, D. Alfè, Phys. Rev. Lett. **104**(18), 185701 (2010). DOI 10.1103/PhysRevLett.104.185701
86. J. Raty, E. Schwegler, S.A. Bonev, Nature **449**, 448 (2007). DOI 10.1038/nature06123
87. M. Martinez-Canales, A. Bergara, J. Phys. Chem. Solids **69**(9), 2151 (2008). http://dx.doi.org/10.1016/j.jpcs.2008.03.022
88. L. Koči, R. Ahuja, L. Vitos, U. Pinsook, Phys. Rev. B **77**(13), 132101 (2008). DOI 10.1103/PhysRevB.77.132101
89. R. Rousseau, K. Uehara, D.D. Klug, J.S. Tse, Chem. Phys. Chem. **6**(9), 1703 (2005). DOI 10.1002/cphc.200500117
90. Y. Ma, A.R. Oganov, Y. Xie, Phys. Rev. B **78**(1), 014102 (2008). DOI 10.1103/PhysRevB. 78.014102
91. C.J. Pickard, R.J. Needs, Phys. Rev. Lett. **102**, 146401 (2009). DOI 10.1103/PhysRevLett. 102.146401
92. N.H. March, G.G.N. Angilella, R. Pucci, Int. J. Mod. Phys. B **27**, 1330021 (2013). DOI 10.1142/S0217979213300211
93. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, A.A. Firsov, Science **306**, 666 (2004). DOI 10.1126/science.1102896
94. A.K. Geim, K.S. Novoselov, Nat. Mater. **6**(3), 183 (2007). DOI 10.1038/nmat1849
95. P.R. Wallace, Phys. Rev. **71**, 622 (1947). DOI 10.1103/PhysRev.71.622
96. A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, Rev. Mod. Phys. **81**(1), 109 (2009). DOI 10.1103/RevModPhys.81.109

97. G. Forte, A. Grassi, G.M. Lombardo, A. La Magna, G.G.N. Angilella, R. Pucci, R. Vilardi, Phys. Lett. A **372**, 6168 (2008). DOI 10.1016/j.physleta.2008.08.014

98. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Phys. Rev. B **80**, 094203 (2009). DOI 10.1103/PhysRevB.80.094203

99. A. Pidatella, R. Mazzarello, in *Correlations in condensed matter under extreme conditions*, ed. by G.G.N. Angilella, A. La Magna (Springer, Berlin, 2017), chap. 14, p. 197. (This volume.)

100. F.M.D. Pellegrino, R. Pucci, G.G.N. Angilella. PRB Kaleidoscope Image for Vol. 80 Issue 9 (2009). http://prb.aps.org/covers/80/9

101. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Eur. Phys. J. B **76**, 469 (2010). DOI 10.1140/epjb/e2010-00228-9

102. B. Uchoa, A.H. Castro Neto, Phys. Rev. Lett. **98**, 146801 (2007). DOI 10.1103/PhysRevLett.98.146801

103. V.M. Pereira, A.H. Castro Neto, Phys. Rev. Lett. **103**(4), 046801 (2009). DOI 10.1103/PhysRevLett.103.046801

104. T.J. Booth, P. Blake, R.R. Nair, D. Jiang, E.W. Hill, U. Bangert, A. Bleloch, M. Gass, K.S. Novoselov, M.I. Katsnelson, A.K. Geim, Nano Letters **8**(8), 2442 (2008). DOI 10.1021/nl801412y

105. K.S. Kim, Y. Zhao, H. Jang, S.Y. Lee, J.M. Kim, K.S. Kim, J.H. Ahn, P. Kim, J. Choi, B.H. Hong, Nature **457**, 706 (2009). DOI 10.1038/nature07719

106. F. Liu, P. Ming, J. Li, Phys. Rev. B **76**(6), 064120 (2007). DOI 10.1103/PhysRevB.76.064120

107. E. Cadelano, P.L. Palla, S. Giordano, L. Colombo, Phys. Rev. Lett. **102**(23), 235502 (2009). DOI 10.1103/PhysRevLett.102.235502

108. S.M. Choi, S.H. Jhi, Y.W. Son, Phys. Rev. B **81**(8), 081407(R) (2010). DOI 10.1103/PhysRevB.81.081407

109. J.W. Jiang, J.S. Wang, B. Li, Phys. Rev. B **81**(7), 073405 (2010). DOI 10.1103/PhysRevB.81.073405

110. G. Gui, J. Li, J. Zhong, Phys. Rev. B **78**(7), 075435 (2008). DOI 10.1103/PhysRevB.78.075435

111. V.M. Pereira, A.H. Castro Neto, N.M.R. Peres, Phys. Rev. B **80**(4), 045401 (2009). DOI 10.1103/PhysRevB.80.045401

112. R.M. Ribeiro, V.M. Pereira, N.M.R. Peres, P.R. Briddon, A.H.C. Neto, New J. Phys. **11**(11), 115002 (2009). DOI 10.1088/1367-2630/11/11/115002

113. G. Cocco, E. Cadelano, L. Colombo, Phys. Rev. B **81**(24), 241412 (2010). DOI 10.1103/PhysRevB.81.241412

114. A.A. Varlamov, G. Balestrino, E. Milani, D.V. Livanov, Adv. Phys. **48**, 655 (1999). DOI 10.1080/000187399243400

115. G. Montambaux, F. Piechon, J. Fuchs, M.O. Goerbig, Phys. Rev. B **80**(15), 153412 (2009). DOI 10.1103/PhysRevB.80.153412

116. X. Wen, *Quantum Field Theory of Many-Body Systems* (Oxford University Press, Oxford, 2007)

117. B. Wunsch, F. Guinea, F. Sols, New J. Phys. **10**, 103027 (2008). DOI 10.1088/1367-2630/10/10/103027

118. S.L. Zhu, B. Wang, L.M. Duan, Phys. Rev. Lett. **98**(26), 260402 (2007). DOI 10.1103/PhysRevLett.98.260402

119. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Phys. Rev. B **81**, 035411 (2010). DOI 10.1103/PhysRevB.81.035411

120. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, High Press. Res. **29**, 569 (2009). DOI 10.1080/08957950903421701

121. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, High Press. Res. **31**, 98 (2011). DOI 10.1080/08957959.2010.525705

122. S.H. Abedinpour, G. Vignale, A. Principi, M. Polini, W.K. Tse, A.H. MacDonald, Phys. Rev. B **84**(4), 045429 (2011). DOI 10.1103/PhysRevB.84.045429

123. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Phys. Rev. B **82**, 115434 (2010). DOI 10.1103/PhysRevB.82.115434

124. Z.Q. Li, E.A. Henriksen, Z. Jiang, Z. Hao, M.C. Martin, P. Kim, H.L. Stormer, D.N. Basov, Nature Phys. **4**, 532 (2008). DOI 10.1038/nphys989
125. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Phys. Rev. B **84**, 195407 (2011). DOI 10.1103/PhysRevB.84.195407
126. S.A. Mikhailov, K. Ziegler, Phys. Rev. Lett. **99**(1), 016803 (2007). DOI 10.1103/PhysRevLett.99.016803
127. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Phys. Rev. B **84**, 195404 (2011). DOI 10.1103/PhysRevB.84.195404
128. F.M.D. Pellegrino, G.G.N. Angilella, R. Pucci, Phys. Rev. B **85**, 195409 (2012). DOI 10.1103/PhysRevB.85.195409
129. J.M. Pereira, V. Mlinar, F.M. Peeters, P. Vasilopoulos, Phys. Rev. B **74**(4), 045424 (2006). DOI 10.1103/PhysRevB.74.045424
130. M. Barbier, P. Vasilopoulos, F.M. Peeters, Phys. Rev. B **80**(20), 205415 (2009). DOI 10.1103/PhysRevB.80.205415
131. M. Barbier, P. Vasilopoulos, F.M. Peeters, Phys. Rev. B **81**(7), 075438 (2010). DOI 10.1103/PhysRevB.81.075438
132. M. Barbier, P. Vasilopoulos, F.M. Peeters, Phys. Rev. B **82**(23), 235408 (2010). DOI 10.1103/PhysRevB.82.235408
133. N.M.R. Peres, J. Phys.: Cond. Matter **21**(32), 323201 (2009). DOI 10.1088/0953-8984/21/32/323201
134. J. Cayssol, B. Huard, D. Goldhaber-Gordon, Phys. Rev. B **79**(7), 075428 (2009). DOI 10.1103/PhysRevB.79.075428
135. S. Gattenlöhner, W. Belzig, M. Titov, Phys. Rev. B **82**(15), 155417 (2010). DOI 10.1103/PhysRevB.82.155417
136. A. Concha, Z. Tešanović, Phys. Rev. B **82**(3), 033413 (2010). DOI 10.1103/PhysRevB.82.033413
137. A. Raoux, M. Polini, R. Asgari, A.R. Hamilton, R. Fazio, A.H. MacDonald, Phys. Rev. B **81**(7), 073407 (2010). DOI 10.1103/PhysRevB.81.073407
138. F.M.D. Pellegrino, in *Correlations in condensed matter under extreme conditions*, ed. by G.G.N. Angilella, A. La Magna (Springer, Berlin, 2017), chap. 12, p. 169. (This volume.)

Chapter 4
Anomalous Magnetism and Superconductivity in Lanthanide Metals at Extreme Pressure

James S. Schilling

Abstract Under ambient pressure the only lanthanide known to superconduct is La, the superconducting state for the remaining lanthanides being suppressed by their strong local-moment magnetism. Except for possibly Ce, this magnetism is conventional and approximately obeys de Gennes scaling. Under high pressure both Ce and Eu exhibit superconductivity that may be unconventional, whereas the magnetic states of Dy, Tb, and Nd become anomalous, the magnetic ordering temperature of Dy surpassing ambient temperature at Mbar pressures. We suggest that these anomalously high magnetic ordering temperatures are an heretofore unrecognized feature of the Kondo lattice state.

Introduction

As well illustrated in the theoretical work of Professor Renato Pucci, high pressure is a thermodynamic variable that is particularly well suited to further our understanding of complex states of matter. The static multi-Mbar pressures available today through diamond anvil cell technology are sufficient to increase the energy per atom by $1–10\,eV$ and thus are capable of significantly altering the ground states of matter. For example, stable magnetic systems eventually destabilize under sufficient pressure, leading to new and unexpected forms of magnetism and/or superconductivity.

In the Periodic Table of Superconductivity (see Fig. 4.1) [1, 2] the 30 elements known to superconduct at ambient pressure appear with a yellow background whereas light green adorns the additional 23 elements that only become superconducting under high pressure. Of particular interest is the lanthanide series of elements where the magnetic $4f$ orbitals are successively filled with electrons, yielding a localized magnetic state on each ion that suppresses superconductivity. Nonmagnetic La, being devoid of $4f$ electrons, is the only lanthanide that superconducts at ambient pressure. La's neighbor, Ce, is magnetic with one $4f$ electron, but becomes superconducting under pressure at $5\,GPa$ [4]. As seen in Fig. 4.2, at $7\,kbar$ ($0.7\,GPa$) Ce

J.S. Schilling (✉)
Department of Physics, Washington University, St. Louis, MO 63130, USA
e-mail: jss@wuphys.wustl.edu

Periodic Table of Superconductivity

(dedicated to the memory of Bernd Matthias; compiled by James S. Schilling)

30 elements superconduct at ambient pressure, 23 more superconduct at high pressure.

Legend: ambient pressure superconductor (yellow): T_c(K), T_c^{max}(K), P(GPa). high pressure superconductor (light green): T_c^{max}(K), P(GPa).

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
H																	He
Li 0.0004 / 14 / 30	Be 0.026											B 11 / 250	C	N	O 0.6 / 100	F	Ne
Na	Mg											Al 1.14	Si 8.2 / 15.2	P 13 / 30	S 17.3 / 190	Cl	Ar
K	Ca 29 / 217	Sc 19.6 / 106	Ti 0.39 / 3.35 / 56.0	V 5.38 / 16.5 / 120	Cr	Mn	Fe 2.1 / 21	Co	Ni	Cu	Zn 0.875	Ga 1.091 / 7 / 1.4	Ge 5.35 / 11.5	As 2.4 / 32	Se 8 / 150	Br 1.4 / 100	Kr
Rb	Sr 19.5 / 50	Y 19.5 / 115	Zr 0.546 / 11 / 30	Nb 9.50 / 9.9 / 10	Mo 0.92	Tc 7.77	Ru 0.51	Rh .00033	Pd	Ag	Cd 0.56	In 3.404	Sn 3.722 / 5.3 / 11.3	Sb 3.9 / 25	Te 7.5 / 35	I 1.2 / 25	Xe
Cs 1.3 / 12	Ba 5 / 18	insert La-Lu	Hf 0.12 / 8.6 / 62	Ta 4.483 / 4.5 / 43	W 0.012	Re 1.4	Os 0.655	Ir 0.14	Pt	Au	Hg-α 4.153	Tl 2.39	Pb 7.193	Bi 8.5 / 9.1	Po	At	Rn
Fr	Ra	insert Ac-Lr	Rf	Ha													

La-fcc 6.00 / 13 / 15	Ce 1.7 / 5	Pr	Nd	Pm	Sm	Eu 2.75 / 142	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu 12.4 / 174
Ac	Th 1.368	Pa 1.4	U 0.8(β) / 2.4(α) / 1.2	Np	Pu	Am 0.79 / 2.2 / 6	Cm	Bk	Cf	Es	Fm	Md	No	Lr

Fig. 4.1 Periodic Table listing elements that superconduct at ambient pressure (*yellow*) or only under high pressure (*light green*) [1, 2]. For each element the upper position gives the value of T_c(K) at ambient pressure; middle position gives maximum value T_c^{max}(K) in a high-pressure experiment at the pressure P(GPa) (lower position). If T_c decreases under pressure, only the ambient pressure value of T_c is given. Sources for T_c values are given in Ref. [3] (color figure online)

suffers a large 16% volume collapse within the fcc structure [5] whereby its magnetic moment appears to be quenched. Near this volume collapse pressure a dilute (2 at.%) concentration of magnetic Ce impurities in La is seen to strongly suppress La's superconductivity over a narrow pressure region [6]. This is the hallmark of Kondo effect phenomena with giant pair breaking [7]. A Kondo resistivity minimum has also been observed in dilute La(Ce) alloys [8]. Ce's volume collapse itself can be accounted for within the Kondo volume collapse model of Allen and Martin [9].

The appearance of Kondo effect phenomena indicates that the magnetic state of Ce is approaching an instability. Besides suppressing superconductivity, the Kondo effect can also quench magnetic order, sometimes leading to exotic forms of

Fig. 4.2 (*upper*) Molar volume versus pressure in kbar (0.1 GPa), showing 16% volume collapse at 0.7 GPa ($\gamma - \alpha$ transition) [5]. (*lower*) Superconducting transition temperature of the dilute magnetic alloy La(2 at.% Ce) versus pressure compared to that of pure La [6]. The very large superconducting pair-breaking in the alloy leads to a "sinkhole-like" depression of T_c near 1 GPa
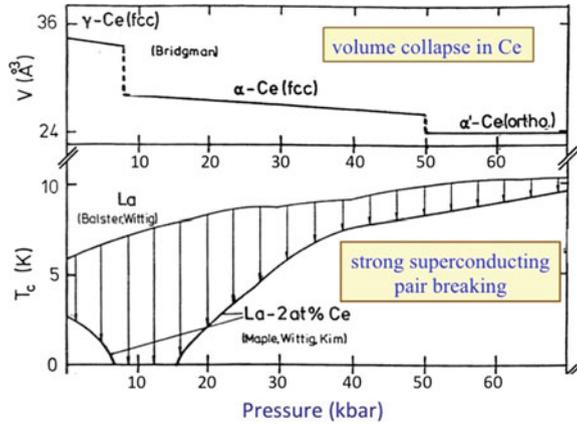
Fig. 4.3 Magnetic ordering temperature T_o plotted versus the absolute value of the negative exchange parameter $|J|$ according the Doniach–Yang model [10, 11]. Since T_o increases only as $|J|^2$, but the Kondo temperature T_K increases exponentially with $|J|$, the latter ultimately dominates and quenches magnetic ordering
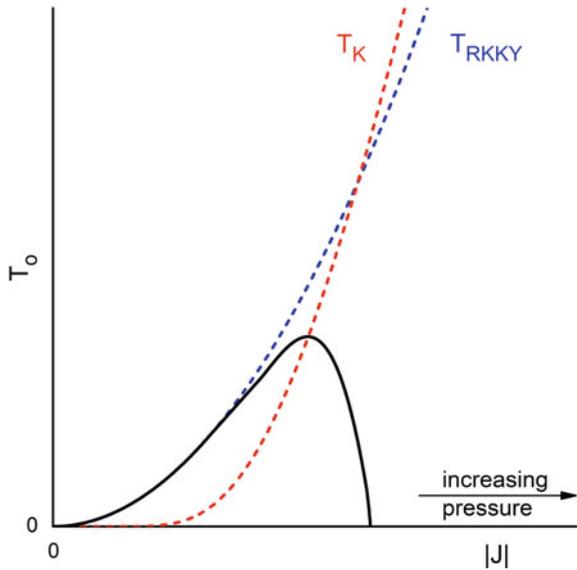
superconductivity. This scenario is illustrated in Fig. 4.3 where the magnetic ordering temperature T_o is plotted versus the magnitude of the negative covalent mixing exchange parameter $|J|$. This is the well-known Kondo lattice model of Doniach [10] and Yang [11]. Since the magnetic ordering strength T_{RKKY} from the Ruderman–Kittel–Kasuya–Yosida (RKKY) exchange interaction [12] initially increases as J^2, but the Kondo temperature T_K increases exponentially with $|J|$, the latter eventually overtakes the former as $|J|$ increases and the magnetic order is quenched. The author, and probably most, if not all, scientists working in the highly correlated electron research area, have tacitly assumed that the value of T_o from the negative covalent

mixing exchange must be weaker than that arising from normal positive exchange. Our present results suggest that the opposite is the case, as we will soon see.

Results and Discussion

Since the normal exchange interaction and the electronic properties of the con- duction electrons change relatively slowly across the lanthanide series, the val- ues of T_o would be expected to approximately obey de Gennes scaling, whereby $T_o \propto (g-1)^2 J_t (J_t + 1)$ (see, e.g., Ref. [13]; here J_t is the total angular momentum quantum number and g is the Landé-g factor. Ce is an exceptional case; T_o for Ce lies at 13.7 K in the hexagonal phase, a value four times higher than that (3.3 K) expected from simple de Gennes scaling compared to Gd, where $T_o = 292$ K [14]. Likewise, the suppression of the superconducting transition temperature ΔT_c in dilute magnetic alloys should also follow de Gennes scaling [7]. Compared to the other magnetic lan- thanides in dilute (1%) alloys with superconducting La, ΔT_c for Ce is much larger than anticipated [15]. The anomalously high value of the magnetic ordering temper- ature $T_o = 13.7$ K for Ce could have been a tip off that anomalously high values of T_o might be possible within the Kondo lattice model.

At ambient pressure Eu is a divalent metal that orders antiferromagnetically near $T_o = 90$ K (see, e.g., Ref. [16]. Previous studies of the dependence of T_o on pressure were restricted to $P \leq 42$ GPa [17]. In recent electrical resistivity and ac magnetic susceptibility measurements to pressures as high as 142 GPa, superconduc- tivity in Eu was found to appear for pressures above 80 GPa [18], as seen in Fig. 4.4. Eu thus became the latest pressure-induced superconducting element. Unfortunately, the magnetic ordering temperature of Eu could not be determined in this experiment so it was not possible to investigate the possible interplay between the magnetic

Fig. 4.4 Real part of the ac susceptibility versus temperature for Eu metal as pressure is increased from 76 to 142 GPa [18]. The superconducting transition appears at 84 GPa and shifts slowly under pressure to higher temperatures
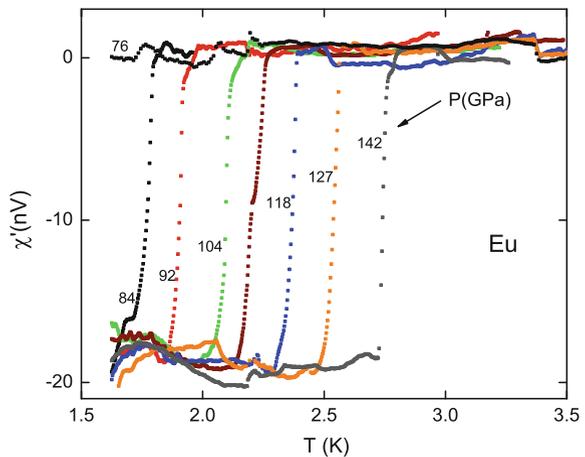
and superconducting states, as indicated for the Doniach–Yang model in Fig. 4.3. Although the complete magnetic/superconducting phase diagram has yet to be mapped out, a very recent Synchrotron Mössbauer Spectroscopy (SMS) experiments to 101 GPa find that magnetic order in Eu vanishes near the same pressure (80 GPa) where the superconducting transition appears [19]. A parallel X-ray Emission Spectroscopy (XES) experiment on Eu to 119 GPa found no evidence for a change in either Eu's local magnetic state or in its valence [19]. It would be very interesting to accurately track in a single experiment the pressure dependence of both the magnetic ordering T_o and superconducting T_c temperatures in Eu to at least 100 GPa pressure. In recent electrical resistivity measurements the pressure dependence of T_o has been determined for Dy [20], Tb [21], Gd [21], and Nd [22] to pressures of ∼150 GPa.

The magnetic ordering of the lanthanides at ambient pressure is well explained [13, 23] by a conduction-band-driven indirect RKKY exchange interaction [12]. As pointed out above, the magnetic ordering temperature T_o, as well as the strength of superconducting pair breaking ΔT_c, are both expected to scale with the de Gennes factor $(g - 1)^2 J_t(J_t + 1)$ modulated by $J N(E_f)$ [13], where J is the exchange interaction between the $4f$ ion and the conduction electrons and $N(E_f)$ is the density of states at the Fermi energy. The heavy lanthanide Dy possesses a large local magnetic moment and orders antiferromagnetically at 178 K. The value of this ordering temperature is approximately consistent with that of Gd at 292 K since the de Gennes factor of Dy is about half that of Gd.

In Fig. 4.5 are shown the pressure-dependent magnetic ordering temperatures of Dy and Gd from recent electrical resistivity measurements in a diamond-anvil cell to 157 GPa for Dy and 105 GPa for Gd [20]. Both lanthanides are seen to undergo
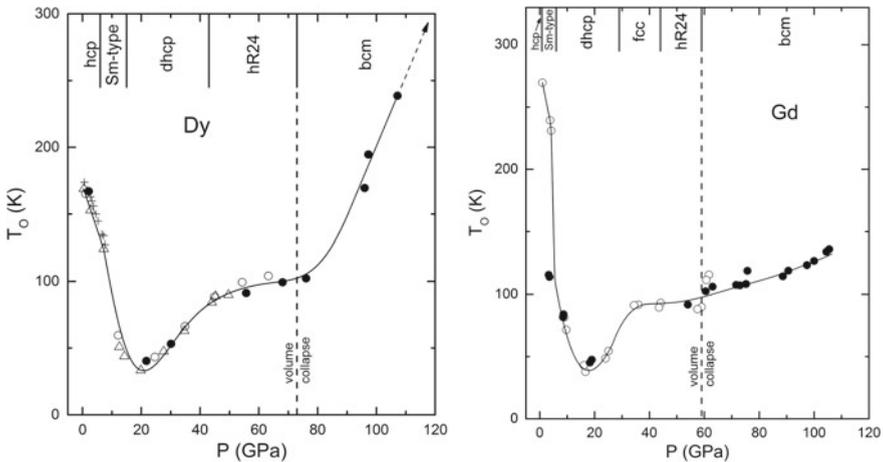
Fig. 4.5 Magnetic ordering temperature T_o of Dy and Gd versus pressure [20]. (+) earlier studies to ∼8 GPa on Dy [24]. References for crystal structures at top of graphs Dy [24], Gd [26]; vertical *dashed line* gives pressure where 5–6% volume collapse occurs. In both plots the extended *solid line* through data points is guide to the eye

similar structural phase transitions over this pressure range and exhibit a $\sim$4% volume collapse near 60–70 GPa, a much smaller value than that of Ce (16%). Up to 70 GPa the pressure dependences $T_o(P)$ for Dy and Gd are seen to be very similar. The same applies for Tb [21]. Since the de Gennes factor would not be expected to change under pressure, unless a valence transition occurs, the very similar pressure dependences $T_o(P)$ of Gd and Dy to 70 GPa point to a common mechanism and likely originate from the pressure dependence of $JN(E_f)$, facilitated by a series of nearly identical phase transitions in Dy [24], Tb [25], and Gd [26] driven by increasing $5d$-electron concentration with pressure [27]. In fact, Fleming and Liu [28, 29] show that the initial decrease in T_o with pressure for Dy, Tb, and Gd is a direct consequence of energy shifts in the valence band structure.

Above 70 GPa, however, the $T_o(P)$ dependences for Dy and Gd are seen in Fig. 4.5 to differ markedly, $T_o(P)$ for Dy showing an extremely rapid increase with pressure, extrapolating to 400 K at the highest pressure of the experiment (157 GPa) [20]. How rapid this increase really is most clearly demonstrated if T_o is plotted versus relative sample volume V/V_o; in such a plot it can be seen that the increase in T_o above 70 GPa is *far* more rapid than the initial decrease with pressure [20]. For lanthanides the compressibility near 70 GPa is approximately 10-times less than that at 1 bar [24, 25].

Note that the $T_o(P)$ dependence for Gd shows only a slow increase with pressure, in contrast to the results for Dy. Since the conduction electron properties of Dy and Gd are so similar, the extremely rapid increase of T_o above 70 GPa would suggest that Dy's magnetic state has changed in a significant way. This would be the case if the magnetic state of Dy at these extreme pressure is becoming unstable leading to an increase in the magnitude of the negative covalent mixing exchange parameter $|J|$ and consequently in the Kondo temperature T_K. The question thus arises whether the rapid increase in T_o under pressure for Dy might be mirroring the left part of the Doniach–Yang phase diagram in Fig. 4.3 where T_o is seen to increase rapidly with $|J|$. If this is the case, then, on the basis of Fig. 4.3, it would be expected that at pressures well above 157 GPa, T_o for Dy would begin to fall rapidly toward 0 K. Unfortunately, pressures appreciably above 157 GPa are extremely difficult to obtain in a true four-point resistivity experiment.

That the temperature-dependent resistivity gives a reliable value for the magnetic ordering temperature T_o has been very recently confirmed through SMS measurements by Bi et al. [30] on Dy to pressures as high as 141 GPa.

Parallel experiments on the lanthanide Tb [21] yield results similar to those for Dy. Above 80 GPa T_o for Tb begins to rise with pressure extremely rapidly to the highest pressure of the experiment at 141 GPa.

Why does the $T_o(P)$ dependence for Gd above 70 GPa differ so markedly from that of Dy and Tb? The absence of magnetic instabilities in Gd, even at extreme pressures, is not surprising since the magnetic state of Gd with its half-filled $4f^7$ shell is by far the most stable of all elements, its $4f^7$ level lying $\sim$9 eV below the Fermi level [31]. A pressure of "only" 100 GPa (1 Mbar) is not nearly enough to destabilize Gd. This is reflected in the slow monotonic increase of T_o with pressure for Gd in the bcm phase seen in Fig. 4.4.

Fig. 4.6 Superconducting transition temperature T_c versus pressure for Y(1 at.% Dy) compared to that for Y [20]. Inset shows similar graph for Y(0.5 at.% Gd) [32]. Vertical *dashed line* marks pressure of volume collapse for Dy at 73 GPa [24] and in inset for Gd at 59 GPa [26]. At top of graph are crystal structures taken on by superconducting host Y [33]. Extended *solid line* through data points is guide to the eye
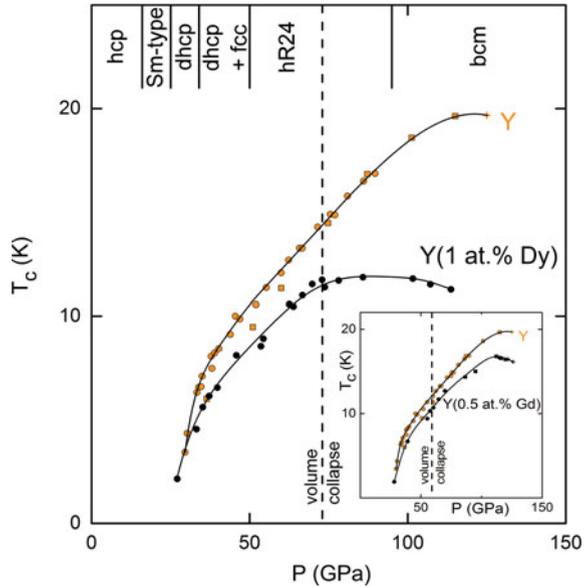
It is well known that the electronic properties across the lanthanide series are closely related. From La to Lu the number of d-electrons per atom decreases monotonically, nicely accounting for the observed changes in crystal structure [27]. Under pressure the d-electron count per atom increases and leads to the same progression of crystal structures under pressure as from right to left across the lanthanide series [27]. The conduction electron environment of the magnetic $4f$ orbitals changes only slowly across the series. A long-standing strategy [7, 15] to probe the magnetic state of a given ion in a concentrated magnetic material is to alloy the magnetic ion in dilute concentration with a superconductor and determine to what extent the superconducting transition temperature is suppressed ΔT_c. If the magnetic state of Dy becomes unstable under pressures above 70 GPa, one would expect that a dilute concentration of Dy in superconducting La or Y would show an anomalously strong suppression of superconductivity near this pressure.

That this actually occurs is seen in Fig. 4.6 where the pressure dependence of T_c for pure Y is compared to that of the dilute magnetic alloy Y(1 at.% Dy). $T_c(P)$ for the alloy begins to markedly pull away from that of Y above 70 GPa, i.e., ΔT_c increases sharply! The pair-breaking near 120 GPa is seen to be very strong, approaching $\Delta T_c \approx 9$ K. A similar result is obtained for dilute alloys of Y with Tb [21].

On the other hand, alloying the ultrastable Gd ion into Y gives a very different result. $T_c(P)$ for both Y and Y(0.5 at.% Gd) track each other over the entire range of pressure [32]. This result supports the scenario that the very rapid increase in T_o with pressure above 70 GPa for Dy and Tb has its origin in the increasing instability of their magnetic states. Very recent X-ray Absorption Near Edge Structure (XANES) studies confirm that no change in valence occurs in Dy to 115 GPa [34], as was found

earlier for Tb to 65 GPa [32]. As the magnetic $4f$ level approaches the Fermi level with increasing pressure, the exchange interaction with the conduction electrons takes on a negative sign, signaling the onset of strong Kondo resonance phenomena. This enhancement of the magnitude of $|J|$ with pressure causes the magnetic ordering temperature $T_o \propto |J|^2$ to increase until $|J|$ becomes so large that the local magnetic moment begins to be compensated through the exponentially increasing Kondo spin screening, as anticipated in the simple Doniach–Yang Kondo lattice model [10, 11] (see Fig. 4.3). This then leads to an anomalously high value of T_o, such as observed for Dy and Tb at extreme pressure, a value surpassing that possible for normal positive exchange interactions.

To establish whether or not a Doniach–Yang-like model is appropriate to interpret the present results for Dy and Tb, much higher pressures would be necessary to search for the sharp downswing in $T_o(P)$ seen in Fig. 4.3. Since this is not currently feasible experimentally, another option is to search for a lanthanide where the anomalous dependence in $T_o(P)$ begins to occur at a much lower pressure. In very recent resistivity measurements on Nd, Song et al. [22] have found a Doniach–Yang-like $T_o(P)$ dependence where T_o first rises slowly, then rapidly, but then passes through a maximum and falls toward 0 K under pressure. For dilute Nd magnetic impurities in Y the suppression of superconductivity is found to be even more dramatic than for Dy or Tb impurities. In addition, for pressures near 100 GPa a Kondo resistivity minimum is found for Y(Nd) alloys. These results are the subject of a future publication [22].

In summary, in contrast to the magnetically ultrastable lanthanide Gd, the magnetic ordering temperature T_o increases dramatically for pressures above 70 GPa for Dy and Tb and 20 GPa for Nd. In the same pressure region dilute magnetic impurities of Dy, Tb, and Nd in superconducting Y show giant pair-breaking effects. These results are consistent with the scenario that under extreme pressures these three lanthanide metals approach a magnetic instability and turn into Kondo lattices.

Lanthanide elements play a central role in many modern technologies, including permanent magnets, computer memories, and applications requiring giant magnetostriction. Unfortunately, their sub-ambient magnetic ordering temperatures T_o force the addition of transition elements to substantially raise T_o, such as Fe or Co in $SmCo_5$ and $Nd_2Fe_{14}B$, the two most powerful permanent magnetic materials. If the mechanism(s) responsible for these anomalously high values of T_o under extreme pressure can be clearly identified, one can try to reproduce these conditions in a suitable compound at ambient pressure and possibly synthesize a superior permanent magnet material. This strategy receives some support from the fact that the Curie temperature of the ternary compound $CeRh_3B_2$ lies at 115 K [35], a value even exceeding that of $GdRh_3B_2$ and 100-times higher than that anticipated from simple de Gennes scaling. We suggest that Dy under extreme pressure and $CeRh_3B_2$ at ambient pressure may share a common mechanism for their anomalous magnetic properties.

References

1. M. Debessai, T. Matsuoka, J.J. Hamlin, W. Bi, Y. Meng, K. Shimizu, J.S. Schilling, J. Phys.: Conf. Ser. **215**(1), 012034 (2010). DOI 10.1088/1742-6596/215/1/012034
2. M. Sakata, Y. Nakamoto, K. Shimizu, T. Matsuoka, Y. Ohishi, Phys. Rev. B **83**, 220512 (2011). DOI 10.1103/PhysRevB.83.220512
3. M. Debessai, J.J. Hamlin, J.S. Schilling, Phys. Rev. B **78**, 064519 (2008). DOI 10.1103/PhysRevB.78.064519
4. J. Wittig, Phys. Rev. Lett. **21**, 1250 (1968). DOI 10.1103/PhysRevLett.21.1250
5. P.W. Bridgman, Proc. Am. Acad. Arts Sci. **76**, 55 (1948)
6. M.B. Maple, J. Wittig, K.S. Kim, Phys. Rev. Lett. **23**, 1375 (1969). DOI 10.1103/PhysRevLett.23.1375
7. M.B. Maple, Appl. Phys. **9**(3), 179 (1976). DOI 10.1007/BF00900605
8. K.S. Kim, M.B. Maple, Phys. Rev. B **2**, 4696 (1970). DOI 10.1103/PhysRevB.2.4696
9. J.W. Allen, R.M. Martin, Phys. Rev. Lett. **49**, 1106 (1982). DOI 10.1103/PhysRevLett.49.1106
10. S. Doniach, in *Valence Instabilities and Related Narrow-Band Phenomena*, ed. by R.D. Parks (Springer, Boston, 1977), pp. 169–176. DOI 10.1007/978-1-4615-8816-0_15
11. Y.f. Yang, Z. Fisk, H.O. Lee, J.D. Thompson, D. Pines, Nature **454**(7204), 611 (2008). DOI 10.1038/nature07157
12. M.A. Ruderman, C. Kittel, Phys. Rev. **96**, 99 (1954). DOI 10.1103/PhysRev.96.99
13. K.N.R. Taylor, M.I. Darby, *Physics of Rare Earth Solids* (Chapman and Hall, London, 1972)
14. R.V. Colvin, S. Legvold, F.H. Spedding, Phys. Rev. **120**, 741 (1960). DOI 10.1103/PhysRev.120.741
15. B.T. Matthias, H. Suhl, E. Corenzwit, Phys. Rev. Lett. **1**, 92 (1958). DOI 10.1103/PhysRevLett.1.92
16. P.H. Barrett, D.A. Shirley, Phys. Rev. **131**, 123 (1963). DOI 10.1103/PhysRev.131.123
17. F.P. Bundy, K.J. Dunn, Phys. Rev. B **24**, 4136 (1981). DOI 10.1103/PhysRevB.24.4136
18. M. Debessai, T. Matsuoka, J.J. Hamlin, J.S. Schilling, K. Shimizu, Phys. Rev. Lett. **102**, 197002 (2009). DOI 10.1103/PhysRevLett.102.197002
19. W. Bi, J. Lim, G. Fabbris, J. Zhao, D. Haskel, E.E. Alp, M.Y. Hu, P. Chow, Y. Xiao, W. Xu, J.S. Schilling, Phys. Rev. B **93**, 184424 (2016). DOI 10.1103/PhysRevB.93.184424
20. J. Lim, G. Fabbris, D. Haskel, J.S. Schilling, Phys. Rev. B **91**, 045116 (2015). DOI 10.1103/PhysRevB.91.045116
21. J. Lim, G. Fabbris, D. Haskel, J.S. Schilling, Phys. Rev. B **91**, 174428 (2015). DOI 10.1103/PhysRevB.91.174428
22. J. Song, W. Bi, D. Haskel, J.S. Schilling, (in preparation)
23. D.D. Jackson, V. Malba, S.T. Weir, P.A. Baker, Y.K. Vohra, Phys. Rev. B **71**, 184416 (2005). DOI 10.1103/PhysRevB.71.184416
24. R. Patterson, C.K. Saw, J. Akella, J. Appl. Phys. **95**(10), 5443 (2004). DOI 10.1063/1.1699489
25. N.C. Cunningham, W. Qiu, K.M. Hope, H.P. Liermann, Y.K. Vohra, Phys. Rev. B **76**, 212101 (2007). DOI 10.1103/PhysRevB.76.212101
26. H. Hua, V.K. Vohra, J. Akella, S.T. Weir, R. Ahuja, B. Johansson, Rev. High Press. Sci. Technol. **7**, 233 (1998)
27. J.C. Duthie, D.G. Pettifor, Phys. Rev. Lett. **38**, 564 (1977). DOI 10.1103/PhysRevLett.38.564
28. G.S. Fleming, S.H. Liu, Phys. Rev. B **2**, 164 (1970). DOI 10.1103/PhysRevB.2.164

29. S.H. Liu, Phys. Rev. **127**, 1889 (1962). DOI 10.1103/PhysRev.127.1889
30. W. Bi, J. Song, Y. Deng, E. Alp, J.S. Schilling, (in preparation)
31. Z.P. Yin, W.E. Pickett, Phys. Rev. B **74**, 205106 (2006). DOI 10.1103/PhysRevB.74.205106
32. G. Fabbris, T. Matsuoka, J. Lim, J.R.L. Mardegan, K. Shimizu, D. Haskel, J.S. Schilling, Phys. Rev. B **88**, 245103 (2013). DOI 10.1103/PhysRevB.88.245103
33. G.K. Samudrala, G.M. Tsoi, S.T. Weir, Y.K. Vohra, High Press. Res. **34**(2), 266 (2014). DOI 10.1080/08957959.2014.903946
34. W. Bi, J. Song, G. Fabbris, D. Haskel, J.S. Schilling, (in preparation)
35. S.K. Dhar, S.K. Malik, R. Vijayaraghavan, J. Phys. C: Solid State Phys. **14**(11), L321 (1981). DOI 10.1088/0022-3719/14/11/008

Chapter 5
Phonon Fingerprints on Low and High-Energy Spectrum of Cuprate Superconductors

R. Citro

Abstract The strongly correlated electron systems with low-dimensionality (one and two) have attracted a great interest in the last 30 years, both for their experimental realization and for the great numbers of theoretical techniques to study their relevant models. To this class of systems belong the cuprate superconductors, whose physics is described by the properties of the CuO_2 planes. Our research has been focused on the theoretical study of the interplay of electron–electron correlations and electron–phonon interaction in these materials. In particular, we have been investigating the low and high-energy kink in the electron spectrum and the anomalous bond-stretching (BS) in-plane phonon softening via a semi-phenomenological model. It describes the coupling of the phonons to a charge mode in the mid-infrared energy range. Experiments of oxygen isotope substitution on the phonon softening and linewidth are suggested as a source of important information on the charge excitations and equations describing isotope effects are derived. We discuss how the mid-infrared scenario could offer an explanation of the high-energy kink and the anomalous BS phonon softening, observed in different perovskites, in a unified framework.

Introduction

The role of phonons in the physics of high-T_c cuprates has attracted considerable interest and many debates have been generated by different experiments in the last decade. Essentially, there appears more evidence for the relevant contribution of phonons to the low-energy electronic properties of these materials. In particular, angular resolved photoemission spectroscopy (ARPES) experiments have revealed a singularity (kink) in the single-particle dispersion at energies $\sim$100 meV [1, 2] which persists both above and below the critical temperature and is observed in nodal as well as in antinodal directions of the Brillouin zone. The kink is associated to the "peak-dip-hump" structure in the spectral density which also shows a significant

R. Citro (✉)

Dipartimento di Fisica "E. R. Caianiello" and Spin-CNR, Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84081 Fisciano (SA), Italy

e-mail: citro@sa.infn.it

isotope effect. Paradoxically, the effect is highest in the high energy, hump region. The velocity ratio, characterizing a change in the dispersion slope above and below the singularity, shows a strong concentration dependence which saturates around optimal doping. It should be mentioned that these features could not be explained within the traditional weak-coupling scheme but can be naturally described within the strong coupling picture. A crucial property consists in tight relation between high and low-energy scales of the excitation spectrum and transfer of spectral weight between them.

In this context, the latest advancement in the ARPES technique has allowed to discover another kink in the electronic dispersion at much larger energy [3–5], around 0.4–0.5 eV. Compared to the low-energy kink, the spectral weight of the other singularity has been found to be even larger. This kink is also characterized by a large vertical drop in energy (comparable to the energy of the kink itself) at the momentum corresponding to the singularity. The origin of this kink is even more controversial, although different "magnetic" scenarios have already been put forward.

In this contribution, we review our previous work explaining the aforementioned features in the cuprate superconductors. Namely, we have been proposing a scenario in which both the high-energy kink in the ARPES spectra and the BS phonon softening could be related to the mid-infrared (MIR) structure of the charge response [6, 7]. The anomalous properties of this region of the spectrum have long been known from optical experiments and have served as motivation for some semi-phenomenological theories, such as marginal Fermi liquid. However existing experimental techniques have not been able to resolve the charge excitations sufficiently in momentum space. Although the energy scale of MIR can be suggestive in itself, most of existing information comes from optical experiments, i.e., from zero momentum, and cannot be directly applied to the analysis of the kink. However we show below that analysis of the bond-stretching phonon mode can be very important for resolving this issue.

Low-Energy Kink in the Single-Particle Spectrum

Before discussing the high-energy features of the quasiparticle spectrum in cuprate superconductors, let us briefly describe a model for the low-energy kink. In Ref. [6] it has been shown that various experimental features of photoemission experiments [1] can be qualitatively explained in a framework that considers both strong correlations effect and the usual electron–phonon interaction. We have considered the Hubbard–Holstein Hamiltonian in which Einstein phonons are coupled to local electron density on a square lattice with nearest-neighbor hopping and have calculated the single-particle Green's function. To calculate the phonon contribution to the electron self-energy, we have assumed the adiabatic approximation for the phonons and we have employed the Hubbard I approximation for the "bare" electronic Green's function [6]. Although the drawbacks of the latter are well known, it well captures the main effect of correlations mentioned above. Namely, it allows to consider the concentration dependence of the dispersion and spectral lineshape reported in experiments,

Fig. 5.1 Quasiparticle dispersion $E - E_F$ at different dopings (indicated in the figure) for the phonon frequency $\omega_0 = 0.035$ and coupling $\lambda = 1.1$, the electron correlation $U = 3$, temperature $T = 0.003$. The energy unit is the bandwidth W. A kink in the energy dispersion is clearly visible
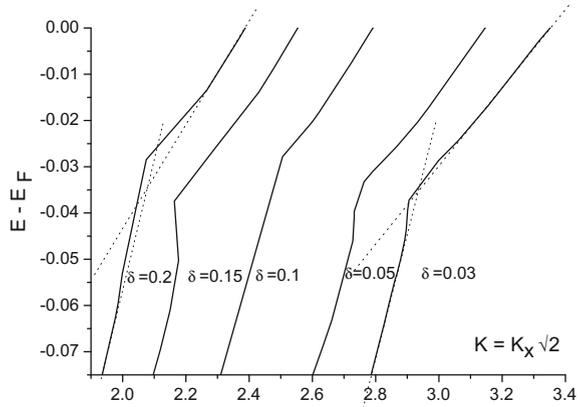
Fig. 5.2 Evolution of the Fermi surface (*continuous lines*) and kink surface (*dashed lines*) with electron concentration in the first quadrant of the Brillouin zone. Note that $c = 1 - \delta$
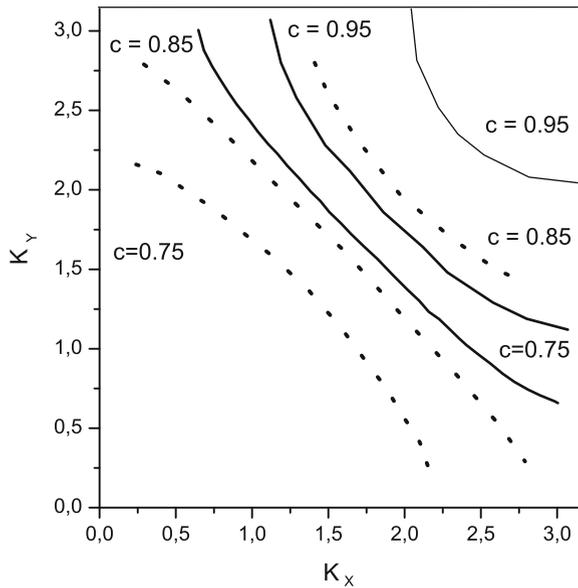
the dependence on isotope substitution, temperature and doping [6]. In particular, we have found a good agreement with experiments of the calculated doping dependence of the velocity ratio, i.e., the ratio of the two slopes characterizing the quasiparticle linear dispersion before and after the kink (see Fig. 5.1).

We have also found an interesting new effect related to a topological change of the kink surface with doping. Our results have also been able to describe the experimentally observed situation in which, at optimal doping, the Fermi surface is hole-like while the kink surface is electron-like (see Fig. 5.2). The kink in the electron spectrum occurs at the energy of the phonon because it separates the quasiparticle state at lower energy, renormalized by interaction with virtual phonons, from states at

higher energy where a real phonon is emitted. Such states correspond to the incoherent part of the ARPES spectra (hump), while the kink is identified with the dip followed by the quasiparticle peak in the vicinity of the Fermi surface.

The Mid-infrared Boson Scenario

The High-Energy Kink

In particular, the LEK is well reproduced by the coupling of the correlated electrons with a phonon (boson mode) whose linewidth is about 5–10% of its energy, however, the situation is quite different for the high-energy kink (HEK). Below we discuss how the features of the HEK can be accounted for by assuming the "overdamped" regime for the "boson", i.e., that the linewidth of this excitation is significantly larger than the characteristic energy. This assumption is motivated by the ubiquity of the incoherent excitations in the mid-infrared (MIR) frequency range of the optical spectra in cuprates as revealed, e.g., by Drude–Lorentz analysis. As explained below, although the bare frequency of such excitation is larger than the HEK, the scale that actually characterizes its effect on electron spectrum falls in the relevant energy range where HEK appears. We consider the self-energy effects within a simple model of electrons interacting with a bosonic mode in the form of damped oscillator. The parameters of the oscillator are chosen to match the known experimental data, in particular those from ARPES. As for instance, its coupling as extracted from the slope of single particle spectrum in the region preceding the HEK (e.g., for BISCO 2212 and LBCO is $\lambda_c = 1 - 1.5$), momentum and energy location of the observed structures etc. [8, 9].

We introduce the retarded bosonic propagator $B(q, \omega)$ in the form of damped oscillator [10] with spectral density $\rho_B(q, \omega)$

$$B(\omega) = \frac{a_q \omega_q}{\omega^2 - \omega_q^2 + i\gamma_q \omega}, \tag{5.1a}$$

$$\rho_B(\omega) = \frac{\omega \gamma_q a_q \omega_q / \pi}{\left(\omega^2 - \omega_q^2\right)^2 + \left(\gamma_q \omega\right)^2}. \tag{5.1b}$$

The choice of Eq. (5.1) matches the generic form of the response function frequently used in analyzing experimental data that, in turn, corresponds to physical requirements, such as Kramers–Krönig relations. Coupling of this mode to electron excitations is considered in the second order of perturbation theory, i.e., non-self-consistently. The scale of the boson frequency is assumed to allow the adiabatic approximation to be reasonable at least for evaluation purposes. Using the spectral representation for the Matsubara Green's function, after carrying out the frequency

summation and analytic continuation to real frequencies one obtains the second-order expression for the retarded electron self-energy, e.g., [11, 12],

$$\Sigma_R(k, \omega) = \frac{1}{N} \sum_q g^2(k, q) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho_F(k - q, x) \rho_B(q, y) \frac{n_B(y) + n_F(x)}{\omega + y - x + i0^+} dx dy,$$
(5.2)

where $g(q)$ is the coupling constant and $n_{B,F}$ are Boson, Fermion distribution functions. When the linewidth of the electronic spectral function $\rho_F(k - q, x)$ is neglected, $\rho_F(k - q, x) = \delta(x - \varepsilon_{k-q})$, and a tight binding parametrization of the band dispersion is used as explained above, that allows to focus on the scattering mechanism due to the "boson". If also the linewidth of $\rho_B(\omega)$ is neglected, the expression Eq. (5.2) transforms into the one that has been often employed in describing the LEK (see, e.g., Ref. [13]) and results in the usual logarithmic singularity of $\mathrm{Re}\,\Sigma(\omega)$ at the boson frequency. To study the effect of the broad linewidth on electron spectrum we neglect the dispersion in spectral function Eq. (5.1) and in the coupling $g(q)$. Then the self-energy becomes a local function

$$\Sigma_R(\omega) = \lambda_c \gamma \omega_0^2 \int_{-\infty}^{\infty} dy \frac{y}{\left(y^2 - \omega_0^2\right)^2 + (\gamma y)^2} \int dE \rho_0(E) \left(\frac{n_B(y) + n_F(E)}{\omega + y - E + i0^+} \right),$$
(5.3)

where $\rho_0(E)$ is the electronic density of states, λ_c is the dimensionless coupling constant. The coupling constant λ determined from the slope of $\mathrm{Re}\,\Sigma(\omega)$ is slightly different from λ_c. The bandwidth of ε_k is taken as the energy unit in numerical calculations.

The spectral density of the damped oscillator, Eq. (5.1a) is peaked at the frequency $\Omega = \left(12 - 6R^2 + 6\sqrt{16 - 4R^2 + R^4}\right)^{1/2} \times \omega_0/6$, where $R = \gamma/\omega_0$. It is clear that the effective frequency Ω is significantly lower than ω_0 due to a large damping typical for the MIR oscillator. As seen from Fig. 5.3, it is this effective frequency that

Fig. 5.3 Real and imaginary parts of the self-energy $\Sigma(\omega)$, Eq. (5.3), resulting from interaction with the overdamped oscillator with spectral density $\rho_B(\omega)$. Parameters are $\omega_0 = 0.18$, $\gamma = 0.3, \lambda = 1.3$, $T = 0.001, \delta = 0.15$. The flattening of $\mathrm{Im}\,\Sigma(\omega)$ corresponds to the incoherent "waterfall" region in ARPES with almost constant linewidth
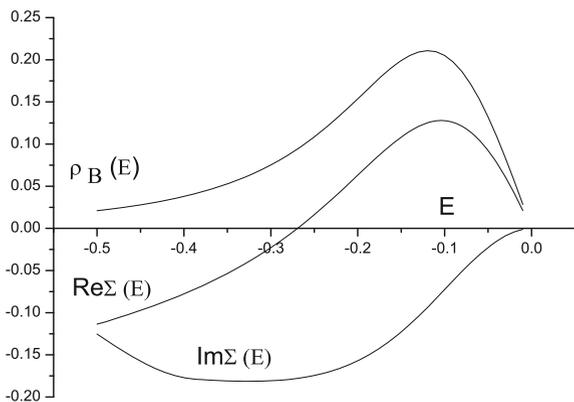
determines the location of the high-energy kink. In the figure we plot the $\mathrm{Re}\,\Sigma(\omega)$ and $\mathrm{Im}\,\Sigma(\omega)$ together with the "boson" spectral function $\rho_B(\omega)$ to make the correlation between the peaks clearly visible. The slope of the $\mathrm{Re}\,\Sigma(\omega)$ corresponds to the coupling constant 1.3 and the ratio of $t'/t = 0.1$ was taken to model the LSCO band ε_k. One can also see that the $(-\mathrm{Im}\,\Sigma(\omega))$ follows a ω^2 at low energies and its inflection point corresponds to Ω. After reaching its maximum at higher energy, $-\mathrm{Im}\,\Sigma(\omega)$ follows an almost flat region until it starts decreasing closer to the band bottom. Such a behavior has been reported in Refs. [4, 14] for the linewidth of the ARPES spectral function. In our calculations we clearly see the feature mentioned in Ref. [15] and related to the finite bandwidth when the renormalized dispersion curve crosses with bare one. The crossing occurs at certain nonzero frequency within the waterfall region where $\mathrm{Re}\,\Sigma(\omega) = 0$.

The calculated dispersion of the spectral function peak along the two main symmetry directions of the Brillouin zone are presented in Fig. 5.4 for doping $\delta = 0.15$ and band structure of BISCO 2212. The dashed curve shows the evolution of the curve at larger doping, $\delta = 0.25$, for the diagonal direction. The HEK is close to $k_0 \simeq \pi/4$ at $E_1 \simeq \Omega = 0.15$ ($\simeq 0.4\,\mathrm{eV}$), the "waterfall" is overcome by parabolic dispersion at $E_2 \simeq 0.33$ ($\simeq 0.85\,\mathrm{eV}$). The momentum of HEK k_0 decreases with doping. These results qualitatively agree with experimental data [4, 8] and appear to suggest that the band structure is of primary importance for the location and the doping dependence of the HEK, as the doping dependence of k_0 follows the shrinking of the Fermi surface. However, it should be taken into account that also the parameters of the MIR oscillator depend on doping. In particular, the quantitatively weaker dependence of k_0, observed in experiments could be related to the decrease of ω_0 with doping. The temperature dependence of $\Sigma(\omega)$ is very weak in the relevant region of energies. We have found a broader linewidth in the $(0, 1)$ and $(1, 0)$ directions as compared to the diagonal in agreement with the measurements presented in Ref. [16], indicating a relevant anisotropy in the structure of the HEK. These results can be described within the present model by introducing the anisotropy into the parameters of the oscillator. For instance, the difference in energy of the kink for momentum cuts along different

Fig. 5.4 Dispersion of the spectral density maximum along the diagonal direction, $K = K_x = K_y$, (*continuous line* $\delta = 0.15$, *dashed line* $\delta = 0.25$) and orthogonal direction, $K = K_x(K_y)$, $K_y(K_x) = 0$, (*dotted line* $\delta = 0.15$) of the Brillouin zone. Parameters are: $\omega_0 = 0.18$, $\gamma = 0.21$, $\lambda = 1.3$, $T = 0.001$
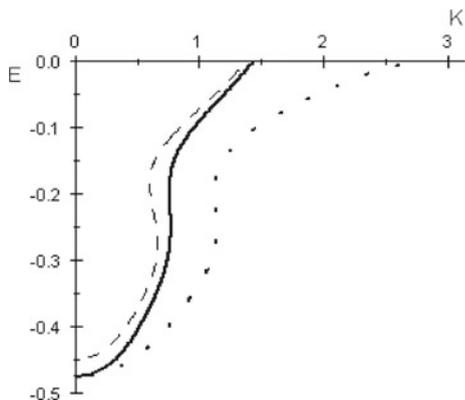
directions of the Brillouin zone found in Ref. [14] can be explained by a larger value
of Ω along the diagonal direction. In a different context, it has been noted that the
MIR peak in optics is affected by lattice oscillations [17]. It has also been suggested
[7] that the anomalous softening of the bond-stretching in-plane phonon mode is
related to an overdamped charge excitation in the same MIR part of the spectrum
and that the behavior of this mode could serve as an energy–momentum resolved
probe for the charge excitation, as we will discuss below. This raises an intrigu-
ing possibility that the same overdamped excitation is responsible for the phonon
anomaly and the HEK. One of its characteristics implied from the phonon analysis
is the anisotropy of the charge response at finite momentum in the energy range that
matches the grid-like pattern mentioned in Ref. [18]. Namely that anomalous phonon
softening occurs in the $(0, 1)$, $(1, 0)$ directions and a much smaller effect is observed
in the $(1, 1)$ direction. This anisotropy is related to a larger charge response located
at lower energies for the orthogonal direction as compared to more dispersive and
weaker response along the diagonal. The strength of the anisotropy depends on the
material and doping. Interestingly, in Refs. [8, 14] the energy of the kink (in absolute
value) in LBCO, BISCO 2212 was found to decrease significantly away from the
diagonal direction of the Brillouin zone. This behavior is also in agreement with the
MIR charge "boson" scenario proposed here. The broad incoherent spectrum above
the high-energy kink emerges due to the composite nature of electronic excitation
(in itself a source of finite linewidth) with emission of an "overdamped boson". This
explains the qualitative difference between the high- and low-energy kinks. In the
latter case, the linewidth of the boson is not more than 10% of its energy scale, while
in the case of MIR it is an order of magnitude larger. At the same time, for a compa-
rable coupling constant, such linewidth accounts for the verticality of the dispersion
and linewidth of the waterfall. As we have seen, the Im Σ remains almost flat in
this region. At energies around $0.9\,\text{eV}$ the Im Σ starts to decrease and the dispersion
recovers the bare band parabolic shape, albeit shifted to higher energies. It should
be mentioned that the optical conductivity data have served only for estimations
and, as discussed above, one needs to have a probe for momentum resolved charge
excitations in the 0.4–$0.5\,\text{eV}$ range to characterize the "overdamped boson".

It follows from the proposed scenario that HEK should exist in the ARPES of
other perovskite materials where the incoherent MIR excitation has been observed
(and vice versa). For instance, in Ref. [19] the ARPES analysis is focused on the
LEK at $60\,\text{meV}$ in bilayer manganite $La_{2-2x}Sr_{1+2x}Mn_2O_7$. However, also the data
for higher energies are presented, where the existence of the other kink at around
$0.45\,\text{eV}$ is clearly visible. It would also be interesting to carry out a similar search for
the HEA in ARPES experiments on a nonmagnetic perovskite such as $Ba_{1-x}K_xBiO_3$
with a relatively high T_c that is also known to have an incoherent charge excitation
in the MIR about 0.4–$0.6\,\text{eV}$ as well as anomalous softening of the bond-stretching
phonon mode [20].

The BS Phonon Softening

Whether the high-energy kink is related to a boson strongly interacting with electrons, in analogy with low-energy kink, or it is due to another mechanism is yet unclear. However, we have noted a striking universality of two anomalous features belonging to the energy scales under question and appearing simultaneously in a large number of metallic perovskite materials (cuprates, bismuthates, manganites, nickelates, ruthenates [20]), i.e., a strong softening of bond-stretching phonon mode and a broad structure in the MIR region of the charge response. One can also note the similarity of their doping dependence when comparing optical absorption experiments [21] and unpolarized neutron scattering [22]. Thus, we have further investigated the possibility of connection between the BS phonon mode and the charge response in the MIR, proposing a semi-phenomenological model in which the charge mode at high energy is coupled to phonons, causing its anomalous softening. Namely, we have assumed that renormalization of the phonon Green's function can be described by a polarization function of the form

$$P(q, \omega) = \frac{\beta_q}{\omega^2 - \Omega_q^2 + i\Gamma_q\omega}, \tag{5.4}$$

with the oscillator strength β_q, dispersion Ω_q, and width Γ_q of the charge mode to be determined from experimental data on the phonon linewidth γ_q and dispersion ω_q. Note that the electron–phonon coupling constant can be absorbed into β_q. Although the damped oscillator form of $P(q, \omega)$ is much more simple than the one found in microscopic approaches [23, 24], it is the most commonly used in analyzing, e.g., optical experiments. However, our aim has not been to fit experimental data, but to validate the use of such parametrization. As one can immediately see from Eq. (5.4), the known behavior of γ_q and ω_q is insufficient even for simple fitting of the three unknown parameters. Nevertheless, one can easily show that even at this stage the model, Eq. (5.4), sets some constraints that provide useful information on the experiments. Indeed, by considering the real and imaginary parts of polarization $P(q, \omega)$ one obtains a relation

$$\Gamma_q = \gamma_q \frac{\Omega_q^2 - \omega_q^2}{\omega_0^2 - \omega_q^2}. \tag{5.5}$$

Taking the typical values for $\omega_q \sim 100\,\text{meV}$ and $\gamma_q \sim 10\,\text{meV}$, one can now check if Eq. (5.5) gives back typical values for the structures observed in EELS or optics. A priori it is not guaranteed that this constraint should be satisfied at all. For instance, by inserting the energy of optical plasmon $\Omega_{pl} \sim 1 - 2\,\text{eV}$ into Eq. (5.5) we find $\Gamma_{pl} \gtrsim 3\,\text{eV}$, in clear contradiction with experiment. When the typical values for the MIR response are substituted ($\Omega \sim 0.5\,\text{eV}$, $\Gamma \sim 0.7\,\text{eV}$), the constraint is satisfied with a surprising accuracy. The same is confirmed by analyzing the doping variation of these parameters in several materials summarized in Ref. [25].

Fig. 5.5 The effect of
isotope substitution on the
dispersion of the
bond-stretching phonon
mode in the (100) or (010)
directions of the Brillouin
zone. In contrast, the
horizontal lines represent the
uniform shift expected for
the diagonal direction (110)
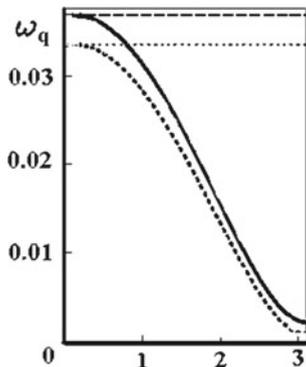
Fig. 5.6 The effect of
isotope substitution on the
phonon linewidth in the
(100) or (010) directions of
the Brillouin zone
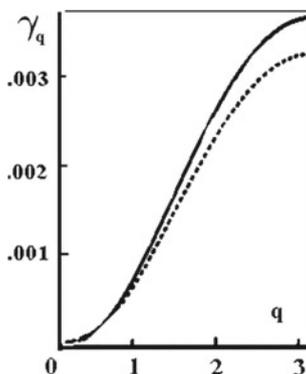
To derive new constraints following from the semi-phenomenological model, we note that closer are in energy the two interacting excitations, the larger is their mutual impact. Thus, if one changes the phonon frequency by isotope substitution, it could be possible to obtain the required equations, for the unknown parameters.

In Figs. 5.5 and 5.6 we show the typical change of the phonon dispersion and linewidth due to isotope substitution $\left(dm = m \left({}^{18}O \right) - m \left({}^{16}O \right), d\omega_q = \omega_q \left({}^{18}O \right) - \omega_q \left({}^{16}O \right) \right)$.

In particular, one can define a momentum dependent isotope coefficient $\bar{\alpha}_q = d\omega_q / (\omega_0 d \ln m)$ which, according to the model, gives

$$\bar{\alpha}_q \simeq \alpha_0 \sqrt{1 - \frac{\beta_q \sin^2 (q_x/2)}{\Omega_q^2}}, \tag{5.6}$$

where we have assumed for simplicity that terms $\sim (\omega_q/\Omega_q)$ can be neglected. It is also easy to obtain the change in phonon linewidth induced by such substitution

$$\frac{d\gamma_q}{\gamma_q} \simeq -2\alpha_0 \frac{dm}{m}. \tag{5.7}$$

However these equations do not provide a third independent relation, although they can serve for the consistency check of the model. As explained above, we expect the main new effect to be the inhomogeneous deformation of the phonon dispersion in the (100) and (010) directions due to isotope substitution. This suggests a new independent quantity defined as

$$\Delta\alpha_q \equiv \alpha_q - \alpha_0, \tag{5.8}$$

which results in the following expression

$$\Delta\alpha_q/\alpha_0 = \beta_q \sin^2(q/2)\,\omega_0^2 \frac{\left(\Gamma_q/\Omega_q\right)^2 - 1}{\Omega_q^4}. \tag{5.9}$$

An interesting consequence of this relation is that the sign of the coefficient corresponds to the ratio of the linewidth and dispersion. A similar coefficient can also be derived for the change of linewidth of the phonon.

One can now see that the number of equations is overcomplete and allows not only to find all the parameters of the charge response, but also test the consistency of the model. Clearly, the model can be further generalized to include additional oscillators. Nevertheless there are good reasons to expect that the MIR structure contributes to the renormalization of the bond-stretching phonon. In this case, the latter could become a unique probe of momentum resolved charge response at these energies in perovskite materials.

Conclusions

We have reviewed a semi-phenomenological model that can explain both the origin of the high-energy kink observed in photoemission experiments and the anomalous bond-stretching phonon mode softening in cuprate superconductors and some other metallic perovskite oxides. It has been argued that both the high-energy anomaly and the phonon softening are caused by the interaction with an overdamped charge bosonic mode in the mid-infrared region of the optical spectrum. We have shown that a large damping of the bosonic mode is an essential feature responsible for the "waterfall" dispersion and linewidth of the spectral weight in ARPES experiments. On the other side, we have proven that the energy scale of this charge mode is several times larger than the phonon frequency, i.e., within 0.2–0.6 eV, and its doping and momentum dependence is discussed in relation to experimentally observed behavior of the BS phonons. Experiments on the effect of oxygen isotope substitution on the phonon softening and linewidth are suggested as a source of important information

on the charge excitations, and equations describing isotope effects have been derived. The set of equations is overcomplete and can therefore serve as a consistency test for the model. The proposed mid-infrared scenario could offer an explanation of the high-energy anomaly and phonon softening in different perovskites in a unified framework.

References

1. A. Lanzara, P.V. Bogdanov, X.J. Zhou, S.A. Kellar, D.L. Feng, E.D. Lu, T. Yoshida, H. Eisaki, A. Fujimori, K. Kishio, J.I. Shimoyama, T. Noda, S. Uchida, Z. Hussain, Z.X. Shen, Nature **412**(6846), 510 (2001). DOI 10.1038/35087518
2. G.H. Gweon, T. Sasagawa, S.Y. Zhou, J. Graf, H. Takagi, D.H. Lee, A. Lanzara, Nature **430**(6996), 187 (2004). DOI 10.1038/nature02731
3. B.P. Xie, K. Yang, D.W. Shen, J.F. Zhao, H.W. Ou, J. Wei, S.Y. Gu, M. Arita, S. Qiao, H. Namatame, M. Taniguchi, N. Kaneko, H. Eisaki, K.D. Tsuei, C.M. Cheng, I. Vobornik, J. Fujii, G. Rossi, Z.Q. Yang, D.L. Feng, Phys. Rev. Lett. **98**, 147001 (2007). DOI 10.1103/PhysRevLett.98.147001
4. J. Graf, G.H. Gweon, K. McElroy, S.Y. Zhou, C. Jozwiak, E. Rotenberg, A. Bill, T. Sasagawa, H. Eisaki, S. Uchida, H. Takagi, D.H. Lee, A. Lanzara, Phys. Rev. Lett. **98**, 067004 (2007). DOI 10.1103/PhysRevLett.98.067004
5. J. Lee, K. Fujita, K. McElroy, J.A. Slezak, M. Wang, Y. Aiura, H. Bando, M. Ishikado, T. Masui, J.X. Zhu, A.V. Balatsky, H. Eisaki, S. Uchida, J.C. Davis, Nature **442**(7102), 546 (2006). DOI 10.1038/nature04973
6. R. Citro, S. Cojocaru, M. Marinaro, Phys. Rev. B **73**, 014527 (2006). DOI 10.1103/PhysRevB.73.014527
7. S. Cojocaru, R. Citro, M. Marinaro, Phys. Rev. B **75**, 220502 (2007). DOI 10.1103/PhysRevB.75.220502
8. T. Valla, T.E. Kidd, W.G. Yin, G.D. Gu, P.D. Johnson, Z.H. Pan, A.V. Fedorov, Phys. Rev. Lett. **98**, 167003 (2007). DOI 10.1103/PhysRevLett.98.167003
9. M.L. Kulic, O.V. Dolgov, ARPES spectra of Bi2212 give the Coulomb coupling $\lambda^C \approx 1$ and the electron-phonon coupling $\lambda^{EP} = 2 - 3$ (2006). arXiv:cond-mat/0610549
10. G.D. Mahan, *Many-Particle Physics*, 2nd edn. (Plenum Press, New York and London, 1990)
11. F. Marsiglio, M. Schossmann, J.P. Carbotte, Phys. Rev. B **37**, 4965 (1988). DOI 10.1103/PhysRevB.37.4965
12. F. Dogan, F. Marsiglio, Electron-phonon vs electron-impurity interactions with small electron bandwidths (2006). arXiv:cond-mat/0603635
13. X.J. Zhou, T. Cuk, T. Devereaux, N. Nagaosa, Z.X. Shen, in *Handbook of High-Temperature Superconductivity: Theory and Experiment*, ed. by J.R. Schrieffer, J.S. Brooks (Springer New York, New York, 2007), pp. 87–144. DOI 10.1007/978-0-387-68734-6_3. arXiv:cond-mat/0604284
14. J. Chang, S. Pailhés, M. Shi, M. Månsson, T. Claesson, O. Tjernberg, J. Voigt, V. Perez, L. Patthey, N. Momono, M. Oda, M. Ido, A. Schnyder, C. Mudry, J. Mesot, Phys. Rev. B **75**, 224508 (2007). DOI 10.1103/PhysRevB.75.224508

15. J. Hwang, E.J. Nicol, T. Timusk, A. Knigavko, J.P. Carbotte, Phys. Rev. Lett. **98**, 207002 (2007). DOI 10.1103/PhysRevLett.98.207002
16. Z.H. Pan, P. Richard, A.V. Fedorov, T. Kondo, T. Takeuchi, S.L. Li, P. Dai, G.D. Gu, W. Ku, Z. Wang, H. Ding, Universal quasiparticle decoherence in hole- and electron-doped high-T_c cuprates (2006). arXiv:cond-mat/0610442
17. P. Piekarz, T. Egami, Phys. Rev. B **72**, 054530 (2005). DOI 10.1103/PhysRevB.72.054530
18. D.S. Inosov, J. Fink, A.A. Kordyuk, S.V. Borisenko, V.B. Zabolotnyy, R. Schuster, M. Knupfer, B. Büchner, R. Follath, H.A. Dürr, W. Eberhardt, V. Hinkov, B. Keimer, H. Berger, Phys. Rev. Lett. **99**, 237002 (2007). DOI 10.1103/PhysRevLett.99.237002
19. Z. Sun, Y.D. Chuang, A.V. Fedorov, J.F. Douglas, D. Reznik, F. Weber, N. Aliouane, D.N. Argyriou, H. Zheng, J.F. Mitchell, T. Kimura, Y. Tokura, A. Revcolevschi, D.S. Dessau, Phys. Rev. Lett. **97**, 056401 (2006). DOI 10.1103/PhysRevLett.97.056401
20. M. Braden, L. Pintschovius, T. Uefuji, K. Yamada, Phys. Rev. B **72**, 184517 (2005). DOI 10.1103/PhysRevB.72.184517
21. D.N. Basov, T. Timusk, Rev. Mod. Phys. **77**, 721 (2005). DOI 10.1103/RevModPhys.77.721
22. L. Pintschovius, Physica Status Solidi B **242**(1), 30 (2005). DOI 10.1002/pssb.200404951
23. G. Khaliullin, P. Horsch, Phys. Rev. B **54**, R9600 (1996). DOI 10.1103/PhysRevB.54.R9600
24. D.N. Aristov, G. Khaliullin, Phys. Rev. B **74**, 045124 (2006). DOI 10.1103/PhysRevB.74.045124
25. F. Hanke, M. Azzouz, Moroccan Journal of Condensed Matter **6**, 1 (2005)

Chapter 6
Electrides and Their High-Pressure Chemistry

Xiao Dong and Artem R. Oganov

Abstract Recently, electrides were discovered in many systems (especially those containing alkali and alkali earth metals) at high pressures. An electride can be defined as an ionic compound where the role of an anion is played by a strongly localized electron density. High-pressure emergence of electrides is due to the Pauli expulsion of valence electrons from the core, while some electrides are better described as originating from multicenter covalent bonds. Not being bound directly to a nucleus, these localized electrons are chemically active, making electrides the strongest reducing agents known, able to interact even with such an extremely inert element as helium.

The Discovery of Electrides

In classical chemical systems, electron density is peaked on nuclei. However, in the 1970–1980s, a new type of compounds was established, in which bare electrons, not bound to any particular nucleus, are concentrated in the interstitial space and behave as anions. Such compounds were named electrides.

X. Dong (✉)
Center for High Pressure Science and Technology Advanced Research,
Beijing 100193, China
e-mail: xiao.dong@hpstar.ac.cn

A.R. Oganov
Skolkovo Innovation Center, Skolkovo Institute of Science and Technology,
3 Nobel Street, Moscow 143026, Russia
e-mail: a.oganov@skoltech.ru

A.R. Oganov
State University of New York, Stony Brook, NY 11794-2100, USA

A.R. Oganov
Moscow Institute of Physics and Technology, 9 Institutskiy Lane, Dolgoprudny,
Moscow Region 141700, Russia

A.R. Oganov
International Center for Materials Design, Northwestern Polytechnical University,
Xi'an 710072, China
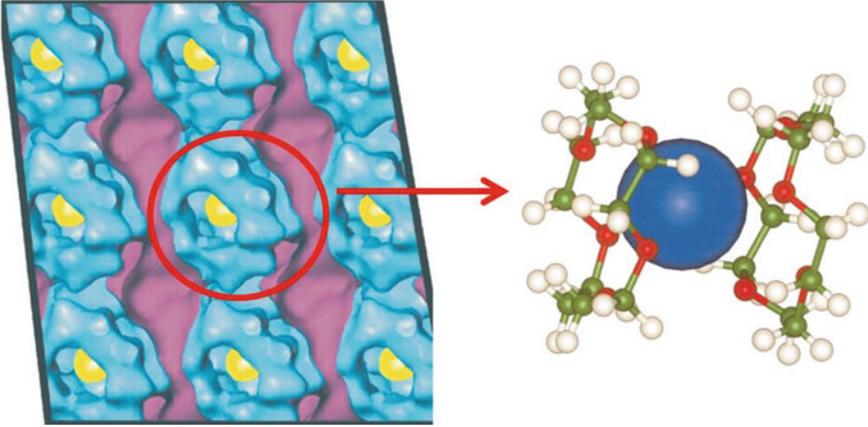
Fig. 6.1 Structure of the organic electride $Cs^+(15\text{-}crown\text{-}5)_2e^-$. Electron-trapping cavities and channels are shown in *pink*, 15-crown-5 molecules in *blue*, and cesium cations in *yellow*. Inset shows the ball-and-stick representation of the $Cs^+(15\text{-}crown\text{-}5)_2$ "sandwich," with the Cs^+ ion drawn to scale [1] (color figure online)

The first clear report of an electride we could find dates back to 1981 [2], followed by many other discoveries of chemically complex compounds (at first, only organic, but then also inorganic [3]). In 1990, Dye [1, 4, 5] discovered that an organic crown ether or cryptand can trap electrons as counterions, in alkali metal complexes, such as $Cs^+(18\text{-}crown\text{-}6)_2e^-$ or $Cs^+(15\text{-}crown\text{-}5)_2e^-$. In such electrides as shown in Fig. 6.1, the localized electrons are usually single electrons occupying the cavities, and there is a weak spin–spin interaction between electrons occupying neighboring cavities. So, electrides known at normal pressure are paramagnetic or antiferromagnetic.

Since high pressure favors spin pairing, such spin-polarized electrides will obviously give way to spin-paired electrides under compression. Another insight into the nature of high-pressure electrides comes from a quantum-mechanical consideration of the interaction between valence and core electrons.

In 2008, Rousseau and Ashcroft [6] modeled the effects of core electrons as a hard potential $v(r) = V_0\theta(r_c - r)$, where r_c is the ionic radius, θ is the Heaviside step function and V_0 a hard screening potential. Using this potential, they wrote the Hamiltonian and the Schrödinger equation:

$$H(r) = -\frac{\hbar^2}{2m}\nabla^2 + \sum_R v(r - R), \tag{6.1a}$$

$$H\varphi = E\varphi, \tag{6.1b}$$

where R is the position of nucleus. To consider the effect of compression, they also employed an average radius of ions, $r_s = \sqrt[3]{3\Omega/4\pi N}$, where N is the number of atoms and Ω the volume. In this way, the parameter r_c/r_s can describe the compression ratio.
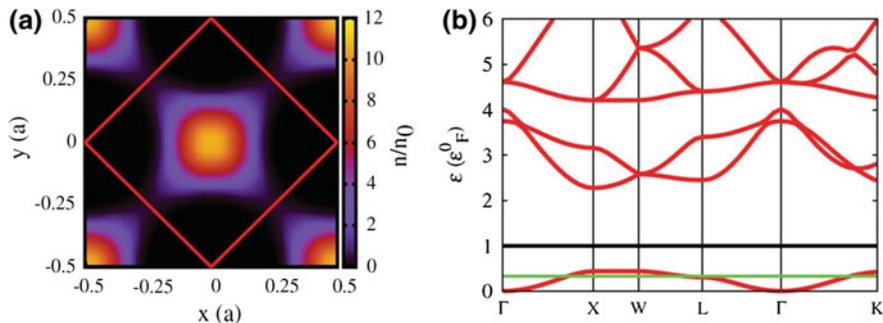
Fig. 6.2 Electronic properties calculated from Eq. (6.1b) with $r_c/r_s = 0.7$. **a** Electron density in the $z = 0$ plane. **b** Band structure (eigenvalue spectrum) for excluding spheres. The Fermi energy (*green line*) and the free electron Fermi energy (*black line*) are indicated [6] (color figure online)

Solving numerically the Schrödinger's equation for the face-center cubic (fcc) structure, they found that when $r_c/r_s > 0.5$, the wave function will deviate significantly from the nearly free electron gas behavior and localize in the interstitial regions (Fig. 6.2a). In the band structure, the band width decreases with compression, which means the electrons are more localized (Fig. 6.2b). Already this oversimplified model gives an enlightening idea that electropositive atoms at strong compression may experience band narrowing, interstitial localization of valence electrons, and metal–insulator transition. The first anticipation of demetallization of alkali metals under pressure came from Professor Pucci's group in 1997 [7], although not through the electride mechanism.

Then, Ma and Oganov [8] found a true high-pressure electride, transparent $hP4$ phase of sodium. In the high-pressure experiment, metallic sodium transformed under compression first to a poorly metallic incommensurate $tI19$ phase (black at 156 GPa, Fig. 6.3a) and then to the transparent phase $hP4$. In the $hP4$ structure, there are two kinds of Na atoms, Na1 and Na2, in total four atoms of Na and two interstitial sites with electron localization in the unit cell. Each interstitial site contains an electron pair, and the chemical formula can be written as $Na_2(2e)$. From structural geometry, one could imply pd and pd^2 hybridizations for Na1 in the octahedral sites and Na2 in the triangular prismatic coordination, respectively (Fig. 6.3b). It has a wide band gap—for example, GW calculations predict the band gap to be over 6 eV at 600 GPa, which implies a colorless transparent insulator.

Similar electrides were also found for Li [9–12]. However, lithium atom has two differences from sodium: (1) the core radius (i.e., the ionic Li^+ radius, 0.76 Å) is much smaller than its covalent radius (1.28 Å) and van der Waals radius (1.82 Å). This means Li is quite compressible and easy to get into the electride state—compared with Na (which becomes an electride at pressures ~ 200 GPa), Li only needs 60 GPa to get into the semiconducting phase $C2cb$–40 [10]. (2) Li is much smaller than Na, thus the interstitial electrons are localized in a smaller space and nearer to each other, which means that electrons can more easily tunnel from one interstitial site to
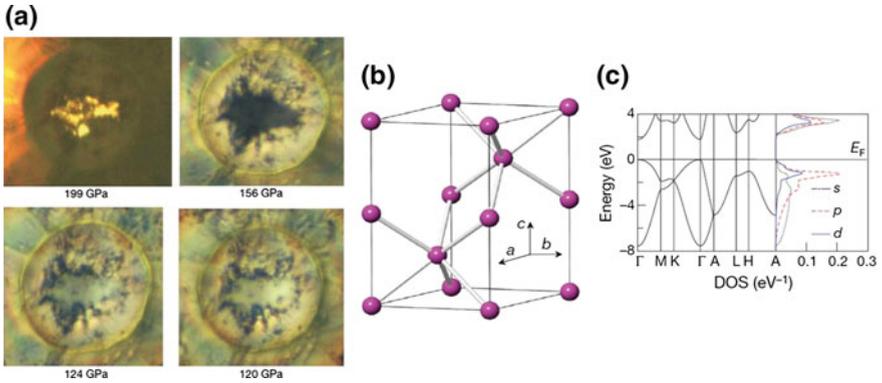
Fig. 6.3 **a** Photographs of the Na sample taken under pressure in the diamond anvil cell (DAC), showing an optically transparent sample at 199 GPa. **b** Crystal structure of $hP4$-Na (space group $P6_3/mmc$). **c** Band structure and partial densities of states (DOS) at 300 GPa. From Ref. [8]

another, and over the potential barrier formed by the repulsion from core electrons. As a result, the electride state is sensitive and easy to break: e.g., the band gap is quite small (0.8 eV) and the external pressure causes a phase transition to the metallic $Cmca$-24 phase at 90 GPa [10].

Besides elementary substances, electride states were reported for Al [13], Mg_3O_2 [14], Na_3Cl [15], and Na_2He [16], SiO [17]. Many of these compounds have chemical formulas very different from those known at zero pressure.

The Effect of Core Electron and Anti-Wilson Phase Transition

The free electron gas model is a convenient tool to show the expected effect of band broadening under pressure. Besides being convenient, this model becomes exact (if one neglects nuclear strong and weak interactions) for any chemical system at infinite compression. In this model, the electronic dispersion relation is $E_k = \hbar^2 k^2 / 2m$. Here, the wave vector is in the momentum space, $k \in (-\pi/a, \pi/a)$, where a is the lattice parameter. So the band width $\Delta E_k = \hbar^2 \pi^2 / 2ma^2$. With pressure reducing the lattice parameter, the range of both k and ΔE will increase, which is ruled by Heisenberg's uncertainty principle. This means that in this simple theoretical model, Heisenberg's uncertainty principle requires that the band width of solids increases with pressure.

For a semiconductor or insulator at high pressure, the broadening band will occupy the energy space of forbidden band, thereby decreasing the band gap. If the pressure is sufficiently high, the conduction band and valence band will finally overlap, and the insulator–metal transition will occur, see Fig. 6.4. This mechanism was discovered by Wilson, after whom it is named [18].
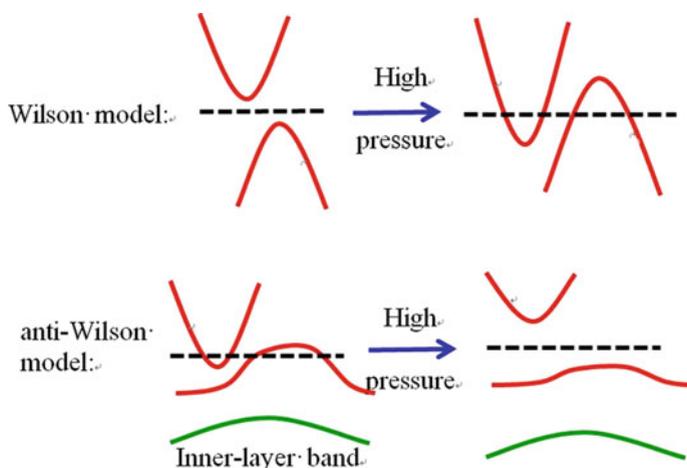
Fig. 6.4 Evolution of the band structure under pressure in Wilson and anti-Wilson models

The band overlap originating from the Wilson mechanism affects the electronic states near Fermi surface, and is normally accompanied with the strengthening of electron–phonon coupling to cause the Peierls distortion, or other more complex phonon and electron effects, often causing complex phase transitions. The related pressure-induced insulator–metal phase transition is usually called the Wilson transition.

Wilson transition is indeed a common phenomenon at high pressure, but based on a crude theoretical model of free electron gas, which ignores the effect of atomic cores—i.e., strong Pauli and electrostatic repulsion of valence electrons from the core orbitals. In quantum mechanics, valence orbitals must be radially orthogonal with core electrons of the same angular momentum—and can be thought to be not "allowed" to get close to the core due to the Pauli exclusion principle. If pressure is sufficiently high, most of the free volume for such electrons will disappear and such "expelled" electrons will have to localize in the intersticial space. This should be accompanied by band narrowing—instead of band broadening, i.e., exactly opposite to the Wilson model. As a result, with pressure increasing, one can often observe the transformation of the initial metallic state into a poor metal with a decreased density of states (DOS) at the Fermi level (N_F), such as tI19-Na [19] and $CaLi_2$ [20] or opening of the band gap, as in elemental Na [8] and Li [10]. However, at sufficiently high pressures a reentrant metallization must occur: for example, Li and Na become metal again at 95 GPa [10] and 15.5 TPa [21], respectively. At pressures of order of 1 a.u. (29.4 TPa), we expect the disappearance of the periodic law, which will be replaced with the physics of the free electron gas to govern the behavior of all compounds. At further compression, when the internuclear distances become much smaller than the electronic de Broglie wavelength, electrons stop feeling nuclear attraction—this is what happens in white dwarfs (pressures of order 10^{22} Pa) [22].

Physical Form: Duality of Pseudoatoms and Multicenter Bonds

Pseudoatoms

In classical crystal chemistry, there are four main types of bonding: covalent, metallic, ionic, and van der Waals. Bader theory [23] is one of the ways to describe bonding quantitatively. Electronic charge density $\rho(x, y, z)$ is continuous throughout the space, and has stationary points with $\partial\rho/\partial x = \partial\rho/\partial y = \partial\rho/\partial z = 0$, i.e., $\nabla\rho = 0$. Considering second derivatives, density minima have $\partial^2\rho/\partial x^2$, $\partial^2\rho/\partial y^2$, $\partial^2\rho/\partial z^2 > 0$, and maxima have negative second derivatives (and maxima coincide with the positions of the nuclei—although in a static picture, density maxima at nuclei have a cusp, rather than zero derivative of the density). Saddle points, called bond critical points in Bader theory, can have negative Laplacian ($\nabla^2\rho < 0$) when the electrons are locally concentrated and shared by both nuclei (usually implying covalent interactions) or positive Laplacian (i.e., $\nabla^2\rho > 0$) when the electrons are concentrated around the atoms (implying closed-shell interactions).

However, for electrides a different picture emerges: they display significant charge density maxima not corresponding to positions of nuclei, as shown in Figs. 6.2a, 6.5 and 6.6a. This implies that electronic concentration in electrides is much stronger than in a normal covalent bond and even reverses the sign of the second derivative to create a local maximum without a nucleus. The effect of electron concentration can be clearly seen also in the deformation charge density [24] and electron localization function (ELF) [25]. Deformation charge density is the difference between the actual charge density and superposition of non-interacting atomic charge densities. As shown in Fig. 6.6b, interstitial voids get the greatest increase of electron density—

Fig. 6.5 Electronic charge density in a normal bond and in an electride
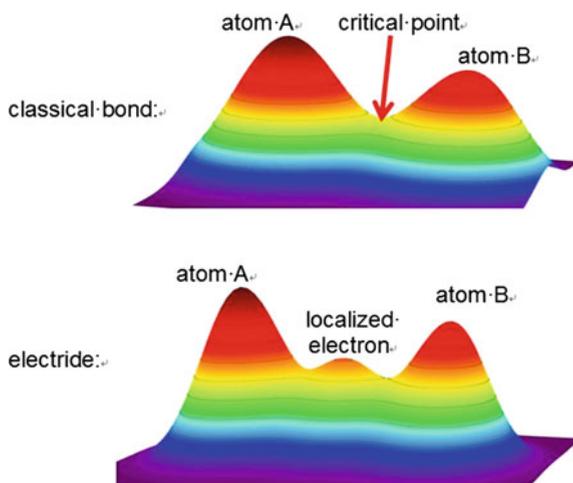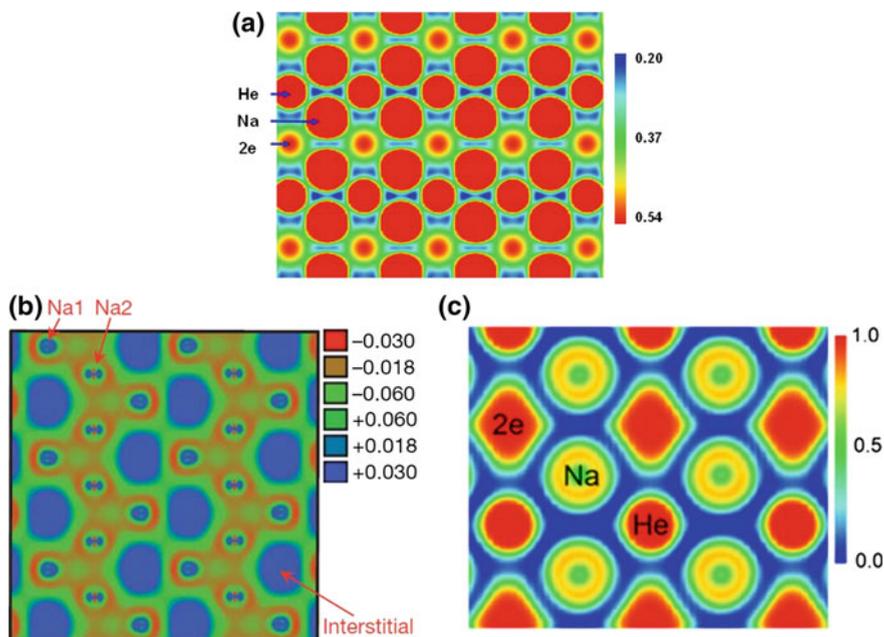
Fig. 6.6 **a** Computed charge density ($e\,\text{Å}^{-3}$) of Na_2He at 300 GPa, plotted in the [110] plane of the conventional cell [16]. **b** Deformation charge density of $hP4$-Na plotted in the [110] plane [8]. **c** ELF of Na_2He in the [110] plane [16]

in other words, when an electride state is formed, valence electrons concentrate in the interstitial space. ELF distributions give a similar picture. ELF is defined in the range [0, 1], with ELF = 1 corresponding to perfect localization, and ELF = 0.5 to electron gas. In the electride Na_2He (Fig. 6.6c), the interstitial electron has almost ELF = 1, even higher than for the core $2s$ and $2p$ electrons of Na.

There are rather many elemental solids and compounds which have nonnuclear charge density maxima. Only some of them have band gaps, while others are metals. In metals, electron gas very effectively screens long-range electrostatic interactions. Whether the system is metallic or not, depends on the height of the potential barrier and the ability of the electrons to tunnel through it. For example, d-electrons penetrate the core regions of K, Rb, and Cs, while core radius is too small to isolate the localized electrons in some Li compounds such as Li_6O. A proper distinction would be between: (1) strong electrides (such as Na, Li, and Mg_3O_2), defined as insulators or semiconductors with localized electrons playing the role of an anion in an ionic crystal. (2) weak electrides (such as K, Ca, Mg, Al, Li_6O), defined as metallic systems with non-nuclear charge density maxima.

The electride state has been proven in many compounds and the origin of its stability is interesting to explore. In quantum mechanics, without nuclear attraction, electronic wave function will delocalize throughout the entire space. This is obviously in conflict with the charge density distributions observed in electrides. To reconcile this contradiction, we add a volumetric constraint in the quantum system, and discover another picture, quite different from the electron gas model. Here we use a simplified wave packet model.

Related to the uncertainty principle, there is the Robertson-Schrödinger inequality,

$$\langle A^2 \rangle \langle B^2 \rangle \geq \frac{1}{4} |\langle [A, B] \rangle|^2, \tag{6.2}$$

so

$$\langle i^2 \rangle \langle p_i^2 \rangle \geq \frac{\hbar^2}{4}, \quad i = x, y, z, \tag{6.3}$$

where p_x, p_y, p_z is the momentum of the particle, and $\langle \bar{x} \rangle = \langle \bar{y} \rangle = \langle \bar{z} \rangle = 0$, $\langle \bar{p}_x \rangle = \langle \bar{p}_y \rangle = \langle \bar{p}_z \rangle = 0$.

If the particle is localized with a volumetric constraint, then

$$\langle x^2 \rangle \langle y^2 \rangle \langle z^2 \rangle \leq V, \tag{6.4}$$

where V is a constant, which means the wave function cannot extend infinitely. So the kinetic energy is

$$
\begin{aligned}
E_k &= \frac{1}{2m} \left(\langle p_x^2 \rangle + \langle p_y^2 \rangle + \langle p_z^2 \rangle \right) \\
&\geq \frac{3}{2m} \sqrt[3]{\langle p_x^2 \rangle \langle p_y^2 \rangle \langle p_z^2 \rangle} \\
&\geq \frac{3\hbar^2}{8m} \left(\langle p_x^2 \rangle \langle p_y^2 \rangle \langle p_z^2 \rangle \right)^{-1/3} \\
&\geq \frac{3\hbar^2}{8m} V^{-1/3}.
\end{aligned}
\tag{6.5}
$$

If and only if $\langle p_x^2 \rangle = \langle p_y^2 \rangle = \langle p_z^2 \rangle$ (spherically symmetric localization), and $\langle i^2 \rangle \langle p_i^2 \rangle = \hbar^2/4$ ($i = x, y, z$) (Gaussian wave packet), the kinetic energy reaches its minimum. Next, we take a Gaussian wave packet as an example. We assume that the wave packet has anisotropy, which means it is an ellipsoid with A, B, C as its principal axes, where $A = 1/2\langle x^2 \rangle$, $B = 1/2\langle y^2 \rangle$, $C = 1/2\langle z^2 \rangle$. So, its wave function is

$$\varphi(x, y, z) = \left(\frac{ABC}{\pi^3} \right)^{1/4} e^{-\frac{1}{2}(Ax^2 + By^2 + Cz^2)}, \tag{6.6}$$

and its kinetic energy is

$$
\begin{aligned}
E_k = \langle \varphi | E | \varphi \rangle &= -\frac{\hbar^2}{2m} \int \varphi^* \nabla^2 \varphi \, dx \, dy \, dz \\
&= \frac{\hbar^2}{4m}(A + B + C) \\
&\geq \frac{3\hbar^2}{4m}(ABC)^{1/3} \\
&= \frac{3\hbar^2}{8m} \left(\langle x^2 \rangle \langle y^2 \rangle \langle z^2 \rangle \right)^{-1/3}.
\end{aligned}
\tag{6.7}
$$

Only if $A = B = C$, the inequality becomes an equality.

In simple words, if there is a limited volume available to the electron, Heisenberg's uncertainty principle requires an increased uncertainty in the momentum space, increasing the kinetic energy of the electron and leading to its localization as a spherical Gaussian wave packet (to decrease this kinetic energy). In a word, volumetric constraint and uncertainty principle are the driving forces behind the formation of nonnuclear charge density maxima, characteristic of electrides.

With the localization as spherical Gaussian wave packets, such bare electrons will behave as atomic particles—anions—and we can consider them as independent particles. Compared to true atoms with central potentials from nucleus attraction, interstitial electrons are compressible and easy to reshape. At high pressure, interstitial electrons are usually spin-paired, which often means that an electron pair occupies one interstitial area. It has a radius somewhat smaller than He atom, which also contains one electron pair. For example, in Na_2He compound, the radius of $2e$ is about 0.9 of the radius of He.

Considering the interstitially localized electrons as pseudoatoms, we can evaluate their interaction with cations and other anions. Normally, electrides have structures similar to either alloys or ionic phases: for example, $hP4$-Na has the Ni_2In structure, where Na atoms occupy the positions of Ni atoms and interstitial electrons occupy the In sites. The $hP4$ structure can be viewed as a strongly ($\sim$ twofold) squeezed along the c-axis double hexagonal close packing ($\dots ABAC \dots$) of Na atoms, or as a nearly perfect hexagonal close packing of the localized electrons. This implies that interstitial electrons have a nearly isotropic interaction, leading to close-packed structure in this and other examples.

In 2014, Miao and Hoffmann [26] discovered a method to calculate the orbital energies in atoms under pressure: they used a sufficiently large ($3 \times 3 \times 3$) supercell of the perfect He fcc structure relaxed at a given pressure, where they consider He as a chemically inert space filler and pressure transmitting medium. In that supercell, they replaced the central He atom by the atom of interest, relaxed the structure, and computed its electronic DOS. Assuming the highest occupied molecular orbital (HOMO) of the helium matrix fixed, irrespective of the inserted atom, one can get the relative orbital energies at high pressure. Considering localized electrons as a new
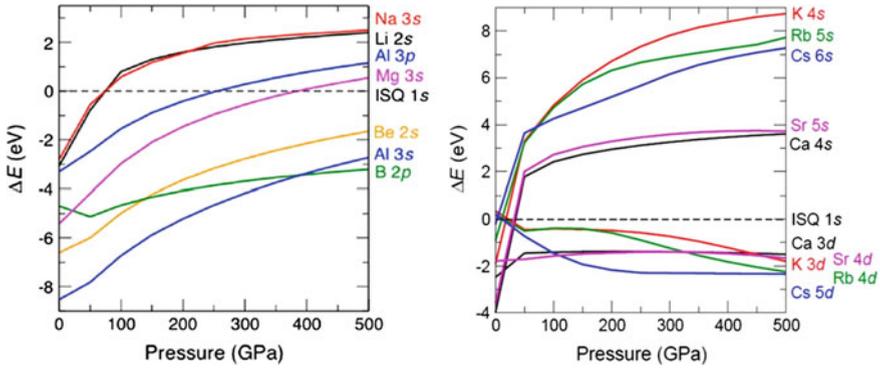
Fig. 6.7 Orbital energy of an electron on the corresponding orbital [26]. ISQ stands for "interstitial quasi-atom", which is set as reference zero energy

type of element, they also replaced the central He atom with an additional electron to represent the electrides.

Using the He matrix method described above, one can get the orbital energies of different elements (Fig. 6.7). At hundreds of GPa, the highest occupied orbitals of Na, Li, Al, and Mg exceed the orbital level of interstitial electrons, which means the donation of electrons to the interstices is energetically favorable.

For heavier akali metals, such as K, Rb, and Cs, high-pressure behavior is quite different from light alkalis Li and Na. The difference comes from d-orbital effects and s-d transfer [27]. At zero pressure, n s-orbitals with $n-1$ nodes have lower energy than $(n-1)$ d-orbitals with $(n-3)$ nodes, but at high pressure, $(n-1)$ d-orbitals become energetically more favorable and become occupied. So, the n s-electron will transfer to the $(n-1)$ d-orbital, making heavy alkali metals effectively transition metals under pressure and, in some way, rearranging the Periodic Table [22]. For example, the electronic configuration of K changes from [Ar]$4s^1$ to [Ar]$3d^1$. Additionally, with less Pauli exclusion from core electrons than ns electrons, $(n-1)$ d-electrons are more localized in the core and easier to tunnel the energy barrier formed by overlapped core electrons at high pressure. As a result, at high pressure, K, Rb, and Cs, although they do have (diminished) nonnuclear charge density maxima, are metallic. The core localization of d-electrons and s-d transfer also lead to changes in the chemical activity under pressure. This s-d transfer happens easily in K and heavier alkalis: $4s$-$3d$ orbital transition requires moderate energy, whereas in Na the $3s$-$3d$ transition is energetically very costly and s-d transition is never complete (even more so in Li). So the light active elements, such as Li, Na, very easily adopt the electride state and can form semiconducting and insulating phases at high pressure, but K, Rb and Cs cannot.

Multicenter Bonds

In the above picture of valence electrons being "squeezed out" by core electrons, the interatomic distances should be sufficiently small—smaller than the sum of the core radius of one atom and valence orbital radius of the other. For example, the shortest Na–Na distance in $hP4$-Na is 1.89 Å at 300 GPa, and the $3s$ orbital radius of Na is 1.71 Å, whereas the size of the core is best modeled by the ionic radius of Na^+ (1.02 Å). This indicates a strong core–valence overlap and core–core overlap. Due to short distances from nonnuclear density maxima to neighboring atoms, these localized electrons can be considered as a multicenter bond formed at high pressure—however, this description seems even more appropriate in those cases where interatomic distances are longer than sum of core and valence radii (for example, in the high-pressure compound SiO [17]).

Here we use a simplified model, Na_8 cluster, to investigate this idea. Put eight Na atoms in the vertices of a cube, then we get a Na_8 cluster with symmetry group O_h. We set a as the edge length, which is also the shortest bond length in the Na_8 cluster. Every Na atom has 1 valence electron on the $3s$ orbital and Na_8 cluster has eight valence electrons and four new hybridized bonding orbitals. Because of the high symmetry of O_h group, these four orbitals can be represented as 1 a_{1g} orbital and 3 triply degenerated t_{1u} orbitals. Specifically, the a_{1g} orbital is nearly spherical and has s-wave symmetry, while the t_{1u} orbital has p-wave symmetry. The whole Na_8 cluster is a quantum dot and behaves like a superatom.

With the bond length corresponding to atmospheric pressure, $a = 3.7$ Å, the wave function is quite flat (around 0.025 a.u.) and its maximum is near the surface of cube, as shown in Fig. 6.8a, b and e. It means that the electrons delocalize in the space, displaying an electron gas-like behavior. However, when we compress the cube to $a = 2.0$ Å (the ionic radius of Na^+ is 1.02 Å, i.e., there is now some core-core overlap), the wave functions are quite different. For a_{1g} orbital, the charge density maximum moves to the cube center, with an obviously strong concentration of the wave function. This orbital can be seen as an eight-center two-electron ($8c$-$2e$) bonds, because two electrons occupy this orbital and are shared by eight nuclei. For t_{1u} orbitals, the Na_8 cluster is more like a single atom with p-wave than that with $a = 3.7$ Å. The charge density is now concentrated far outside the Na_8 cube, and can be thought to be suitable for forming bonds with outside entities, while the $8c$-$2e$ bond stabilizes the cluster itself.

Now we can expand this simplified model into a true system, Na_2He, formed at high pressure (>113 GPa) [16]. The structure of Na_2He belongs to the very dense Heusler alloy structure ($AlCu_2Mn$-type, related to Fe_3Al-type): He atoms form a cubic close packing, in which all tetrahedral voids are filled by Na atoms, and $2e$ fill all octahedral voids. Every He atom (and every $2e$) is coordinated by eight Na atoms. Note that $2e$ form a cubic close packing of their own—just like in $hP4$-Na, where they form a nearly perfect hexagonal close packing. In this geometry, localized electrons are inside the Na_8 cluster and occupy the a_{1g} orbitals. Every Na has sp^3-hybridization and participates in four $8c$-$2e$ bonds. The hybridization of Na and the existence of $8c$-$2c$ bonds is proven by solid-state adaptive natural density
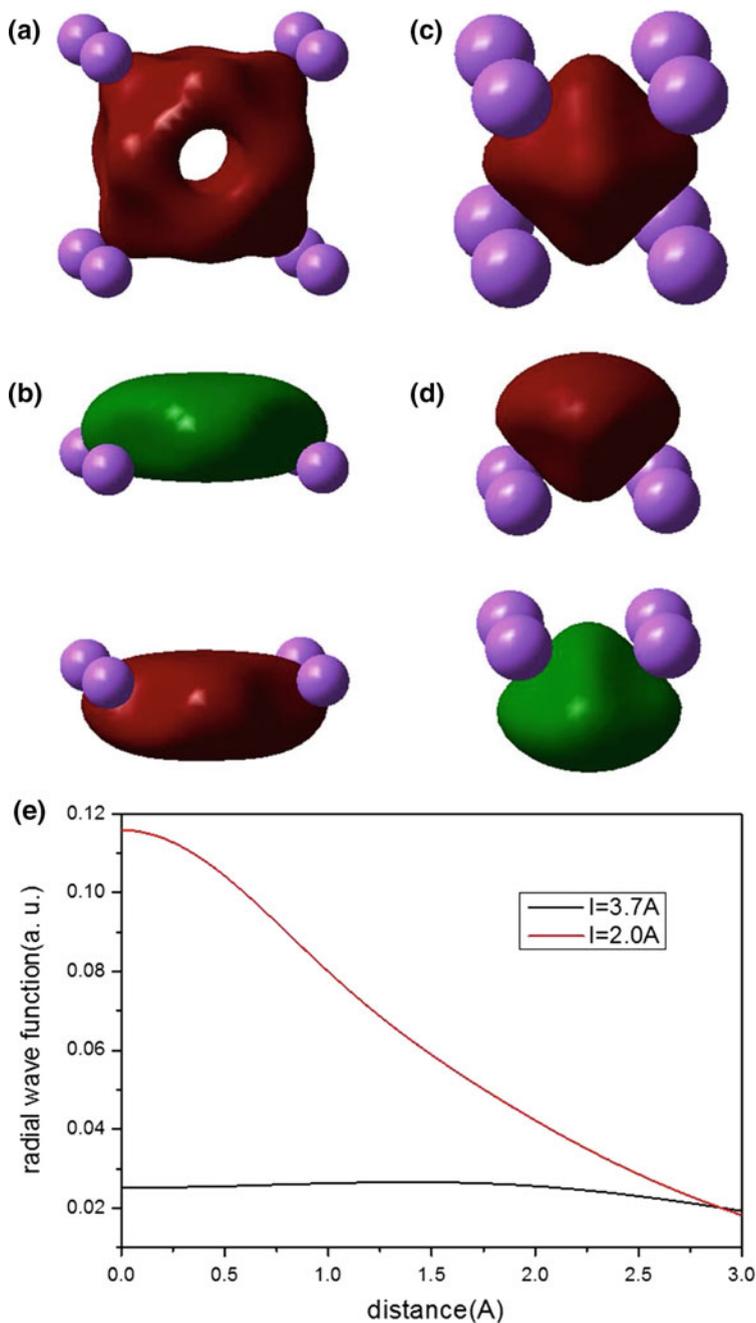
Fig. 6.8 Orbital wave functions of Na_8 cluster. ℓ is the edge length of Na_8 cube. **a** the a_{1g} orbital with cube side $a = 3.7$ Å. **b** the t_{1u} orbital with $a = 3.7$ Å. **c** the a_{1g} orbital with $a = 2.0$ Å. **d** the t_{1u} orbital with $a = 2.0$ Å. **e** the radial wave function as a function of the distance from cube center for $a = 3.7$ Å and 2.0 Å
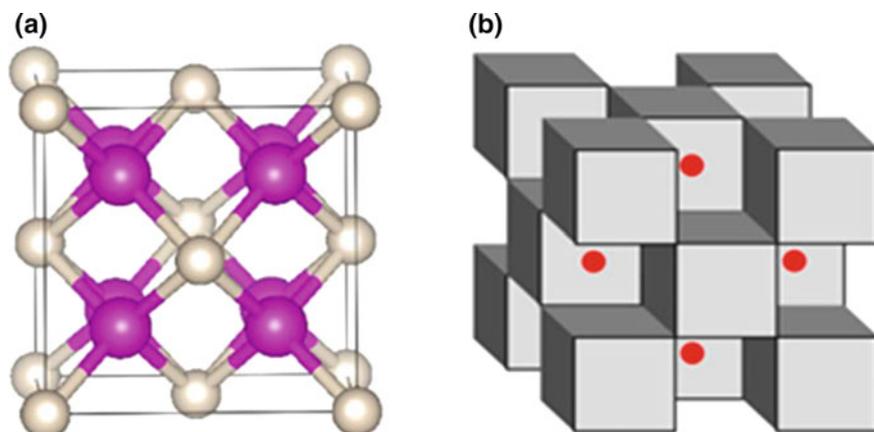
Fig. 6.9 Crystal structure of Na_2He at $300\,GPa$. **a** ball-and-stick representation (*pink* and *gray* atoms represent Na and He, respectively) and **b** polyhedral representation, where half of Na_8 cubes are centered by He atoms (shown by polyhedra), and half by $2e$ (shown by *red spheres*) [16] (color figure online)
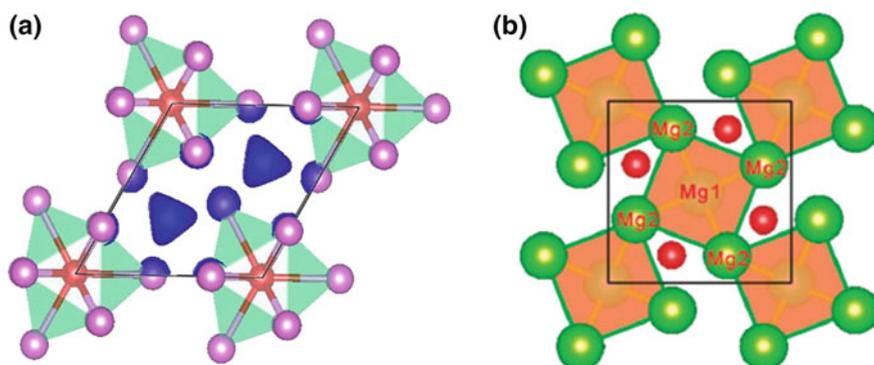
Fig. 6.10 Crystal structure of **a** Li_6O [32] with *blue* isosurface of ELF $= 0.9$ and **b** Mg_3O_2 [14]. In Mg_3O_2, there is a strong charge density localization between four Mg2 and two Mg1 atoms (color figure online)

partitioning (SSAdNDP) [28, 29] and periodic natural bond orbital (NBO) methods [30, 31] (Figs. 6.9 and 6.10).

The Novel Chemistry of Electrides

The interstitially localized electrons greatly change the electronic state of the elements and compounds, leading to altered chemical properties. It is known that Dye's zero-pressure electride has very low work function and is a strong reducing agent [5]. Similarly, high-pressure electrides possess unusual chemical properties.

The first chemical phenomenon originating from the electride state is the formation of clusters of electropositive metals. Normal stoichiometries are governed by valence ratios, and once compounds have more metal atoms than prescribed by valence ratio, there must be bonds among the metal atoms: this happens in the suboxides of caesium and rubidium: Rb_9O_2, Rb_6O, $Cs_{11}O_3$, Cs_4O, Cs_7O, $Cs_{11}O_3Rb$, $Cs_{11}O_3Rb_2$, and $Cs_{11}O_3Rb_3$ [33]. It is remarkable that in these zero-pressure metal-rich compounds, big electropositive elements form a metallic framework and electronegative elements fill the interstices. Many of these are electrides.

At high pressure, smaller electropositive metals, such as Li, Na, and Mg, form metal-rich compounds, which are unknown at zero pressure. The representative cases are Li_6O [32] and Mg_3O_2 [14]. In these compounds, there is hollow space inside the clusters made of metal atoms, and these interstices are occupied by localized electrons. These metal clusters are stabilized by two factors: (1) localized electrons carry a negative charge and behave as anions, binding together the surrounding metal cations. In this way, the whole system is built by Madelung field with two different kinds of anions. (2) if we consider the localized electrons as multicenter bonds, the cluster is combined and stabilized by this electron concentration.

Another chemical peculiarity of high-pressure electrides is their likely extreme reducing ability. In fact, interstitial localization of the electrons is a compromise solution: these electrons have nowhere else to go. Once an electride is combined with other elements that have unoccupied low-energy orbitals, the electrons will move into these orbitals. Like for zero-pressure electrides, we expect very low work functions. A special feature of high-pressure electrides is their extreme chemical activity: at high pressure, one of the most inert elements, He, can receive $\sim$0.1 electrons from Na [16]. We have found [22] that under pressure, Na has the lowest electronegativity (closely followed by Mg) among all elements—even lower than Cs, and this is directly related to its propensity to form the electride state.

Actually, electrides with metal clusters can be considered as normal compounds reduced by elemental electrides. For example, MgO is a saturated system with normal stoichiometry defined by the valence ratio. At high pressure, the electride Mg can reduce it: reaction $Mg + 2MgO = Mg_2O_3$ occurs at pressures above 500 GPa. In this way, the formation of metal clusters is a manifestation of the extreme reducing ability of electride.

In summary, we systematically surveyed high-pressure electrides, using theory, model systems, and real compounds, and considering the nature of their stability, as well as their physical and chemical properties. Electrides, with strongly localized interstitial electron density maxima, which can be considered as pseudoatoms or multicenter bonds, possess extreme chemical activity, reducing ability and can be used as a path to a new branch of high-pressure chemistry.

References

1. J.L. Dye, Science **301**(5633), 607 (2003). DOI 10.1126/science.1088103
2. J.S. Landers, J.L. Dye, A. Stacy, M.J. Sienko, J. Phys. Chem. **85**(9), 1096 (1981). DOI 10.1021/j150609a004
3. S. Matsuishi, Y. Toda, M. Miyakawa, K. Hayashi, T. Kamiya, M. Hirano, I. Tanaka, H. Hosono, Science **301**(5633), 626 (2003). DOI 10.1126/science.1083842
4. J.L. Dye, Science **247**(4943), 663 (1990). DOI 10.1126/science.247.4943.663
5. J.L. Dye, Acc. Chem. Res. **42**(10), 1564 (2009). DOI 10.1021/ar9000857
6. B. Rousseau, N.W. Ashcroft, Phys. Rev. Lett. **101**, 046407 (2008). DOI 10.1103/PhysRevLett.101.046407
7. F. Siringo, R. Pucci, G.G.N. Angilella, High Press. Research **15**(4), 255 (1997). doi:10.1080/08957959708244246. Also available as preprint arXiv.org/abs/cond-mat/9512011
8. Y. Ma, M. Eremets, A.R. Oganov, Y. Xie, I. Trojan, S. Medvedev, A.O. Lyakhov, M. Valle, V. Prakapenka, Nature **458**, 182 (2009). DOI 10.1038/nature07786
9. C.L. Guillaume, E. Gregoryanz, O. Degtyareva, M.I. McMahon, M. Hanfland, S. Evans, M. Guthrie, S.V. Sinogeikin, H. Mao, Nat. Phys. **7**, 211 (2011). DOI 10.1038/nphys1864
10. M. Marqués, M.I. McMahon, E. Gregoryanz, M. Hanfland, C.L. Guillaume, C.J. Pickard, G.J. Ackland, R.J. Nelmes, Phys. Rev. Lett. **106**(9), 095502 (2011). DOI 10.1103/PhysRevLett.106.095502
11. C.J. Pickard, R.J. Needs, Phys. Rev. Lett. **102**, 146401 (2009). DOI 10.1103/PhysRevLett.102.146401
12. J. Lv, Y. Wang, L. Zhu, Y. Ma, Phys. Rev. Lett. **106**(1), 015503 (2011). DOI 10.1103/PhysRevLett.106.015503
13. C.J. Pickard, R.J. Needs, Nat. Mater. **9**(8), 624 (2010). DOI 10.1038/nmat2796
14. Q. Zhu, A.R. Oganov, A.O. Lyakhov, Phys. Chem. Chem. Phys. **15**, 7696 (2013). DOI 10.1039/C3CP50678A
15. W. Zhang, A.R. Oganov, A.F. Goncharov, Q. Zhu, S.E. Boulfelfel, A.O. Lyakhov, E. Stavrou, M. Somayazulu, V.B. Prakapenka, Z. Konôpková, Science **342**(6165), 1502 (2013). DOI 10.1126/science.1244989
16. X. Dong, A.R. Oganov, A.F. Goncharov, E. Stavrou, S. Lobanov, G. Saleh, G.R. Qian, Q. Zhu, C. Gatti, X.F. Zhou, V. Prakapenka, Z. Konôpková, H.T. Wang, Stable compound of helium and sodium at high pressure (2013). arXiv:1309.3827
17. H. Niu, A.R. Oganov, X.Q. Chen, D. Li, Sci. Rep. **5**, 18347 (2015). DOI 10.1038/srep18347
18. N.F. Mott, Z. Zinamon, Rep. Progr. Phys. **33**(3), 881 (1970). DOI 10.1088/0034-4885/33/3/302
19. E. Gregoryanz, L.F. Lundegaard, M.I. McMahon, C. Guillaume, R.J. Nelmes, M. Mezouar, Science **320**, 1054 (2008). DOI 10.1126/science.1155715
20. Y. Xie, A.R. Oganov, Y. Ma, Phys. Rev. Lett. **104**, 177005 (2010). DOI 10.1103/PhysRevLett.104.177005
21. Y. Li, Y. Wang, C.J. Pickard, R.J. Needs, Y. Wang, Y. Ma, Phys. Rev. Lett. **114**, 125501 (2015). DOI 10.1103/PhysRevLett.114.125501
22. X. Dong, A.R. Oganov, G. Qian, X.F. Zhou, Q. Zhu, H.T. Wang, How do chemical properties of the atoms change under pressure (2015). arXiv:1503.00230
23. R.F.W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, Oxford, 1994)
24. F.L. Hirshfeld, Theor. Chim. Acta **44**(2), 129 (1977). DOI 10.1007/BF00549096
25. A.D. Becke, K.E. Edgecombe, J. Chem. Phys. **92**(9), 5397 (1990). DOI 10.1063/1.458517
26. M.S. Miao, R. Hoffmann, Acc. Chem. Res. **47**(4), 1311 (2014). DOI 10.1021/ar4002922
27. M. Winzenick, V. Vijayakumar, W.B. Holzapfel, Phys. Rev. B **50**, 12381 (1994). DOI 10.1103/PhysRevB.50.12381
28. T.R. Galeev, B.D. Dunnington, J.R. Schmidt, A.I. Boldyrev, Phys. Chem. Chem. Phys. **15**, 5022 (2013). DOI 10.1039/C3CP50350J

29. D.Y. Zubarev, A.I. Boldyrev, Phys. Chem. Chem. Phys. **10**, 5207 (2008). DOI 10.1039/B804083D
30. B.D. Dunnington, J.R. Schmidt, J. Chem. Theo. Comp. **8**(6), 1902 (2012). DOI 10.1021/ct300002t
31. J.P. Foster, F. Weinhold, J. Am. Chem. Soc. **102**(24), 7211 (1980). DOI 10.1021/ja00544a007
32. X. Dong, Y.L. Li, A.R. Oganov, K. Li, H. Zheng, H. kwang Mao, Novel chemistry of lithium oxides and superconducting low-pressure LiO_4 (2016). arXiv:1603.02880
33. A. Simon, Coord. Chem. Rev. **163**, 253 (1997). DOI 10.1016/S0010-8545(97)00013-1

Chapter 7
Electron Correlation Effects Reflected in Thermodynamic Properties of Light Actinides

C.C. Matthai and N.H. March

Abstract Experimentally available thermodynamic data on the light actinides is used to display the ratio $B\Omega/k_\mathrm{B}T_m$ from Th through to Cm. Dramatic variation through the series by a factor of 4 contrasts strongly with that found for the five heavy rare-earth solids, also involving f-electrons, where the same ratio is approximately 50 to within a few percent.

Heavy rare-earth materials, involving f-electrons, are characterized by strong electron correlations [1–3]. Evidence of such correlations may be already found in some thermodynamic properties of these metals. This idea has been developed for $3d$-transition metals by Friedel and Sayers [4], who studied the role of electron correlations in the stability of phases in these transition metals in a semi-quantitative way using a second-order perturbation formula related to Gutzwiller's approximation. Friedel and Sayers [4] found that the maximum effect on the surface tension and cohesion occurs for a half-filled d-band.

Important background to this study can be found in the work by Bhatia and March [5]. We shall need to give first a fairly substantial review of this work, in order to gain appreciation to the progress on electron correlation effects reflected in thermodynamic properties of light actinides.

In Faber's fine book on liquid metals [6], he records a compact **r**-space formula for the vacancy formation energy E_v^f (best for close-packed crystals, as it neglects atomic relaxation due to creating a vacancy).

Young et al. [7] have expressly given an alternative **r**-space expression of Faber's formula [see Eq. (7.4) below].

C.C. Matthai (✉)
Department of Physics and Astronomy, Cardiff University, Cardiff, UK
e-mail: Clarence.Matthai@astro.cf.ac.uk

N.H. March
EMAT, Department of Physics, University of Antwerp, Antwerp, Belgium

N.H. March
Abdus Salam International Centre for Theoretical Physics, Trieste, Italy

N.H. March
Oxford University, Oxford, UK

Below, we shall concern ourselves, following a quite brief summary of presently existing work restricted to pair potentials, with demonstrating the importance of electron correlation in determining theoretically some of the thermodynamic properties of light actinides.

First, it will be somewhat illuminating to display the dimensionless ratio for the quantity $B\Omega$, B being the bulk modulus and Ω the atomic volume in relation to the thermal energy of melting, $k_B T_m$. Table 7.1 records the experimental results for the ratio $B\Omega / k_B T_m$ (see also Fig. 7.1).

Some useful background can be found in the early theoretical work of Bhatia and March [5]. These authors gave a pair potential theory involving the ratio $E_v^f / k_B T_m$, where E_v^f is the mono-vacancy formation energy. Their main result was that this ratio could be found from the simple equation

Table 7.1 The last column shows the ratio of bulk modulus B times atomic volume Ω and the thermal energy of melting $k_B T_m$ for the actinides. The somewhat dramatic variation of the ratio $B\Omega / k_B T_m$ through the seven elements listed above is to be contrasted strongly with that found for the five heavy rare-earth solids, also involving f-electrons, where the same ratio is approximately 50 to within a few percent

Actinide	Atomic number	Ω (Å^3)	B (GPa)	$B\Omega / k_B T_m$
Th	90	32	61	67
Pa	91	25.5	129	129
U	92	21	126	136
Np	93	19	96	185
Pu	94	19.8	48	97
Am	95	29	30	44
Cm	96	30	37	50

Fig. 7.1 Variation of the atomic volume Ω for the actinide series [8]
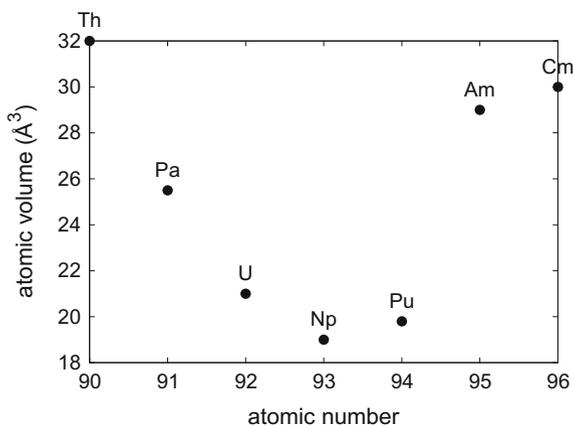
$$\frac{E_v^f}{k_B T_m} + \frac{1}{2} \frac{B\Omega}{k_B T_m} \simeq 20. \tag{7.1}$$

Since materials scientists have a reliable 'rule of thumb' that

$$\frac{E_v^f}{k_B T_m} \simeq 10, \tag{7.2}$$

this yields from Eq. (7.1)

$$\frac{B\Omega}{k_B T_m} \simeq 20, \tag{7.3}$$

whereas the experimental ratio $B\Omega/k_B T_m$ ranged from $\sim$45 to $\sim$190 (cf. Table 7.1). Therefore, light actinides cannot be treated by pair potentials. Of course, it is clear empirically that the f-electron correlations are large in this class of materials. However, the question of what, if any, account of electron correlations can be 'embraced' (if somewhat artificially) via pair potentials may perhaps be worthy of some further study.

An alternative **r**-space expression of Faber's formula has been provided by Young et al. [7] as

$$E_v^f + p\Omega = -\frac{\rho_a}{2} \int g(r)\phi(r)\, d^3\mathbf{r} - \frac{\rho_a}{6} \int r \frac{\partial \phi}{\partial r}\, d^3\mathbf{r}, \tag{7.4}$$

where $g(r)$ is the liquid pair correlation function at the melting temperature T_m. (Experimentally, E_v^f is found to vary unimportantly with temperature.) Because of known 'rule of thumb' correlation $E_v^f/k_B T_m \sim 10$ (accurate for a wide range of metallic crystals, from experimental input), Bhatia and March [5] focussed on the approximate evaluation of Eq. (7.4) in a hot solid near melting.

Of course, for any crystal, the calculation of the pair function $g(r)$ in this case is a major problem necessitating a full treatment of large effects due to anharmonicity. To make progress in further understanding the status of the above empirical correlations, Bhatia and March [5] specifically considered the case of condensed phases of rare gases, where there are known to be no major changes of local coordination on melting—in sharp contrast with, say, tetrahedrally bonded covalent solids like Ge or Si, which melt into a dense metallic phase. They therefore argued that E_v^f can be evaluated in the hot solid by inserting the appropriate liquid structure, described by the pair function $g(r)$ at T_m.

Making this assumption, Bhatia and March [5] were able to approximately evaluate the right-hand side of Eq. (7.4) for E_v^f by inserting the appropriate liquid structure, described by the pair function $g(r)$ at T_m.

Invoking the virial equation for the pressure p, viz.

$$p = \rho_a k_B T - \frac{\rho_a^2}{6} \int r \frac{\partial \phi}{\partial r} g(r)\, d^3\mathbf{r}, \tag{7.5}$$

with pair potential $\phi(r)$, we note that putting $p = 0$ yields the second term of this equation as $-k_B T$. Since E_v^f is known to be much larger than this quantity, it is apparent that the term involving $\partial \phi / \partial r$ can be neglected. Thus, the problem is reduced to an approximate evaluation of the integral $\int g\phi \, d^3\mathbf{r}$ from the above theory of Bhatia and March [5].

The appropriate tool to employ for describing the structure turns out to be the so-called Ornstein–Zernike direct correlation function $c(r)$, which is related to $g(r) = h(r) + 1$ by

$$h(r) = c(r) + \rho_a \int c(|\mathbf{r} - \mathbf{r}'|) h(r') \, d^3\mathbf{r}'. \tag{7.6}$$

At this point, Bhatia and March [5] appealed to the available work of Woodhead-Galloway et al. [9] on $c(r)$ for liquid argon. These authors approximated $c(r)$ by

$$c(r) = c_{\text{hard sphere}}(r) - \frac{1}{k_B T} \phi_{\text{long range}}(r), \tag{7.7}$$

where $\phi_{\text{long range}}(r) = \phi(r)$ outside a hard core of diameter σ. But Percus and Yevick [10] early gave a solution of $c_{\text{hard sphere}}(r)$ and noting that it is identically zero outside $r = \sigma$ in this approximate model, yields from Eq. (7.7) that

$$c(r) = -\frac{\phi(r)}{k_B T}, \qquad r > \sigma. \tag{7.8}$$

Straightforward evaluation then yields

$$\frac{E_v^f}{k_B T_m} = \frac{1}{2} \left[\rho_a \int c(r) \, d^3\mathbf{r} + \rho_a \int h(r)c(r) \, d^3\mathbf{r} - 2 \right]. \tag{7.9}$$

Defining $\tilde{c}(\mathbf{q})$ as the Fourier transform of $c(r)$ through

$$\tilde{c}(q) = \rho_a \int c(r) e^{i\mathbf{q} \cdot \mathbf{r}} d^3\mathbf{r}, \tag{7.10}$$

is shown readily to yield

$$\rho_a \int h(r)c(r) \, d^3\mathbf{r} = -[1 + c(r = 0)]. \tag{7.11}$$

One then finds the desired pair potential result

$$\frac{E_v^f}{k_B T_m} \simeq \frac{1}{2} \left[\tilde{c}(q_L^0) - c(r_L^0) - 3 \right]_{T=T_m}. \tag{7.12}$$

This is Bhatia and March [5] result for E_v^f in units of $k_B T_m$ and the direct correlations function $c(r)$ at the melting temperature T_m.

But $\tilde{c}(q = 0)$ is from the definitions of the direct correlation function immediately related to the long-wave limit $S(0)$ of the liquid structure factor $S(q)$, via the definition of $\tilde{c}(q)$ as

$$\tilde{c}(q) = \frac{S(q) - 1}{S(q)}, \tag{7.13}$$

while $S(0)$ is related to κ_T through the well-established result

$$S(0) = \rho_a k_B T \kappa_T = \frac{k_B T}{B \Omega}, \tag{7.14}$$

where κ_T is the isothermal compressibility. Hence, we have the basic result of Bhatia and March [5] that, at the melting temperature T_m the ratio $E_v^f / k_B T_m$ satisfies

$$\frac{E_v^f}{k_B T_m} + \frac{1}{2} \frac{B \Omega}{k_B T_m} \simeq -\frac{1}{2} \, c(r = 0)|_{T = T_m} - 1. \tag{7.15}$$

For many simple liquids near the melting temperature, this is well approximated by

$$\frac{E_v^f}{k_B T_m} + \frac{1}{2} \frac{B \Omega}{k_B T_m} \simeq 20. \tag{7.16}$$

References

1. N.H. March, G.G.N. Angilella, R. Pucci, Int. J. Mod. Phys. B **28**, 1430019 (2014). DOI 10.1142/S0217979214300199
2. A. Ayuela, N.H. March, Phys. Chem. Liq. **52**(5), 650 (2014). DOI 10.1080/00319104.2014.909918
3. C.C. Matthai, N.H. March, Phys. Chem. Liq. **54**(5), 680 (2016). DOI 10.1080/00319104.2016.1164947
4. J. Friedel, C.M. Sayers, J. Phys. **38**(6), 697 (1977). DOI 10.1051/jphys:01977003806069700
5. A.B. Bhatia, N.H. March, J. Chem. Phys. **80**(5), 2076 (1984). DOI 10.1063/1.446972
6. T.E. Faber, *Introduction to the theory of liquid metals* (Cambridge University Press, Cambridge, 1972)
7. N.H. March, W.H. Young, S. Sampanthar, *The Many-Body Problem in Quantum Mechanics* (Dover, New York, 1995)

8. A.C. Lawson, B. Cort, J.A. Roberts, B.I. Bennett, T. Brun, R.B. Von Dreele, in *Electron correlations and material properties*, ed. by A. Gonis, N. Kioussis, M. Ciftan (Springer, New York, 1999), pp. 75–96
9. J. Woodhead-Galloway, T. Gaskell, N.H. March, J. Phys. C: Solid State Phys. **1**(2), 271 (1968). DOI 10.1088/0022-3719/1/2/301
10. J.K. Percus, G.J. Yevick, Phys. Rev. **110**, 1 (1958). DOI 10.1103/PhysRev.110.1

Chapter 8
Equations of State for Solids Under Strong Compression with Fingerprints for Electronic Anomalies

Wilfried B. Holzapfel

Abstract Changes in the electronic structure of solids under strong compression can be noticed in many cases by close inspection of the Equations of State (EOS), which provide special information on the given thermodynamic system as partial derivatives of the total energy. However, effects of anomalous changes in the electronic structure may be often observed only, when the regular behaviour is well defined. On the basis of an EOS-form interpolating with an Adapted Polynomial (APL) between very strong and moderate compression differences between 'regular' and 'anomalous' compressive behaviour are illustrated with discussions of the EOS for Al, Cu, Ag, and Au as regular in comparison with the EOS of Ce and Pr representing anomalous cases.

Introduction

Thermodynamics of solids provide a link between the microscopic theoretical models describing all the possible quantum-mechanical states of the system and the macroscopic properties, like free energy, entropy, heat capacity, thermal expansion, and the EOS, which is given typically in the form $p = p(V, T, N)$ or inverted as $V = V(p, T, N)$, where p represents the pressure, T the temperature, N the particle number, and V the volume of the system. Most commonly not the total volume of the system is used in general discussions but the reduced or specific volume $v = V/N$, the volume per particle or the volume per mole.

The modeling of the 'system' starts usually with the partition function:

$$Z(V, T, N) = \langle e^{-H/k_B T} \rangle = \sum_n e^{-E_n(V,N)/k_B T}, \qquad (8.1)$$

where H represents the Hamiltonian of the system, $\langle \cdots \rangle$ refers to the expectation value, $E_n(V, N)$ stands for the energy of one specific eigenstate of the system, and

W.B. Holzapfel (✉)
Department-Physik, University Paderborn, Paderborn, Germany
e-mail: holzapfel@physik.upb.de

the sum runs over all the possible eigenstates numerated here in short just by the index n. The partition function reflects all the thermodynamic properties of the system and provides the link to the free energy: $F(V, T, N) = -k_B T \ln Z(V, T, N)$.

This means, that the free energy as a function of the variables V, T, N determines also all the thermodynamic properties of the system. The pressure, for instance, is given as partial derivative: $p(V, T, N) = -\partial F(V, T, N)/\partial V$ and, as mentioned before, one uses mostly the reduced volume and separates the thermal pressure from the pressure at zero temperature in the form: $p(v, T) = p_{zp}(v) + p_{th}(v, T)$.

Usually the thermal pressure $p_{th}(v, T)$ is dominated by contributions from lattice vibrations, which are treated mostly in the quasi-harmonic approximation with various modifications of the Debye-Model [1–3] and further contributions from anharmonicity [3–6], lattice defects [7], electronic excitations [8], and, in some cases, additional contributions from magnetic spin interactions have to be taken into account.

Although these different contributions to the thermal pressure represent a wide field of current interest, I will restrict the present discussion primarily to the difference between 'regular' and 'anomalous' behaviour noticed in the cold pressure related to the ground state of the solid without any further excitations.

EOS for Regular Solids

Let us consider an 'ideal' (metallic) solid with an almost spherical Fermi surface, not touching the Brillouin zone anywhere. In this case, we expect a smooth variation of the volume with pressure or the pressure with respect to volume over very wide ranges in compression with no anomaly at any pressure.

Most of the EOS forms proposed for wide ranges of pressure are based on special assumptions leading to divergences with respect to well-established asymptotic high pressure behaviour of a free electron gas presented by Fermi (1927) and Thomas (1927) and discussed in many papers also more recently as for instance in [9].

As an example for the most commonly used EOS forms with wrong constraints at strong compression I consider here at first the approach of Birch [10], who used the classical finite strain theory to derive a series expansion in Eulerian strain in the form

$$p_{BEL}(x) = \frac{3}{2} K_0 x^{-7} (1 - x^2) \left(1 + \sum_{k=2}^{L} (x^{-2} - 1)^{k-1} \right), \tag{8.2}$$

with $x = (V/V_0)^{1/3}$, where V represents the volume (or reduced specific volume) and V_0 stands for its value at zero (or ambient) pressure at the given temperature with the (isothermal) bulk modulus K_0. In this Birch equation with the label BEL, L denotes the order of this form and mostly this isotherm BEL is used just in second order with the notation BE2:

$$p_{BE2}(x) = \frac{3}{2} K_0 x^{-7} (1 - x^2) \left(1 + c_{BE2}(x^{-2} - 1)\right), \tag{8.3}$$

where $c_{BE2} = \frac{3}{4}(K_0' - 4)$ provides the relation to the isothermal pressure derivative K_0' of the bulk modulus at ambient pressure and at the given temperature for this isotherm. The use of Eulerian strain results in the fact that typical values for $3 < K_0' < 8$ lead to rather small values of this parameter $-1 < c_{BE2} < 3$. However, the exponent -7 of the leading term in x and the negative exponents -2 of x in the series expansion result in a strong divergences with respect to the well-known behaviour of the Fermi gas at very strong compression. As we will see later, this constraint effects also the evaluation of accurate values for K_0 and K_0' from experimental EOS data, when we considers K_0 and K_0' not only as fitting parameters but as well-defined thermodynamic entities. For instance, fitted values of K_0 and K_0' depend on the range of the data, especially when the range in pressure exceeds the value of K_0.

On the other hand, Stacey et al. [11] started from a microscopic model using a Rydberg potential [12] for the nearest-neighbour interaction in a close-packed solid not noticing that the Rydberg potential was designed for molecular vibration around the minimum of the potential with unrealistic constraint at very short distances given by a finite energy at zero distance. The resulting effective Rydberg form has been promoted by Vinet et al. [13, 14] as 'universal equation of state' without any reference to theses earlier studies. Since this form corresponds to a second order approximation I refer to it as ER2:

$$p_{ER2}(x) = 3 K_0 \frac{1 - x}{x^2} \exp\left(c_{ER2}(1 - x)\right), \tag{8.4}$$

with $c_{ER2} = \frac{3}{2}(K_0' - 1)$. Due to the finite energy of the Rydberg potential the calculated pressures of this form diverge with respect to any realistic behaviour at very strong compression, and the evaluation of K_0 and K_0' from experimental EOS data with this form may give again fitted parameters, which differ significantly from the correct thermodynamic values due to the constraints build into this EOS form.

The unreasonable discussions of the ER2 form by Vinet et al. in [13] and in many of their later publications made me so angry that I had to correct their mistakes by a better form.

At first I corrected the leading exponent of x from -2 to -5 fitting to the limiting value constraint by the Fermi gas: $p_{FG}(x) = a_{FG}(Z/V_0)^{5/3} x^{-5} = p_{FG0}/x^5$, where $a_{FG} = 0.02337$ GPa $\cdot$ nm^5 is a universal constant and Z refers to the total electron number in the (atomic) volume V_0 at zero pressure. At first I used a series expansion in the exponential term [9]. Since this series expansion could not be integrated easily in closed form to get the total energy, I proposed the following Adapted Polynomial expansion with the 'order' L:

$$p_{APL}(x) = 3 K_0 \frac{1 - x}{x^5} e^{c_0(1-x)} \left(1 + x \sum_{n=2}^{L} c_{APn} (1 - x)^{n-1}\right), \tag{8.5}$$

where the exponential term with $c_0 = -\ln(3K_0/p_{FG0})$ constrains the limiting behaviour to an asymptotic value of the Fermi gas. The parameter $c_{AP2} = \frac{3}{2}(K_0' - 3) - c_0$ takes care of the value for K_0', and higher order terms could be used to represent the higher order derivatives of the bulk modulus at ambient conditions.

In second order we get the form AP2:

$$p_{\mathrm{AP2}}(x) = 3K_0 \frac{1-x}{x^5} e^{c_0(1-x)} \left(1 + c_{AP2}x(1-x)\right), \tag{8.6}$$

which provides the correct constraint at strong compression, and can be taken as reference for 'regular' solids, when c_{AP2} is rather small. One may notice that for large values of $K_0' \approx 8$ the contribution from c_0 leads typically still to small values of c_{AP2}. In the next section, I apply this form AP2 to the EOS of Al, Cu, Ag, and Au as regular solids and illustrate the problems with the other EOS forms like BE2 and ER2.

Comparison of Different EOS Forms for Regular Solids

If we want to compare experimental data and different EOS forms over wide ranges in compression, we can see by comparison of the data for aluminium (Al) in Fig. 8.1
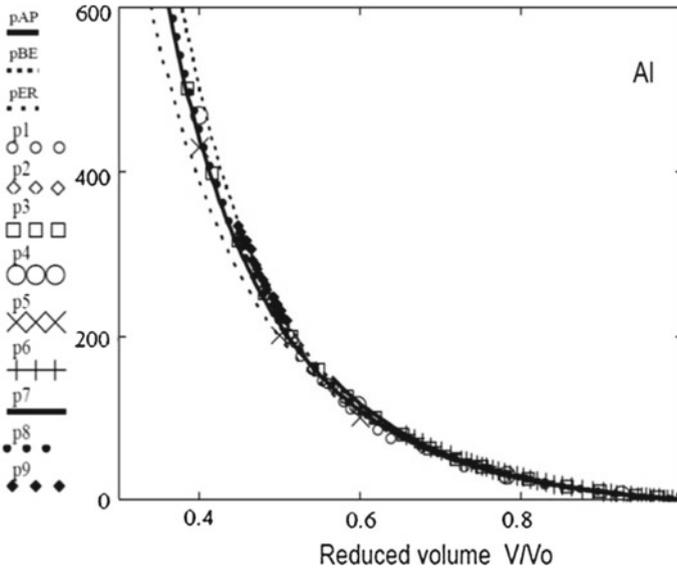
Fig. 8.1 Comparison of different EOS forms (p_{AP2}, p_{BE2}, p_{ER2}) with experimental and theoretical data (p1, p2, p3, p4, p5, p6, p7, p8, p9 referring to [15–23] respectively) for Aluminium (Al) in the range up to 600 GPa plotted versus reduced volume
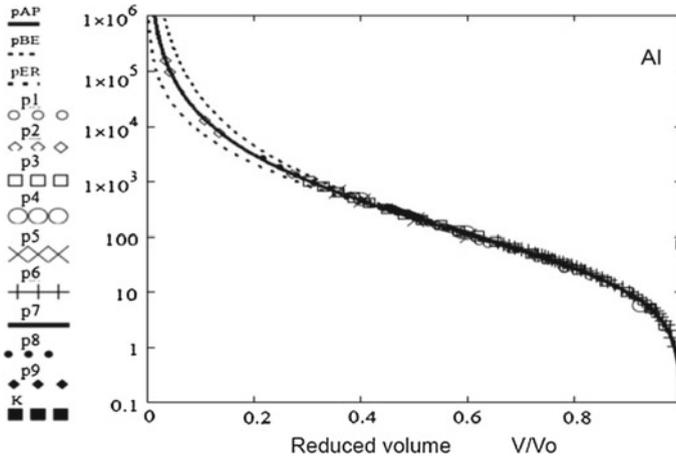
Fig. 8.2 Comparison of different EOS forms (p_{AP2}, p_{BE2}, p_{ER2}) with the same experimental and theoretical data (p1, p2, p3, p4, p5, p6) for Aluminium (Al) in the range up to 10^6 GPa on logarithmic scale plotted versus reduced volume

with 8.2 that it is favourable to use a logarithmic representation of the pressure to show the regions of low and high pressure with similar resolution.

Both figures illustrate that the p_{AP2} fits best all the data and it should be noted that I use here the ultrasonic values for K_0 and K'_0. On the other hand, both p_{BE2} and p_{ER2} deviate from the data at higher pressure. The black mark on the right-hand side of Fig. 8.2 represents the value of the bulk modulus at ambient pressure and shows that this deviation becomes very significant when the pressure exceeds this value of K. If we divide all the pressures-data in this figure by the pressure of the Fermi gas, the ratio $R(x) = p(x)/p_{FG}(x)$ enhances the resolution as seen in Fig. 8.3, and we can compare now this variation with what we expect from the Thomas–Fermi theory at very strong compression as shown by the thin straight line.

Here we see that both the theoretical data from SEAME [16] represented by R2 and extending to extremely high pressure as well as the ratio RAP for $p_{AP2}(x)$ both approach very smoothly the Thomas–Fermi behaviour marked RTF. Now it becomes even more clearly evident that both $p_{BE2}(x)$ and $p_{ER2}(x)$ are constraint in such a way that they diverge with respect to any reasonable variation and cannot represent the EOS behaviour of 'regular' solids under strong compression. Since I consider also copper, silver, and gold as regular solids without any special electronic anomalies under strong compression and we have good EOS data for theses elements, I illustrate with these additional examples what I would like to call a 'regular solid'.

What can be noticed here for Copper (Fig. 8.4) is the fact that RAP falls nicely in between the data R6 and R7 [27, 28] but some of the older shock-wave-reduced-data R2 [25] deviate significantly due to the well-known problems in estimating the thermal pressure in the shocked state [30–33]. On the other hand RAP approaches RTF more steeply than in the case of Al.
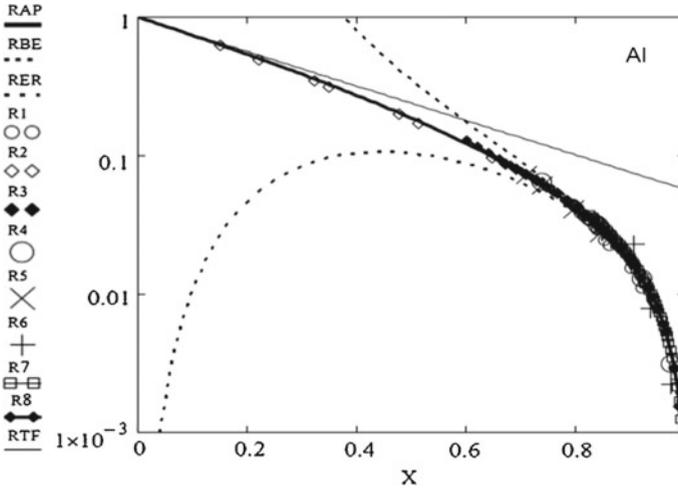
Fig. 8.3 Comparison of different EOS forms (p_{AP}, p_{BE}, p_{ER}) represented by *solid*, *dotted*, and *short dashed lines* with experimental and theoretical data (p1, p2, p3, p4, p5, p6) for Aluminium (Al) using pressure ratios $R(x)$ with respect to pressure of the corresponding Fermi gas plotted on logarithmic scale versus x
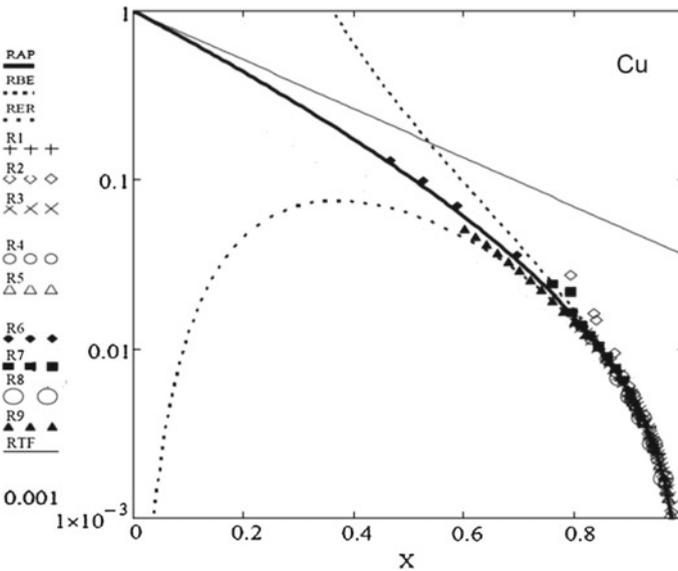
Fig. 8.4 Comparison of different EOS forms (p_{AP}, p_{BE}, p_{ER}) represented by *solid*, *dotted* and *short dashed lines* with experimental and theoretical data (p1, p2, p3, p4, p5, p6, p8, p9, [17, 20, 22, 24–29]) for Copper (Cu) using pressure ratios $R(x)$ with respect to pressure of the corresponding Fermi gas plotted on logarithmic scale versus x
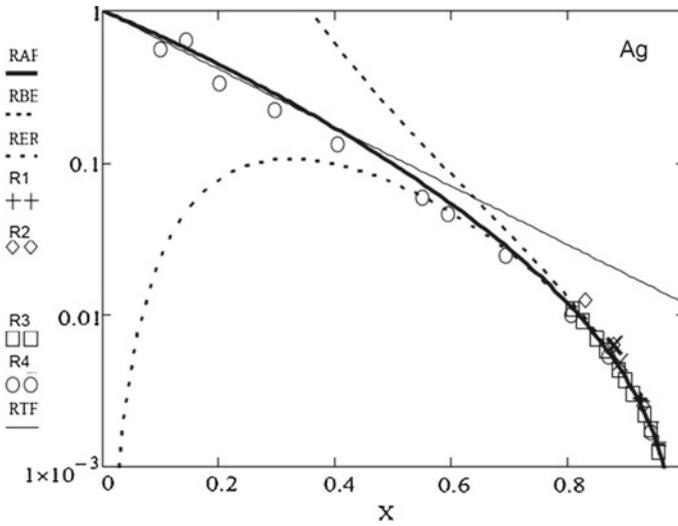
Fig. 8.5 Comparison of different EOS forms (p_{AP}, p_{BE}, p_{ER}) represented by *solid*, *dotted* and *short dashed lines* with experimental and theoretical (p1, p2, p3, p4 [16, 24–26]) data for Silver (Ag) using pressure ratios $R(x)$ with respect to pressure of the corresponding Fermi gas plotted on logarithmic scale versus x

In the case of Silver (Fig. 8.5) the theoretical data R4 from SESAME [16] seem to show a large uncertainty or misprint at very strong compression by the strong deviation of the secondlast data-circle. RAP2 overshoots slightly the theoretical data and the Thomas–Fermi line in this region. With the fit of a third-order form $p_{AP3}(x)$ I could take into account this deviation. However this difference reflects for me just the present uncertainty of any extrapolation or calculation for this extreme range in compression.

For Gold (Fig. 8.6) there are many data for the pressure range below 150 GPa and all these data cluster around the RAP curve derived from $p_{AP2}(x)$ with input data Z, V_0, K_0, and K_0'. only from ambient pressure. The theoretical data R8 for the range up to 600 GPa [22] correspond to a softer behaviour and suggest the interpolating curve RAP which fits here very smoothly also to the Thomas–Fermi line at very strong compression.

I could add similar figures for many other 'regular' metals like Ni, Pd, Pt, Mo, W, and others, but I believe the four examples Al, Cu, Ag, and Au shown here may be sufficient to support my view that the form $p_{AP2}(x)$ reproduces the behaviour of regular metals under strong compression very well and may serve as reference for studies on the influence of special electronic anomalies on EOS.
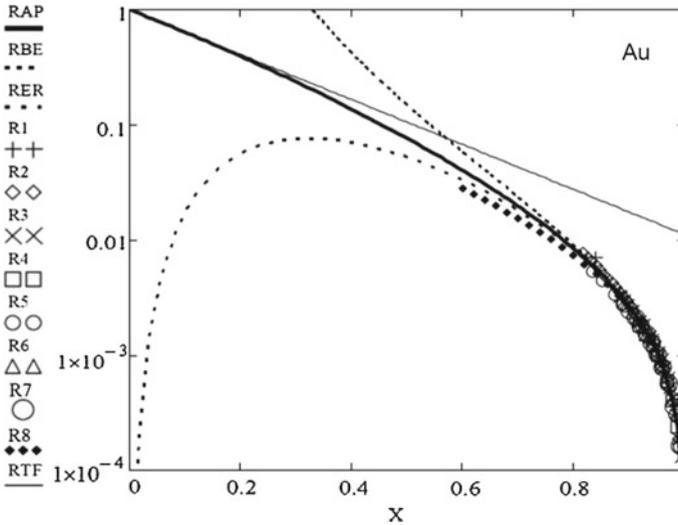
Fig. 8.6 Comparison of different EOS forms (p_{AP}, p_{BE}, p_{ER}) represented by *solid*, *dotted* and *short dashed lines* with experimental and theoretical data (p1, p2, p3, p4, p5, p6, p7, p8 [20–22, 25, 26, 30, 31, 34]) for Gold (Au) using pressure ratios $R(x)$ with respect to pressure of the corresponding Fermi gas plotted on logarithmic scale versus x

Effects of Electronic Anomalies on the EOS of Cerium

Perhaps the best example for the influence of electronic anomalies on the EOS of solids is presented by Cerium (Ce) already at moderate pressures, where not only the EOS shows an unusual behaviour (at elevated temperatures) but more strikingly, one observes a volume collapse transition similar to the transition between liquid and gas without any change in symmetry. Both the low- and the high pressure face have fcc structure [35]. The critical point of this phase transition and EOS data have been studies by high pressure X-ray diffraction [36] and the results are reproduced as circles in Fig. 8.7.

Many different models have been discussed in the literature [35] based on the idea that the f-electrons somehow delocalize. This does not mean that the f-electrons flip out from some inner localized orbitals completely into the conduction band but they may flip or mix between two different 'configuration'. Such a model of two configurations now for the total energy of the solid with the same overall symmetry would allow for configuration crossing with configuration interaction [37] and can be used to model the EOS in the transition region. Like the formation of bonding and anti-bonding orbitals in molecules, configuration interaction may effect the crossing of the energy levels $E_1(a)$ and $E_2(a)$ of the two configurations indicated by the crossing of the thin dashed and dash-dotted curves in Fig. 8.8, where a represents the lattice parameter of the cubic unit cell. If we would take into account that the total energy depends also on temperature, we would need a three-dimensional model
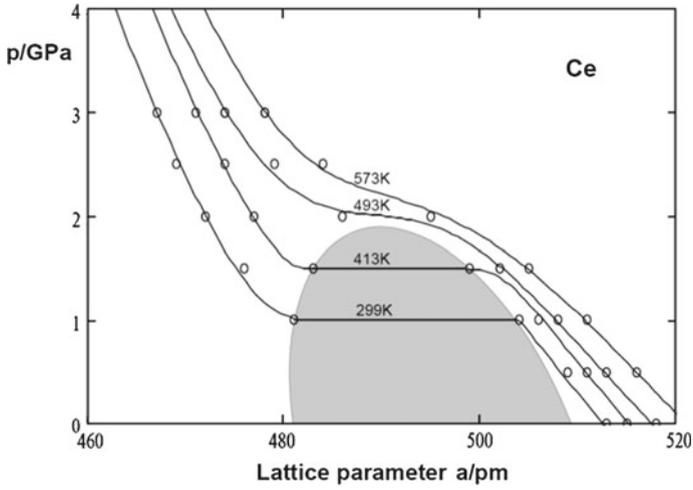
Fig. 8.7 Experimental data (*open circles* [36]) and fitted theoretical curves of the present model for four isotherms of Cerium (Ce) plotted versus the lattice parameter of the fcc unit cell. First order phase transition marked by the *grey zone*
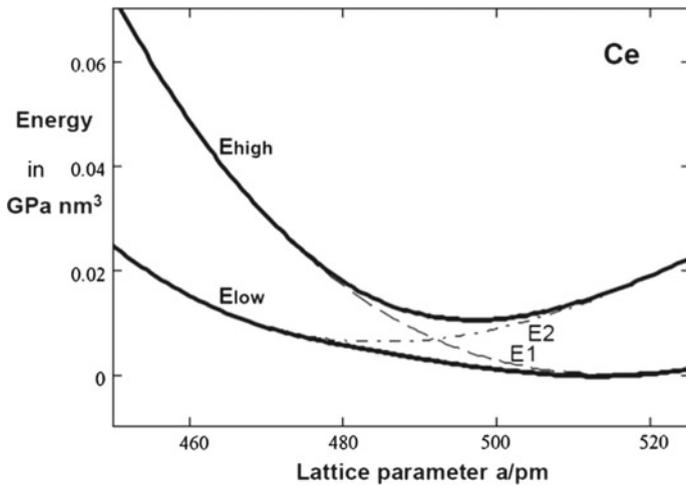
Fig. 8.8 Crossing of energy levels (*dashed* and *dash-dotted curves*) and level splitting by config-uration interaction of the present model (*solid curves*) for Cerium (Ce) at 299 K plotted versus the length of the fcc unit cell

with crossing of the energy planes, but I will restrict my discussion just to individual isotherms.

With configuration interaction the levels $E_1(a)$ and $E_2(a)$ will mix and repulse each other as shown schematically by the curve labelled E_{high} and E_{low} in Fig. 8.8. I use for the configuration interaction the following form:

$$E_m(x) = \frac{E_r E_w}{\sqrt{(E_1(x) - E_2(x))^2 + E_w^2}}. \tag{8.7}$$

This gives for the two interacting levels

$$E_{\text{high}} = \frac{1}{2} \left(E_1(x) + E_2(x) + \sqrt{(E_1(x) - E_2(x))^2 + E_m^2(x)} \right), \tag{8.8a}$$

$$E_{\text{low}} = \frac{1}{2} \left(E_1(x) + E_2(x) - \sqrt{(E_1(x) - E_2(x))^2 + E_m^2(x)} \right). \tag{8.8b}$$

If I take $E_2(x) > E_1(x)$ in the low pressure region as shown in Fig. 8.8, I get for the corresponding EOS forms

$$p_{\text{high}} = \frac{1}{2} \left(p_1(x) + p_2(x) + (p_1(x) - p_2(x)) \frac{E_1(x) - E_2(x)}{\sqrt{(E_1(x) - E_2(x))^2 + E_m^2(x)}} \right), \tag{8.9a}$$

$$p_{\text{low}} = \frac{1}{2} \left(p_1(x) + p_2(x) - (p_1(x) - p_2(x)) \frac{E_1(x) - E_2(x)}{\sqrt{(E_1(x) - E_2(x))^2 + E_m^2(x)}} \right). \tag{8.9b}$$

If I consider that the unperturbed pressures, $p_1(x)$ and $p_2(x)$, and energies, $E_1(x)$ and $E_2(x)$, correspond to 'regular' behaviour represented by $p_{AP2}(x)$ and corresponding EAP2(x) forms, I must select first Z, V_0, K_0, and K_0' for the two unperturbed isotherms and an appropriate difference E_r between the two energy levels at very large x to shift the level crossing to the appropriate value of x. With the parameters E_r and E_w, I can adjust the strength and the width of the configuration interaction to get the 'theoretical' curves in Fig. 8.7.

After our low pressure study of Cerium in its region of the volume collapse more recent X-ray diffraction work [38] provided additional structural and compression data on Ce up to 208 GPa. I have fitted their data for the bct phase in the range from 15 to 208 GPa by the 'regular' AP2 form and used this form also for the unperturbed EOS at lower pressures assuming that the second-order fcc—bct transition has no strong effect on the compressive behaviour. In this way, I obtained the room temperature isotherm for Ce in the extended pressure range shown in Fig. 8.9.

The curve p_{F1} represents the theoretical curve in Fig. 8.7 for 299 K with the transition between the unperturbed pressures $p_1(x)$ and $p_2(x)$ by configuration interaction. p_{F1} approaches, therefore, continuously p_{F2} already around 10 GPa. We can now
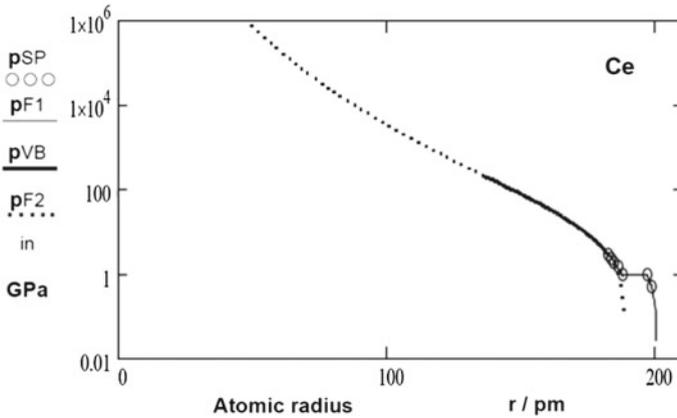
Fig. 8.9 The 299 K isotherms of Cerium (Ce) with the low pressure data for the fcc phase from Fig. 8.8 shown as *circles*, p_{SP}, and *thin curve*, p_{F1}, together with the high pressure data for the bct phase [4] up to 208 GPa shown by the *solid curve*, p_{VB}, with the data from the fitted AP2 form represented as *dotted line*
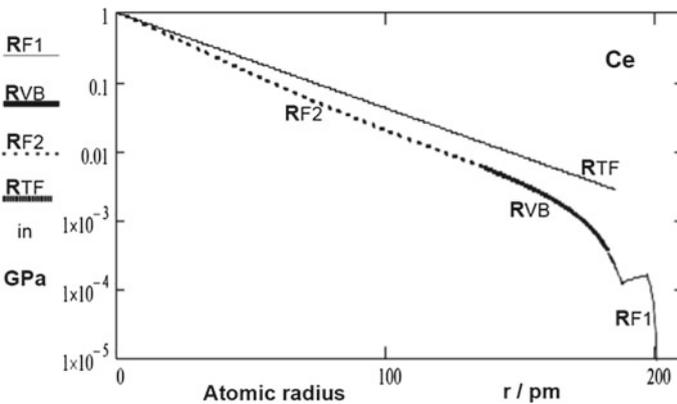
Fig. 8.10   Room temperature isotherm of Cerium (Ce) from Fig. 8.9 scaled by the Fermi gas pressure as shown previously with the ratio $R(x)$ of regular metals covering here the full range of compression from 0 to 200 pm with the volume collapse included

ask the question: Does Ce behave like a regular solid in the bct phase? On the bases of Fig. 8.9, it is difficult to make a decision.

However, the 'regular' behaviour for Al, Cu, Ag, and Au illustrated in the Figs. 8.3, 8.4, 8.5, and 8.6 gives us a background for similar considerations for Ce when we use the corresponding scaling by the Fermi gas pressure shown in Fig. 8.10.

First of all, we see that the anomaly at low pressure is very well recognized also in the expanded view. On the other hand, we notice that RF2 approaches rather smoothly the Thomas–Fermi behaviour at strong compression. However, this approach is not as rapid as for Al, Ag, and Au. Does this reflect an anomaly?

If we look at Fig. 8.4 for Cu, we see some similarity. This would mean that also Ce behaves very regular after the (low pressure) volume collapse. Furthermore, we may take into account that the high pressure X-ray diffraction data [38] may be compatible also with a more rapid approach towards the Thomas–Fermi behaviour due to the uncertainties still involved in the pressure standards used in this type of experiments [3, 39, 40].

Effects of Electronic Anomalies on the EOS of Praseodymium

Praseodymium may be considered as prototype for a lanthanide element with volume collapse phase transition where a significant volume collapse occurs together with a structural phase transition in contrast to cerium, where this transition is isostructural. We may ask here especially whether the phase transition leads directly to regular compressive behaviour or still to a state with inter-configuration interaction. The plot of the pressure ratio for the 299 K-isotherm (Fig. 8.11) shows at first the volume collapse at the phase transition between the curve RF1 for the low pressure fcc- and distorted-fcc phase and the high pressure $\alpha - U$-phase. The best fit of an AP2-EOS to the extended data for the $\alpha - U$-phase [42] is illustrated by the curve RF2, which extrapolates at higher pressures beyond the strait line for the Thomas–Fermi behaviour. Since the Thomas–Fermi line is considered as upper boundary for any
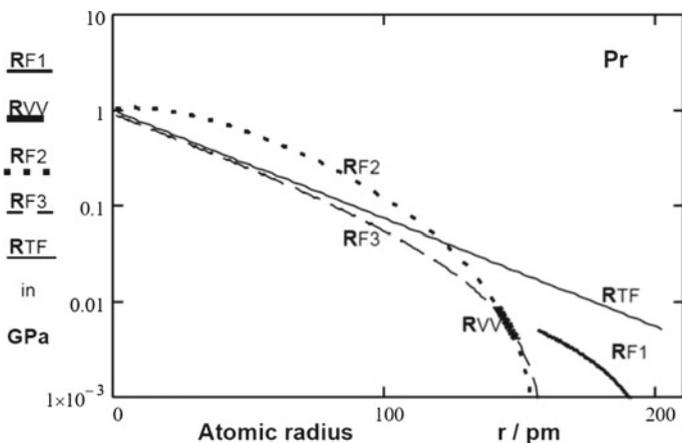
Fig. 8.11 Compression data for Praseodymium (Pr) at room temperature [4, 41–43] scaled in $R(x)$ by pressure of the corresponding the Fermi gas (*thick solid curves* RF1 [41], RVV [42]) compared with the limiting curve from the Thomas–Fermi model (*thin line* RTF) and the pressure expected by extrapolation of the best fitting $p_{AP2}(x)$ form (*dotted curve* RF2) and by $p_{AP3}(x)$ (*slashed curve* RF3) expected for most regular behaviour

EOS, we can only conclude that this extrapolation gives to high pressure values due to the fact that Pr behaves still slightly anomalous in this range and should be modelled by a third-order form p_{AP3} with the slope at ultimate compression fitted to the Thomas–Fermi behaviour shown by the dashed line RF3 or by a model including explicitly configuration interaction.

Since the differences in the EOS for different phases with 'regular' EOS should become rather small at elevated pressure due to the fact that the structural energy differences do not increase as much as the total energy under strong compression, we can ask ourselves, whether it is still possible to observe inter-configuration-crossing under these conditions.

Inter-configuration-crossing at Strong Compression

To test this question I have modelled a situation (Fig. 8.12), where this crossing occurs at much higher compression than in the case of Cerium. Given this model, I can calculate the corresponding EOS for the ground state level E_{low} of Fig. 8.12 as shown by the thick curve p_{M3} in Fig. 8.13. This curve looks so smooth that an anomaly is difficult to notice. Different calculations indicated in Fig. 8.13 by the thin lines p_{M2} and p_{M1} use the same strength E_r in the coupling but much smaller parameters E_w for the width. In fact the sharp change in p_{M1} corresponds to almost zero width, which is unrealistic but illustrates the position of the level crossing.

Since the anomaly in p_{M3} is not easily noticed in Fig. 8.13, I plot the same data in Fig. 8.14 as pressure ratios (RM1, RM2, RM3) with respect to the corresponding
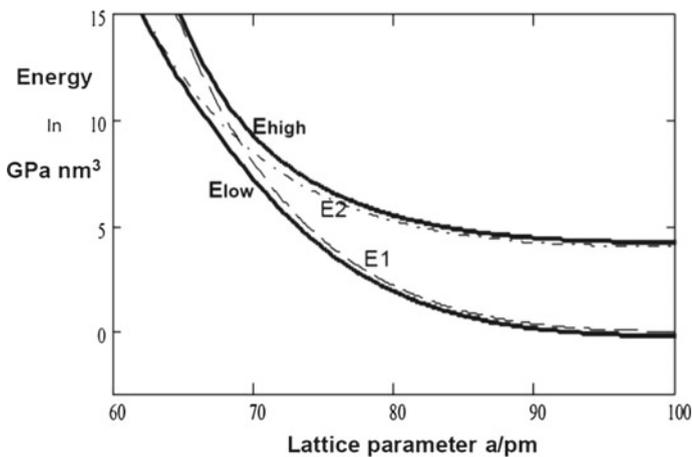
Fig. 8.12 Model calculation for level crossing at strong compression, here at 69 pm with respect to the minimum (zero pressure) at 100 pm and with moderate strength of the inter-configuration coupling
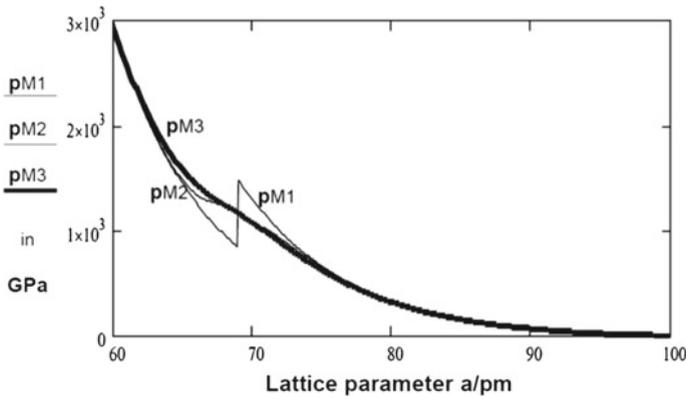
Fig. 8.13 Calculated EOS (*solid curve p_{M3}*) for the ground state energy E_{low} of Fig. 8.12 compared with the results for narrower coupling widths E_w shown by the *thin lines sp_{M2}* and p_{M1}
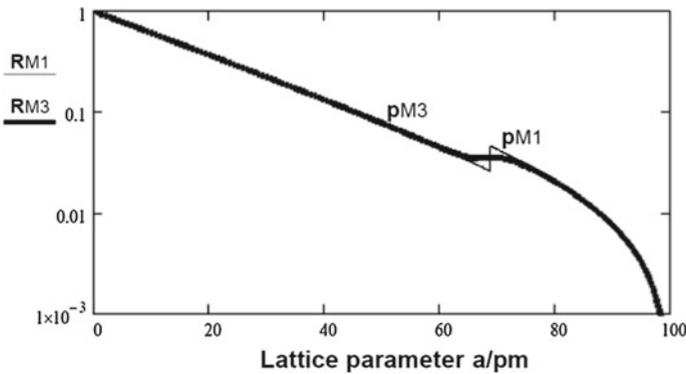
Fig. 8.14 Calculation of pressure ratios RM1 and RM3 for the EOS p_{M1} and p_{M3} of Fig. 8.13 for the extended pressure range to very strong compression

Fermi gas. RM1 from p_{M1} shows the sharp crossover. RM3 from pM3 exhibits a very clear anomaly and RM2 is hidden under RM3.

It should be noticed that the range in compression in Fig. 8.14 extends to $a/pm = 0$ in contrast to the moderate compression in Fig. 8.13. Nevertheless the anomaly is very clearly noticed!

Conclusions

I hope to have shown that an EOS-form like pAP2(x) with the correct constraint for very strong compression avoids the problems of the more commonly used EOS form of Birch [10] or the effective Rydberg form [11, 12] and can be used to distinguish

'regular' from 'anomalous' compressive behaviour of solids under strong compression. The corresponding differences are most clearly noticed, when the pressure ratio $R(x)$ of the actual pressure divided by the pressure of the Fermi gas for the same total electron density is plotted for wide ranges in compression. Although I illustrated anomalous behaviour only for the level crossing typical for lanthanides, also other anomalies for instance from magnetic interaction may by studied in a similar way.

References

1. W.B. Holzapfel, in *High-pressure crystallography*, ed. by A. Katrusiak, P. McMillan (Kluwer Academic Publisher, Dordrecht, 2004), pp. 217–236
2. J.S. Tse, W.B. Holzapfel, J. Appl. Phys. **104**(4), 043525 (2008). DOI 10.1063/1.2969909
3. W.B. Holzapfel, J. Phys.: Cond. Matter **14**(44), 10525 (2002). DOI 10.1088/0953-8984/14/44/327
4. W.B. Holzapfel, High Press. Res. **25**(3), 187 (2005). DOI 10.1080/08957950500235900
5. W.B. Holzapfel, High Press. Res. **26**(4), 313 (2006). DOI 10.1080/08957950601106596
6. W.B. Holzapfel, Z. Naturforsch **63b**, 718 (2008)
7. A.I. Karasevskii, W.B. Holzapfel, Low Temp. Phys. **29**(9), 711 (2003). DOI 10.1063/1.1614174. [Fizika Nizkikh Temp. **29**, 951 (2003)]
8. W.B. Holzapfel, Z. Kristallogr. **216**, 473 (2001). DOI 10.1524/zkri.216.9.473.20346
9. W.B. Holzapfel, in *Molecular systems under high pressure: Proceedings of the II Archimedes workshop on Molecular solids under pressure, Catania, Italy, May 28–31, 1990*, ed. by R. Pucci, G. Piccitto (North-Holland, Amsterdam, 1991), pp. 61–68
10. F. Birch, Phys. Rev. **71**, 809 (1947). DOI 10.1103/PhysRev.71.809
11. F.D. Stacey, B.J. Brennan, R.D. Irvine, Geophysical surveys **4**(3), 189 (1981). DOI 10.1007/BF01449185
12. R. Rydberg, Z. Physik **73**(5), 376 (1932). DOI 10.1007/BF01341146
13. P. Vinet, J. Ferrante, J.R. Smith, J.H. Rose, J. Phys. C: Solid State Phys. **19**(20), L467 (1986). DOI 10.1088/0022-3719/19/20/001
14. W.B. Holzapfel, High. Press. Res. **16**, 81 (1998)
15. R.G. Greene, H. Luo, A.L. Ruoff, Phys. Rev. Lett. **73**, 2075 (1994). DOI 10.1103/PhysRevLett.73.2075
16. D. Young (1997). Private communication
17. X. Wang, U. Grimm, M. Schreiber, Phys. Rev. B **62**(21), 14020 (2000). DOI 10.1103/PhysRevB.62.14020
18. J.C. Boettger, S.B. Trickey, Phys. Rev. B **29**, 6434 (1984). DOI 10.1103/PhysRevB.29.6434
19. A.C. Mitchell, W.J. Nellis, J.A. Moriarty, R.A. Heinle, N.C. Holmes, R.E. Tipton, G.W. Repp, J. Appl. Phys. **69**(5), 2981 (1991). DOI 10.1063/1.348611
20. G.C. Kennedy, N. Keeler, in *AIP handbook*, ed. by D.E. Gray (McGraw-Hill, New York, 1972), p. 4.1721
21. A.D. Chijioke, W.J. Nellis, A. Soldatov, I.F. Silvera, J. Appl. Phys. **98**(11), 114905 (2005). DOI 10.1063/1.2135877
22. A.B. Alchagirov, J.P. Perdew, J.C. Boettger, R.C. Albers, C. Fiolhais, Phys. Rev. B **63**, 224115 (2001). DOI 10.1103/PhysRevB.63.224115
23. Y. Akahama, M. Nishimura, K. Kinoshita, H. Kawamura, Y. Ohishi, Phys. Rev. Lett. **96**, 045505 (2006). DOI 10.1103/PhysRevLett.96.045505
24. J.M. Walsh, M.H. Rice, R.G. McQueen, F.L. Yarger, Phys. Rev. **108**, 196 (1957). DOI 10.1103/PhysRev.108.196
25. L.W. Alt'shuler, K.K. Krupnikov, B.N. Ledenev, W.I. Zhuchikhin, M.I. Brazhnik, Sov. Phys. JETP **34**, 606 (1958)

26. L.V. Alt'shuler, S.E. Brusnikin, J. Appl. Mech. and Tech. Phys. **28**, 192 (1987)
27. H.K. Mao, P.M. Bell, J.W. Shaner, D.J. Steinberg, J. Appl. Phys. **49**(6), 3276 (1978). DOI 10. 1063/1.325277
28. R.C. Albers, A.K. McMahan, J.E. Müller, Phys. Rev. B **31**, 3435 (1985). DOI 10.1103/ PhysRevB.31.3435
29. W.J. Nellis, J.A. Moriarty, A.C. Mitchell, M. Ross, R.G. Dandrea, N.W. Ashcroft, N.C. Holmes, G.R. Gathers, Phys. Rev. Lett. **60**, 1414 (1988). DOI 10.1103/PhysRevLett.60.1414
30. W.B. Holzapfel, in *Physics of extreme state of matter—2010*, ed. by V.E. Fortov, B.S. Karamur-zov, A.I. Temrokov, V.P. Efremov, K.V. Khishchenko, V.G. Sultanov, P.R. Levashov, G.I. Kanel, I.L. Iosilevski, V.V. Milyavskiy, V.B. Mintsev, O.F. Petrov, A.P. Savintsev, G.V. Shpatakovskaya (Russian Academy of Sciences, Chernogolovka, 2010), pp. 6–8
31. W.B. Holzapfel, High Press. Res. **30**(3), 372 (2010). DOI 10.1080/08957959.2010.494845
32. B.K. Godwal, A. Ng, R. Jeanloz, High Press. Res. **10**(5-6), 687 (1992). DOI 10.1080/ 08957959208225320
33. J.C. Jamieson, J.N. Fritz, M.H. Manghnani, in *High pressure research in geophysics*, ed. by S. Akimoto, M.H. Manghnan (Center for Acad. Publ., Tokyo, 1992)
34. S.M. Dorfman, V.B. Prakapenka, Y. Meng, T.S. Duffy, J. Geophys. Res.: Solid Earth **117**(B8), B08210 (2012). DOI 10.1029/2012JB009292
35. D.A. Young, *Phase diagrams of the elements* (University of California, Berkeley, 1991)
36. A. Schiwek, F. Porsch, W.B. Holzapfel, High Press. Res. **22**(2), 407 (2002). DOI 10.1080/ 08957950212799
37. D. Wohlleben, J. Röhler, J. Appl. Phys. **55**(6), 1904 (1984). DOI 10.1063/1.333516
38. Y.K. Vohra, S.L. Beaver, J. Akella, C.A. Ruddle, S.T. Weir, J. Appl. Phys. **85**(4), 2451 (1999). DOI 10.1063/1.369566
39. W.B. Holzapfel, J. Appl. Phys. **93**(3), 1813 (2003). DOI 10.1063/1.1525856
40. W.B. Holzapfel, High Press. Res. **25**(2), 87 (2005). DOI 10.1080/09511920500147501
41. W.A. Grosshans, W.B. Holzapfel, Phys. Rev. B **45**, 5171 (1992). DOI 10.1103/PhysRevB.45. 5171
42. N. Velisavljevic, Y.K. Vohra, High Press. Res. **24**(2), 295 (2004). DOI 10.1080/ 08957950410001717150
43. G.K. Samudrala, Y.K. Vohra, in *Handbook on the physics and chemistry of rare earths*, vol. 43 (Elsevier, Amsterdam, 2013), chap. 257, p. 275

Chapter 9
Band Gaps and Effective Oscillator Models for Solid Hydrogen and H$_2$O Ice at High Pressure

Wai-Leung Yim, Hongliang Shi, Yunfeng Liang, Russell J. Hemley and John S. Tse

Abstract The validity of simple effective oscillator models for extracting band gap energies from frequency-dependent refractive index data of high pressure ice and solid hydrogen was assessed through comparisons of theoretical dielectric response functions obtained from Bethe–Salpeter equation and band gaps calculated from the GW method. For model structures of solid hydrogen, the 'effective' band gap trend with pressure estimated from experimental refractive index correlate well with theoretical predictions. However, this is not the case for ice at high pressure. The success in the former can be attributed to the lack of variation in the optical absorption spectral profile with pressure due to the molecular-like localized electronic excitations. In comparison, the width of the absorption band in ice broadened continuously with compression. Single oscillator models are not expected to be reliable when the spectral profile changes significantly with increasing pressure as we predict for dense phases of solid hydrogen, where the molecules exhibit stronger intermolecular interactions than in the lower pressure structures of solid hydrogen.

W.-L. Yim
Institute of High Performance Computing, Agency for Science, Technology and Research,
1 Fusionopolis Way, #16-16, Connexis 138632, Singapore

H. Shi
Key Laboratory of Micro-Nano Measurement-Manipulation and Physics,
Department of Physics, Beihang University, Beijing 100191, China

Y. Liang · J.S. Tse (✉)
Department of Physics and Engineering Physics, University of Saskatchewan,
Saskatoon, SK S7N 5E2, Canada
e-mail: john.tse@usask.ca

Y. Liang
Research into Artifacts, Center for Engineering, University of Tokyo, Kashiwa Campus,
Kashiwa, Japan

R.J. Hemley
Department of Civil and Environmental Engineering, The George Washington University,
Washington, DC 20052, USA

R.J. Hemley
Lawrence Livermore National Laboratory, Livermore, CA 94551, USA

Introduction

The energy of the gap separating the top of the valence band and bottom of the conduction band of an insulator is an important parameter characterizing the electronic properties of the material [1]. Traditionally, band gap energies can be determined from the absorption edge from the optical spectra. This technique, however, cannot be used for the measurement of large band gap materials under pressure when the absorption edge is higher than that of the anvil (e.g., in diamond anvil cells) as the incident radiation is blocked by the strong ultraviolet absorption (i.e., above 4 eV or lower depending on the state of stress of the diamond anvils). The pressure dependence of the refractive index and its dispersion of a material at visible wavelengths contains information on the energy and intensity changes in electronic excitations at higher energies, including excitations across the band gap. These shifts have been constrained by analyses of visible refractive index data estimated using simple effective oscillator models. One such model is that proposed by Wemple and DiDomenico [2], which was shown to give satisfactory results in more than 100 widely different covalent and ionic solids and even liquids under ambient conditions [2]. This model has thus been used in several studies to provide constraints on the evolution of the electronic structure of several materials under pressure, including estimates of predicted pressures of band gap closure.

The Fabry–Perot interference fringe technique for determination of the refractive index of materials under pressure was pioneered by Grimsditch et al. [3] and applied to solid H_2 by van Straaten and Silvera [4] to estimate the high pressure equation of state, a result that significantly underestimated the density later determined by accurate X-ray diffraction measurements [5]. The method has also been used to measure the refractive indices of dense rare gas solids He, Ne, and Xe [6], and molecular solids H_2O [7], and H_2 [8, 9] at high pressure. Effective oscillator models (EOM) have often been used to extract band gap information from the experimental refractive index data. It is pertinent to test the validity of EOM predictions using direct electronic structure calculations. As will be presented below, the reliability of EOMs strongly dependent on the profile of the empty states of the compressed materials. In solid hydrogen, in which refractive indices have been measured at room and low temperature to multimegabar pressures (up to 320 GPa) and onset of absorption in the visible has been observed at the highest pressures [10, 11], the trend in band gap energies derived from fitting the refractive index dispersion data to EOM is in good agreement with electronic structure theory. In comparison, in ice VIII, even though the theoretical refractive indices at high pressure are in good agreement with experiment, the directly calculated band gaps were found to increase with increasing pressure in contradiction with the values extracted from the EOM analysis.

Solid H_2 holds a special position in high pressure science [12]. It was proposed in 1935 that at sufficiently high pressure, hydrogen would transform from an insulating

molecular solid to a metallic atomic solid [13]. Later, it was hypothesized that metallic hydrogen will be a high temperature superconductor [14]. At low temperature, the metallic phase was proposed to behave like a quantum solid with exotic properties, such as superfluidity [15]. Most theoretical calculations, albeit without the detail knowledge of the crystal structures, suggested hydrogen can be metallized at pressure higher than 300 GPa and subsequent calculations based on model structures have predicted that metallic hydrogen can be superconducting with very high critical temperatures [16, 17]. Because the realization and characterization of metallized solid hydrogen remains an experimental challenge. Direct measurement of band gap at high pressure is technically challenging. Significant progress has been made in recent years that provide constraints on band gap closure [18–20]. The metallization pressure also may be deduced from band gap closure by extrapolation of the gap energies obtained from refractive index measurements at lower pressure [11]. It was estimated that the band gap should close at 14-fold compression or approximately 275 GPa depending on the assumed equation of state [12]. It was also pointed out that is it crucial to distinguish between the behavior of the direct and indirect gaps.

Zha et al. [7] measured the dispersion in refractive index of H_2O ice VII in the visible spectral range [7] up to 120 GPa (i.e., through the ice VII–X transition). The effective band gaps estimated from the EOM was found to decrease with pressure. In addition, the band gap shift shows a quasi-linear relationship with density; from extrapolation into the highly compressed state gives gap closure around 4 TPa, which is well beyond the expected stability field of ice X. Theoretical estimates of the insulator–metal transition range from 700 GPa to 2 TPa [21, 22]. Furthermore, density functional theory (DFT) calculations on a model structure for ice VII found the band gap should increase with pressure to about 12 eV at 300 GPa [22]. Understanding these electronic properties of high pressure ices is relevant to planetary physics. It has been speculated that water exist in high pressure crystalline forms in some planetary bodies may affect the electrodynamics of these planets. Therefore, it is essential to understand and resolve the different trends of the band gap energy derived from refractive index and theoretical DFT calculations.

The purpose of this paper is to assess the reliability of band gap energies obtained from EOM analysis from experimental refractive index with accurate theoretical calculations [23–27]. The dielectric functions [1] of selected model structures of solid H_2 at 100–250 GPa and of H_2O ice at 2.85–88 GPa were calculated from the Bethe–Salpeter equation (BSE) method [27]. The refractive indices were then computed and compared to experiment. The energy gaps were also calculated with the one-shot GW ($G0W0$) method [23–26]. We first present computational details and the selection of the structural models used for solid H_2 and dense ice. This is followed by a discussion of the computational results and comparisons with available experimental data. The paper ends with an assessment of the EOM approach for constraining electronic properties of initially insulating materials at high pressure.

Theoretical Background

Effective Oscillator Model

Relating the calculated frequency-dependent dielectric functions ($\varepsilon(\omega)$) with the measured refractive index, we note that the refractive index is a complex quantity and the optical spectrum is just the imaginary part (ε'') of the frequency-dependent dielectric function [1]. In practice, only this information is needed since the real part $\varepsilon'(\omega)$ can be evaluated through the Kramers–Kronig transformation. The frequency-dependent refractive index ($\tilde{n}(\omega)$) is also a complex quantity and is related to the dielectric function, $\tilde{n}(\omega)^2 = \varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega)$. Wemple and DiDomenico [2] showed that at low frequencies, a large class of solids can be represented by an EOM, $\tilde{n}(\omega)^2 - 1 = E_d E_0 / [E_0^2 - (\hbar\omega)^2]^2$ [2]. This expression defines an "average energy gap" (E_g) of a single-oscillator with energy E_0 in conjunction with a "dispersion-energy" (E_d) that describes the energy-dependent dielectric constant with a threshold absorption energy E_T. These energy parameters are empirical terms determined from the fits to refractive index dispersion data. The parameters, however, are interrelated and determined by the profile of the optical absorption spectrum ($\varepsilon''(\omega)$), which in turn can be characterized by spectral moments (nth moment) [2]

$$M_n = \frac{2}{\pi} \int_{E_t}^{\infty} \omega^n \varepsilon''(\omega) d\omega.$$

It can be shown that, $E_0^2 = M_{-1}/M_{-3}$, $E_d^2 = M_{-1}^3/M_{-3}$, and $E_g^2 = M_1/M_{-1}$. Note that $(E_d/E_0)^{1/2} = (M_{-1})^{1/2}$. The relevance of the spectral moments will become clear in the ensuing discussion.

Many-Body Perturbation Theory

It is well known that the DFT method in the local density approximation (LDA) or gradient corrected density approximation (GGA) often underestimates the electronic band gap energy of solids. This deficiency arises from the neglect of the electron self-energies due to Coulomb screening and many-body correlation effects. A solution to this problem is to employ the GW (single particle Green function G and screened interaction W) [24] approximation on the top of DFT calculations within LDA for the exchange-correlation functional. Recent studies have shown that this combined approach corrects the largest part of the band gap error and improves the agreement of theoretical estimates with experiment. Additional complications are involved in the calculation of the optical absorption spectrum. In the absorption process, a photon is absorbed, exciting an electron from the valence band to the conduction band to create an electron–hole pair. In insulating solids, particularly large band gap materials, this

electron–hole interaction is spatially localized leading to bound state excitations with large oscillator strengths (excitons) below the valence-conduction energy gap. The GW approximation deals only with single-particles state and does not take into account the electron–hole interactions. To this end, a modern approach is to solve the Bethe–Salpeter equation (BSE) [27] of an effective two-particle Hamiltonian describing neutral excitations. In passing and pertinent to the present study, using this technique, Hahn et al. [28] have successfully studied the water system upon condensation and characterized the nature of the blue shift of the absorption edge from the gas phase ∼7.4–8.7 eV in ice I$_h$. Garbuio et al. [29] have reported that the first two peaks of liquid water can only be reproduced with the inclusion of excitonic effects.

Computational Details

The Vienna ab intio simulation package (VASP) [30, 31] and Quantum Espresso (QE) [32] electronic structure codes were used for the hydrogen calculations. For VASP, the 'hard' H_h projected augmented potential (PAW) [33] with Perdew–Bueke–Ernzerhof (PBE) [34] generalized gradient functional (GGA) was used. The energy cutoff for the plane wave expansion was 850 eV. For QE, a norm-conserving H PBE pseudopotential (NCPP) [35] was generated using a radius cutoff of 0.6 a.u. and an energy cutoff of 85 Ry was used in the calculations. The quality of the pseudopotentials and the choices of plane wave cutoff energies have been verified against all electron calculations using the ELK code [36] with the PBE functional on a hypothetical solid hydrogen crystal with a face center cubic structure. As shown in Fig. 9.1, the calculated equations of states by the three different computer codes (*viz.* VASP, QE, and NCPP) are almost identical up to 300 GPa. The k points mesh for Brillouin zone sampling for $Pca2_1$, $Pa3$, $C2/c$, and $Cmca$ were $10 \times 6 \times 6$, $10 \times 10 \times 10$, $8 \times 8 \times 8$ and $8 \times 8 \times 8$, respectively (the description of the structural models will be discussed below). For GW calculations, both the VASP [29] and Yambo [37] codes were used. Since solid molecular hydrogen initially has a large energy gap, excitonic effects are expected to be important. For this reason, the dielectric functions were calculated by the BSE method with the Yambo code where the correlation effects between electron and hole pairs were taken into consideration. Calculations with the random phase approximation (RPA) with and without local field effects did not agree with the more accurate BSE results. It was found calculations including 180 and 60 bands, for GW and BSE receptively, were needed to achieve converged results. In the BSE calculations, corrections of quasiparticle energies from the GGA values were mimicked by a scissor operator [38] with a rigid energy derived from the fits of GW corrected *vs* GGA energies for the valence and conduction bands. Pilot GW calculations on the $Pca2_1$ at selected pressures using VASP and Yambo show the average quasiparticle energy corrections are consistent with each other within 0.25 eV. Calculations were performed from 100 to 50 GPa and from 150 to 250 GPa for model structures considered for phases II and III, respectively.

Fig. 9.1 Equation of states of a hypothetical FCC H lattice (see text) by pseudopotential (NCPP), projector augmented potential (PAW), and all electrons full potential plane wave methods compared with experiment (see Ref. [39] and references therein)
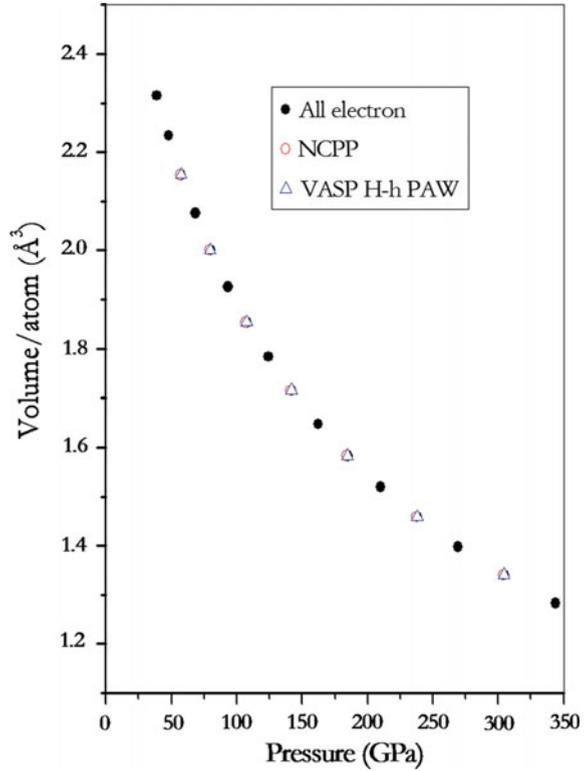
Ice was modeled by a periodic system consists of 16 water molecules with the proton-disorder structural parameters taken from Ref. [40]. Troullier–Martins pseudopotentials [35] were used to describe valence electron–nuclei interactions for H and O. Electronic orbitals were expanded in a plane wave basis set using an energy cutoff of 80 Ry. The PBE [34] exchange-correlation energy functional, which is known to reproduce ground state properties of ice and water [41], was used. At each pressure, the calculations proceeded in three steps: (i) Kohn–Sham eigenvalues and eigenfunctions that enter into the single- and two- particle Green's functions were computed after geometry optimization using the Quantum Espresso package [32]; (ii) the electronic quasiparticle eigenvalues were calculated using GW within the plasmon-pole model within the random phase approximation (RPA); and (iii) the BSE equation was solved for coupled electron–hole excitons. Steps (ii) and (iii) were performed using Yambo, an ab initio plane wave code for MBPT and optical properties calculations [37]. 64 valence and 86 conduction bands were used in the calculation of the Green's functions with a $2 \times 2 \times 2$ k-point set. The selected computational parameters were similar to those used previously [28, 29]. To test the convergence, a larger $4 \times 4 \times 4$ k-point set was used on selected systems and the results were very similar. From preliminary calculations, it was found

that 16,000 plane waves were sufficient to obtain converged self-energies. For the plasmon-pole dielectric function [37], an 800×800 plane wave basis is necessary to achieve convergence converged. In ensuing calculations, 16,000 plane waves and 1000×1000 plasmon-pole dielectric functions were employed.

Results and Discussion

Solid Hydrogen

The lack of precise experimental information on the crystal structures of the high pressure phases of hydrogen have been a challenge to understand the behavior of this important system. At room temperature, hydrogen freezes at 5.4 GPa and adopts a hexagonal close-packed structure with rotationally disordered H_2 molecules (phase I) [12]. Above 100 GPa and temperature below 100 K the 'classical' rotational motions ceased and solid H_2 enters phase II, an orientationally ordered or broken symmetry phase. On further compression at low temperature to 150 GPa, phase II transforms to phase III [42] which is notable for its significantly enhanced infrared H_2 vibron absorption. The strong vibron absorption is attributed to charge transfer between H_2 molecules [42], which further increases with pressure within phase III [42, 43]. When compressed at room temperature the phase I structure persists up to 200 GPa and then transforms to phase III [43]. Several different structures have been proposed for phase II and III, and the exact structure under all conditions (e.g., ortho–para ratio and temperature) is not fully settled. The present work is based largely on calculations reported in Refs. [44–46] In this study, four different rotationally ordered structures known by their space groups were considered: $Pa3$, $Pca2_1$, $C2/c$, and $Cmca$ (Fig. 9.2). The most stable low temperature structures predicted by theoretical calculations in the pressure range of interest here are $Pca2_1$ [44] and $C2/c$ for phases II and III, respectively [46]. Based on spectroscopic constraints [47], the $Pa3$ structure cannot be ruled out for p-H_2 in phase II, and this structure was also assumed in a previous study of the dielectric properties of solid H_2 [6]. The energetically competitive $Cmca$ structure is energetically competitive with $C2/c$ in the region of phase III stability and therefore also considered [46].

The lowest energy electronic transition for the H_2 molecule in the gas phase appears near 14 eV [39] and for the solid at zero pressure appears at about 11 eV [48]. In phase II the energy gap may be expected to be lower but remains indirect due to weak interactions between neighboring H_2. The Wannier function interpolated GW band structure for $Pca2_1$ calculated at 150 GPa supports this expectation (Fig. 9.3a). Figure 9.4 shows the calculated imaginary part of the dielectric functions (ε'') of the $Pca2_1$ structure at selected pressures between 100 and 150 GPa. In the orthorhombic $Pca2_1$, the H_2 molecules are aligned parallel to the xy plane. Therefore, the xx and yy components of $\varepsilon''(\omega)$ are similar but the zz component differs (Fig. 9.3b). Apart from a shift to lower energy of the main excitation (peak maximum) from 10 to
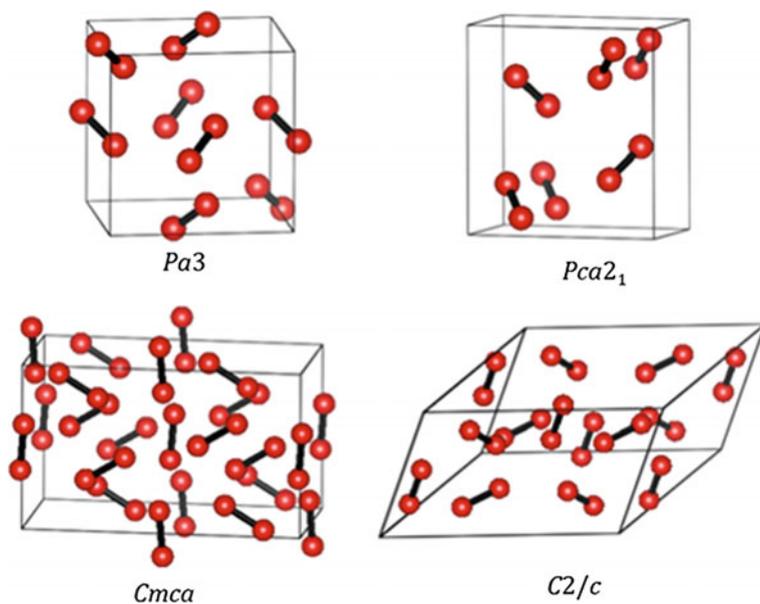
Fig. 9.2 Structural models calculated here for solid hydrogen: $Pa3$, $Pca2_1$, $Cmca$, $C2/c$
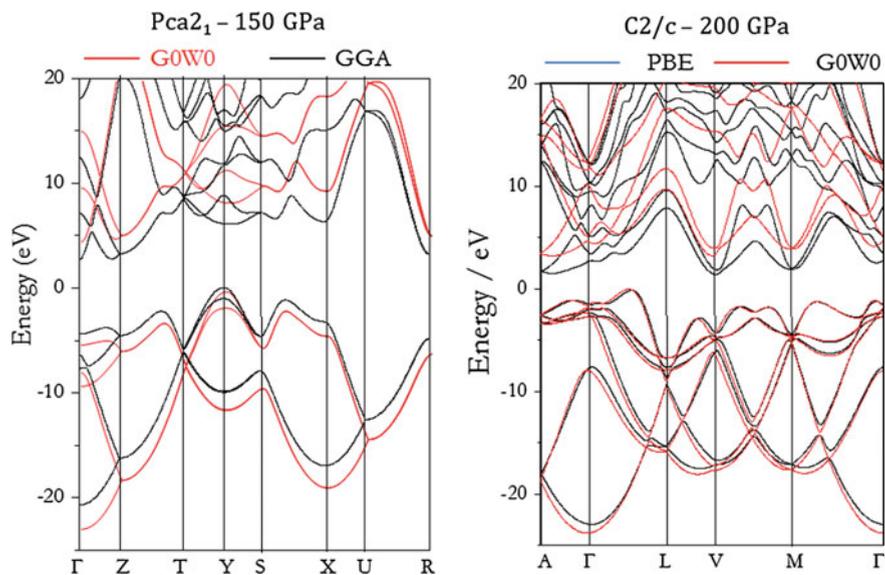
Fig. 9.3 GW and PBE electronic band structure of $Pca2_1$ hydrogen 150 GPa (*left*) and $C2/c$ hydrogen at 200 GPa (*right*)
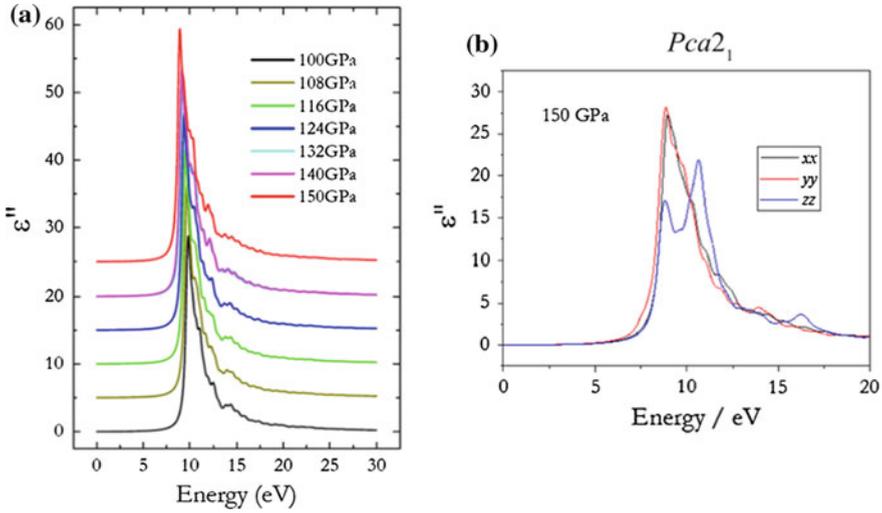
Fig. 9.4 a Calculated imaginary part of the dielectric function ($\varepsilon''(\omega)$) of $Pca2_1$ hydrogen at selected pressures. **b** the xx, yy and zz components of $\varepsilon''(\omega)$ at 150 GPa

8 eV as the pressure increases, the spectral profiles are surprisingly similar. In fact, the calculated full widths at half maximum of the absorption bands increase by only <0.5 eV over the indicated pressure range. Hemley et al. [8] showed that the predicted band gaps and refractive indices of hydrogen are linear in density over a broad range. The refractive index calculated from the theoretical dielectric function between 100 and 150 GPa are compared with experimental results in Fig. 9.5 where both pressure and density behavior are shown. The refractive index increases with pressure in an almost linear manner from 1.9 at 100 GPa to 2.1 at 150 GPa. The rate $dn(\omega)/dP$ is calculated to be 0.006 GPa^{-1}. Within the same pressure region, the experimental refractive index measured at 300 K increased from 2.14 to 2.38 at a rate of 0.005 GPa^{-1}. The theoretical values are about 10% lower than the observed results but the trend is similar. It should be noted that even though the refractive index was measured at room temperature, it was shown that the estimated band gap energies with the EOM are very similar to the low temperature data. To the first approximation, the neglect of the temperature dependence on the density at these pressure is not unreasonable for the present purposes [5, 49].

The relevant indirect and direct band gap energies taken from the GGA and GW calculations are depicted in Fig. 9.6. Between 100 and 150 GPa, the GGA indirect gap decreases from 3.8 to 2.8 eV while the GW gap changes from 6.4 to 5.1 eV or ~2.8 eV cm^{-3} mol. In comparison, the GGA direct gap shows a small increase in energy compared to a gradual decrease from 9.3 to 9.0 eV predicted by GW calculations. In the same pressure region, the effective band gaps derived from the fits to the EOM decrease from 6.6 to 5.2 eV with a rate of 3.0 eV cm^{-3} mol ($dE_g/d(\rho/\rho_0) = 0.96$ eV) [7]. We note that it is not appropriate to compare directly
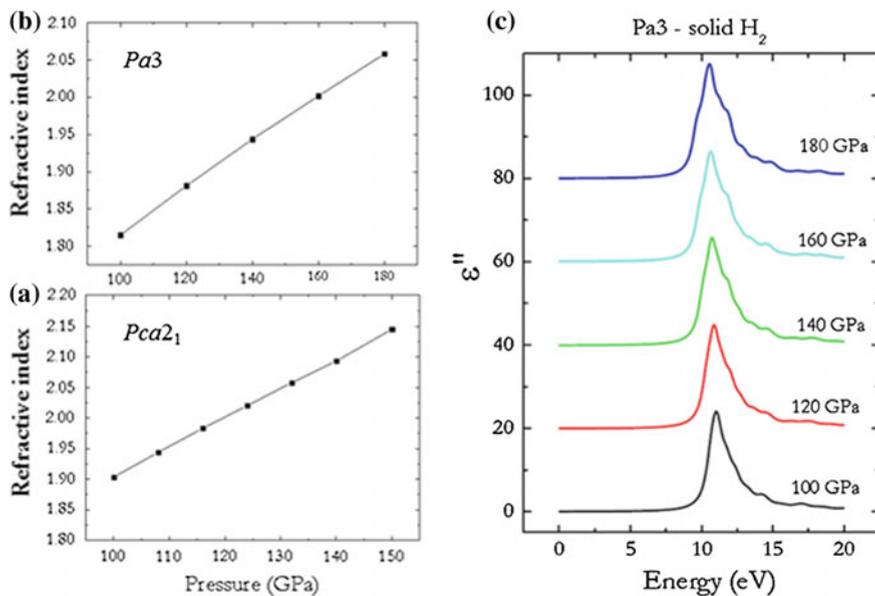
Fig. 9.5 Comparison of the theoretical refractive index for hydrogen in the $Pca2_1$ and $Pa3$ structures (**a**) and (**b**). **c** Calculated imaginary part of the dielectric function ($\varepsilon''(\omega)$) for $Pa3$ hydrogen at selected pressures

Fig. 9.6 GGA and GW direct and indirect band gaps for solid hydrogen in the $Pca2_1$ and $C2/c$ structures as a function of pressure
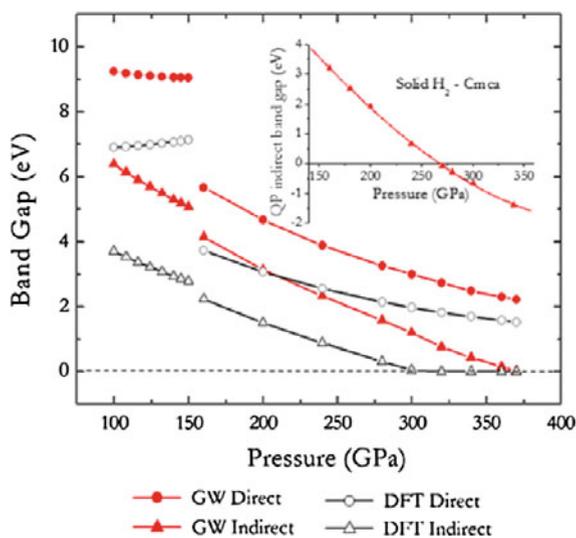
Table 9.1 Spectral moments (M_n, $n = 1, -1$ and -3) of the calculated optical spectrum and dispersion parameters (E_T, E_0, E_d, E_g) related the effective oscillator model for $Pca2_1$ structured solid hydrogen

P (GPa)	M_1	M_{-1}	M_{-3}	E_0	E_d	E_g	$\sqrt{E_d/E_0}$
100	584	4.08	11.6	0.592	3.79	12.0	2.53
108	601	4.31	12.4	0.588	3.97	11.8	2.60
116	618	4.52	13.3	0.584	4.15	11.7	2.67
124	634	4.74	14.0	0.581	4.32	11.6	2.73
132	650	4.95	14.9	0.577	4.49	11.5	2.79
140	667	5.16	15.6	0.575	4.65	11.4	2.85
150	683	5.46	16.9	0.569	4.88	11.2	2.93

the effective gap with the theoretical GW indirect gap. Nevertheless, the decrease in the effective gap energy of ~0.6 eV from 100 to 150 GPa is consistent with the GW indirect band gap values. The rate of 3.0 eV cm^{-3} mol also compares well with the corresponding GW indirect gap value of 2.8 eV cm^{-3} mol. Therefore, the energy gaps extracted from fits to EOM are in good accord with theoretical predictions.

The agreement between the experimental effective band gap energies with the GW indirect gaps can be traced to the lack of variation on the calculated optical spectrum profiles (Fig. 9.4). This fact is exemplified from the calculations of the spectral moments as a function of pressure (Table 9.1). As expected, the dispersive energy, E_d, which governs the distribution of the excitations in the optical spectrum, only changed slightly from 3.79 to 4.88 eV from 100 to 150 GPa. This lack of variation may not affect the simple oscillator model significantly; thus, the results obtained from the dispersion of the refractive indices are reasonable. Incidentally, the effective band gap E_g determined from the analysis of the theoretical optical spectra from 12.0 to 11.2 eV is also in good agreement with the variation of the exciton peaks in the optical spectra (Fig. 9.4). The optical spectra at high pressure become broader, thereby increasing the dispersive energy E_g so that it cannot be treated as a constant parameter in the fit to the EOM. As will be shown later (*vide supra*), similar results are observed in ice VII–X. The smaller overall band width of the excitations in the calculated optical spectra for $Pca2_1$, and $Pa3$ hydrogen can be attributed to small intermolecular interactions and the electronic transitions do not deviate too much from isolated H_2 molecules.

The calculated $\varepsilon''(\omega)$ and refractive index of the $Pa3$ H_2 at 100–150 GPa are shown in Fig. 9.5. The optical spectra and the calculated refractive index are remarkably similar to that of the theoretically predicted $Pca2_1$ structure. Between 100 and 150 GPa, the refractive index again increases linearly from 1.82 to 1.98 with a pressure derivative $dn(\omega)/dP$ of 0.003 GPa^{-1}. In comparison with the $Pca2_1$ model both the absolute value and the pressure derivative of the refractive index are slightly smaller. Once again the profiles of the optical spectra change little with pressure. These results are to be expected since interactions in the $Pa3$ structure at the
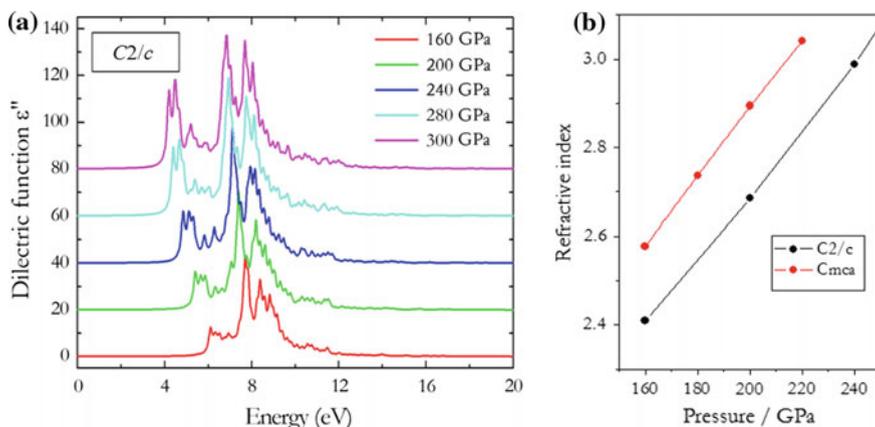
Fig. 9.7 **a** Calculated Imaginary part of the dielectric function ($\varepsilon''(\omega)$) of $C2/c$ solid H_2 at selected pressures. **b** Comparison of the theoretical refractive index for solid H_2 in the $C2/c$ and $Cmca$ structures

densities considered are relatively weak. We further infer that the refractive index is not affected adversely by the details of the crystal structure. Therefore, the band gap trend derived from the EOM is reliable and the gaps energies are close to the GW indirect gap or the absorption band edges. The hexagonal variant of the $Pa3$ structure was considered in a previous theoretical study [9] using the semi-empirical X_α [50] model for the exchange-correlation potential. The dielectric functions were evaluated from first-order perturbation theory. The results are qualitatively similar to the present study. At approximately 93 GPa, the imaginary part of the dielectric function also shows a narrow peak with an absorption edge at 9 eV. The calculated refractive indices were also lower than the experimental data.

Since there is experimental evidence indicating significant intermolecular interactions between H_2 molecules [8, 45] in phase III, one would expect differences in the as-yet unmeasured ultraviolet optical spectra compared to lower pressure phases (see also Ref. [51]). Calculated optical spectrum for the H_2 in the $C2/c$ structure at several pressures are presented in Fig. 9.7a. The calculated onset of the optical excitations is significantly lower than for $Pca2_1$. At 160 GPa, the absorption edge is ~5.6 eV. At the highest pressure studied here (300 GPa), the onset is lowered to 3.6 eV. The theoretical results suggested the absorption edge drops into the visible range at below that of the diamond anvils at the highest pressures and thus observable in experiments at these conditions. Visible absorption is in fact observed at these pressures for phase III, though measurements to date indicate a broad, featureless absorption [10, 11, 20, 43, 52]. In this regard, we note that the absorption edge may depend on the crystal structure. In addition, the optical spectra calculated for are more complex strains and absorptions. Instead of a single strong peak, several peaks spreading over 4–5 eV are predicted and the width of the bands; thus the dispersive

energy E_g is on longer constant. In this regime, the effective band gap determined
from the EOM is not reliable.

The calculated refractive indices of the $C2/c$ and $Cmca$ models are compared in
Fig. 9.7b. At 160 GPa, the theoretical refractive index for $C2/c$ is 2.41 and 2.58 for
the $Cmca$ structure, a sudden jump relative to the $Pca2_1$ or $Pa3$ structures. A larger
refractive index is the consequent of lower onset energy of electronic excitations. It
is important to note that changes in the refractive index with pressure, $dn(\omega)/dP$,
of both structural models are quite similar and close to 0.006 GPa^{-1}. There are
no experimental data available for direct comparison. However, the change of the
refractive index $(1/n(\omega))dn(\omega)/d\omega$ over a broad pressure range up to 220 GPa has
been reported. A discontinuity is noticeable at 150 GPa ($\rho = 0.41$ mol cm^{-3}) [9].

H_2O Ice

Before proceeding to the dense ices, the optical spectra of ice and water at ambi-
ent pressure were computed to verify the computational procedures. The results are
compared with experiment [53–56] and previous theoretical calculations (Fig. 9.8)

Fig. 9.8 Calculated optical spectra for ice I$_h$, and liquid water at ambient pressure and ice VII at selected pressures. The theoretical spectra for ice I$_h$ and water are compared with experimental data (shown on the top of the respective theoretical pattern). For liquid water, the *solid curve* is the measurement by inelastic X-ray scattering [53] and the *short dash curve* from vacuum ultraviolet absorption [54]
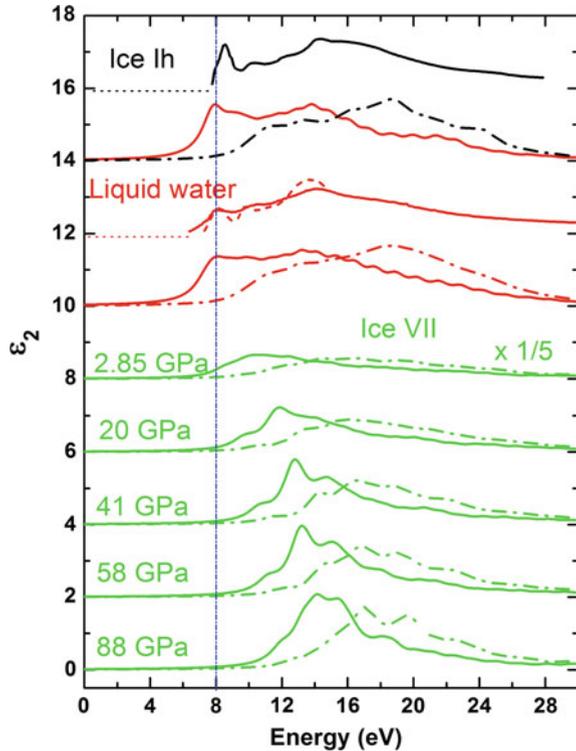
Fig. 9.9 Calculated DFT band gap (*squares*), *GW* band gap (*circles*) and optical band gap (*triangles*) of ice at different volumes compared with results deduced from refractive index data (*solid line*). The dashed lines are guides to the eye for calculated *GW* and *BSE* band gaps. Hydrostatic compression is assumed
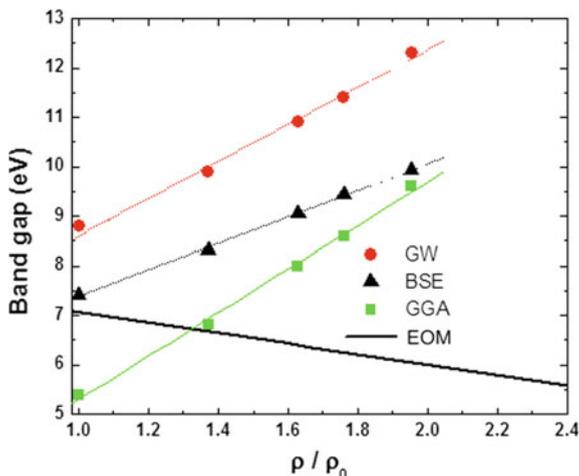
[16, 28]. The *GW* band gap of ice I$_h$ of 8.9 eV is very close to the experimental estimate of ~9.0 eV [57]. The band gap of liquid water averaged over eight configurations randomly selected from the molecular dynamics trajectory of 8.4 eV is within the experimental estimated range of 8.7 ± 0.5 eV [58], consistent with previous *GW* calculations [29, 59]. From the differences between the RPA and BSE spectra, it is clear that excitonic effects are is substantial in these H$_2$O phases.

The calculated optical spectra for ice VII using BSE and *GW*-RPA at selected pressures are shown in Fig. 9.8. The band gap widens due to Madelung effects as the system becomes more ionic at high pressure [22]. It is obvious that excitonic effects are also important in ice at these pressures. The onset of the absorption edge of the calculated optical spectra was found to increase with increasing pressure. The band structure calculations predict that, ice VII has a direct band gap. The *GW* band gap can be determined from the difference between the quasiparticle energies at the top of the valence band and the bottom of the conduction band at Γ. The *GW* band gaps at different pressure are shown in Fig. 9.9. A BSE band gap may be defined as the difference of the energy at the onset of the exciton band and the quasiparticle *GW* band gap (Fig. 9.9). As the pressure increases, the difference between the *GW* and BSE gap grows. Thus, the electron–hole pairs become more strongly bound with density, leading to stronger interactions. The theoretical *GW* band gap increases almost linearly from 8.8 eV at 2.85 GPa to 12.5 eV at 88 GPa.

The above calculations contrast with the results of simple, effective band gap energies derived from refractive index data, which show a decrease from 7 to 5.6 eV in the same pressure range. To explain this discrepancy, the refractive indices of the ice model under pressure were calculated. As discussed above, the optical spectrum is derived from the imaginary part $\varepsilon''(\omega)$ of the frequency-dependent dielectric function, $\varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega) = \tilde{n}(\omega)^2$, and the real part $\varepsilon'(\omega)$ can be calculated through the Kramers–Kronig relations. To assess the numerical accuracy, we also

Table 9.2 Calculated and observed refractive indices (n_{cal}, n_{exp}) at 589.29 nm for ice I_h and water and at 630 nm for ice, moments (M_n, $n = 1$, -1 and -3)) of the optical spectrum and dispersion parameters (E_T, E_0, E_d, E_g) of the effective oscillator model

	n_{cal}	n_{exp}	M_1	M_{-1}	M_{-3}	E_T	E_0	E_d	E_g	$\sqrt{E_d/E_0}$
ice I_h	1.28	1.31	178	0.95	0.008	6.09	10.62	10.11	13.68	0.98
water	1.29	1.33	202	1.01	0.008	6.48	10.98	11.10	14.12	1.00
ice VII (2.85 GPa)	1.44	~1.52	368	1.78	0.012	7.42	11.93	21.20	14.41	1.33
ice VII (41 GPa)	1.60	~1.77	645	2.63	0.013	9.07	14.13	37.17	15.66	1.62
ice X (88 GPa)	1.67	~1.87	792	3.00	0.013	9.94	15.09	45.30	16.25	1.73

Fig. 9.10 Comparison of calculated refractive index dispersion functions of ice at five pressures with experimental data (*inset*) [7]. The *open green triangles* are calculated refractive indices at 88 GPa employing the effective oscillator model with energy parameters extracted from the moments of the theoretical optical spectrum at 88 GPa (color figure online)
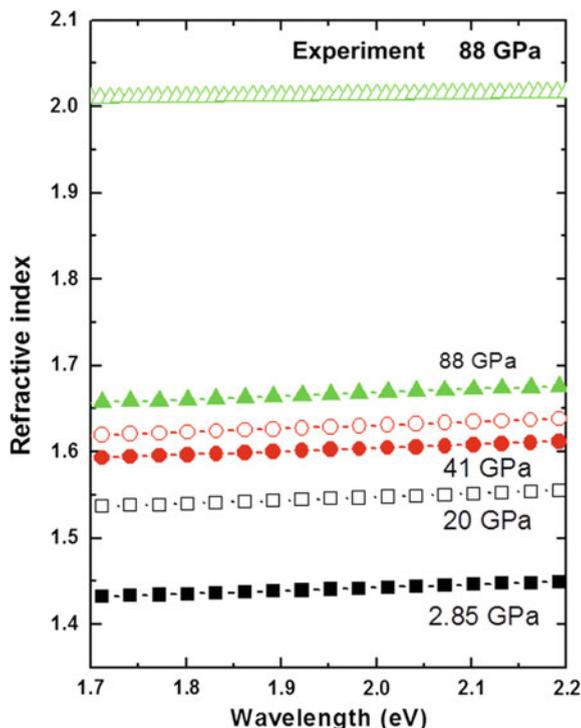
computed the refractive index of ice Ih and water at ambient pressure (Table 9.2). For water at ambient pressure, the calculated value at 589.29 nm from the average of 8 liquid water[1] configurations of 1.29 is close to the experimental value of 1.33 [60]. Similarly, the refractive index of ice I_h at 589.29 nm is calculated to be 1.28, again very close to the observed value of 1.31 [61]. The refractive indices calculated from the theoretically optical spectra of ice VII are compared with the observed values in Fig. 9.10. The predicted pressure trend is similar to experiment, though the calculated refractive indices are systematically lower. The origin of this difference remains to be determined. Notably, the refractive indices are found to increase with pressure. For example at 1.7 eV, from 2.85 to 40 GPa, the refractive index changed from 1.43 to 1.59, a net difference of 0.16 (Fig. 9.10). In comparison, the observed refractive index increased by 0.13, from 1.59 to 1.72 at 6.4 to 34.9 GPa. Thus, the pressure

---

[1]The structures of water were selected from the trajectory of an NVT ab initio molecular dynamics simulations (CPMD) calculation at 320 K using a model consisted of 316 water molecules in a cubic unit cell. The starting structure was taken as ice VII (from Ref. [37]) but the density was scaled to 321.320 g/cm³. A fictitious electron mass of 200 a.u. and a time step of 326 a.u. were used for the integration of ionic motions. Brillouin zone sampling was restricted to the Γ point. After randomization of the coordinates and thermal equilibration, MD was performed for 327.322 ps. Eight snapshots were taken every 320.327 ps.

dependence or the predicted refractive indices are in satisfactory agreement with the experimental analysis, yet the infrared band gap energies differ.

To understand this discrepancy, we consider the assumptions underlying the effective oscillator model. As discussed above, effective oscillator models typically define an "average energy gap" (E_g) of a single-oscillator with energy E_0 in conjunction with a "dispersion-energy" (E_d) to describe the energy-dependent dielectric constant with a threshold absorption energy E_T. To test the validity of the EOM, E_0, E_g and E_d were evaluated for water, ice I$_h$ and ice VII from the moment relationships from the theoretical optical spectra ($\varepsilon''(\omega)$). Results are summarized in Table 9.2. E_T was estimated from the onset of the optical spectrum. The calculated dispersion parameters, E_0 and E_d, for water of 10.98 and 11.10 eV, respectively can be compared with the empirical values of 13.0 and 9.9 eV [7]. The "average" energy gap E_g using the theoretical parameters is 13.86 eV is much larger to the GW energy gap of 8.4 eV. It should be noted that E_g is a hypothetical average gap energy and no specific physical significance can be attached to it. The most revealing result is the effect of pressure on the dispersion-energy E_d. As the pressure increases, E_d becomes larger. The value of E_d at 88 GPa almost double that at 41 GPa. Since, there is a large redistribution of oscillator strengths in the optical spectra. In fact, this is borne out from inspection of the $\varepsilon''(\omega)$ profiles (Fig. 9.10). As pressure increases, $\varepsilon''(\omega)$ becomes broader and the spectral weight shifts to higher energy, and M_{-1} becomes larger than M_{-3}. To a lesser extent, E_0 also increases from 11.93 eV at 2.85 GPa to 15.09 eV at 88 GPa. Within the EOM, the frequency-dependent refractive index ($n(\omega)$) is related to E_0 and E_d, as $n(\omega)^2 - 1 = E_d E_0 / (E_0^2 - \hbar^2 \omega^2)^2$. The refractive index at 88 GPa was computed using the theoretical E_d and E_0 values listed in Table 9.2. The calculated $n(\omega)$ is ∼2.02, which is significantly higher than the first-principles result of ∼1.65. Moreover, the theoretical ratio $(E_d/E_0)^{1/2}$ ($(M_{-1})^{1/2}$) is found to increase with pressure (Table 9.2), in contrast with the almost constant value of 1.5 obtained from the fitting of experimental refractive indices [7]. The analysis shows that in this case the optical spectrum cannot be adequately described by a few energy parameters and the EOM is insufficient to describe the electronic spectrum of ice over this range. A similar result was recently reported by Pan et al. [62].

We emphasize that purely hydrostatic conditions were considered in the present calculations. Whereas these conditions may be reasonable for hydrogen, which remains comparatively weak over a broad range of pressures, ice is a strong solid and supports a significant uniaxial stress beginning at modest pressures [62]. Furthermore, it has been established that the direct band gap of diamond, the strongest known material, increases with pressure, but it decreases as a function of uniaxial strain in different directions [63]. This result is consistent with the decreasing effective absorption edge of strained diamond anvil windows in high pressure experiments [64], and thus the increase in the effective index of refraction that would be inferred if it were assumed to be isotropic.

Conclusions

The electronic band structures and optical spectra of model structure for H_2O ice and solid molecular hydrogen high pressure have been calculated using a combination of first principle methods. Specifically, refractive indices were evaluated from the theoretically determined frequency-dependent dielectric function for the structures. Excitonic effects were found to be very important. In both cases, the predicted trends on the pressure dependence of the refractive index are consistent with experiments. However, in ice the band gap energy is found to increase with pressure, which contrasts with that determined from the effective oscillator model. This discrepancy is shown to be inadequacies in the description of the ultraviolet optical spectrum because of limited empirical energy parameters in the model. In solid hydrogen, the predicted decreasing band gap with pressure agrees with the results obtained from the analysis of experimental refractive index by EOM. The success of the EOM model is due to the electronic excitations that remain molecular-like, with little dispersion in the profile of the optical spectra with pressure for the model lower pressure structure. In contrast, the profile of the optical spectra changes with pressure for denser structures having more strongly interacting molecules; here the EOM is not expected to be valid. In this regard, the EOM is expected to work well for rare gas solids at high pressure. Further improvement on the theoretical approximations may be needed in order to bring the theoretical results into better quantitative agreement with experiment. Moreover, the effects of deviatoric stresses and uniaxial strains need to be included in both the analysis of the experimental data as well as explicitly in the calculations.

References

1. P. Yu, M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 2010)
2. S.H. Wemple, M. DiDomenico, Phys. Rev. B **3**, 1338 (1971). DOI 10.1103/PhysRevB.3.1338
3. M. Grimsditch, R. Letoullec, A. Polian, M. Gauthier, J. Appl. Phys. **60**(10), 3479 (1986). DOI 10.1063/1.337597
4. J. van Straaten, I.F. Silvera, Phys. Rev. B **37**, 6478 (1988). DOI 10.1103/PhysRevB.37.6478
5. R.J. Hemley, H.K. Mao, L.W. Finger, A.P. Jephcoat, R.M. Hazen, C.S. Zha, Phys. Rev. B **42**, 6458 (1990). DOI 10.1103/PhysRevB.42.6458
6. A. Dewaele, J.H. Eggert, P. Loubeyre, R. Le Toullec, Phys. Rev. B **67**, 094112 (2003). DOI 10.1103/PhysRevB.67.094112

7. C.S. Zha, R.J. Hemley, S.A. Gramsch, H.K. Mao, W.A. Bassett, J. Chem. Phys. **126**(7), 074506 (2007). DOI 10.1063/1.2463773
8. R.J. Hemley, M. Hanfland, H.K. Mao, Nature **350**(6318), 488 (1991). DOI 10.1038/350488a0
9. A. García, M.L. Cohen, J.H. Eggert, F. Moshary, W.J. Evans, K.A. Goettel, I.F. Silvera, Phys. Rev. B **45**, 9709 (1992). DOI 10.1103/PhysRevB.45.9709
10. H.K. Mao, R.J. Hemley, Science **244**(4911), 1462 (1989). DOI 10.1126/science.244.4911. 1462
11. P. Loubeyre, F. Occelli, R. LeToullec, Nature **416**(6881), 613 (2002). DOI 10.1038/416613a
12. H.K. Mao, R.J. Hemley, Rev. Mod. Phys. **66**, 671 (1994). DOI 10.1103/RevModPhys.66.671
13. E. Wigner, H.B. Huntington, J. Chem. Phys. **3**(12), 764 (1935). DOI 10.1063/1.1749590
14. N.W. Ashcroft, Phys. Rev. Lett. **21**, 1748 (1968). DOI 10.1103/PhysRevLett.21.1748
15. E. Babaev, A. Sudbø, N.W. Ashcroft, Nature **431**, 666 (2004). DOI 10.1038/nature02910
16. E.G. Brovman, Y. Kagan, A. Kholas, Sov. Phys. JETP **35**, 783 (1972). [Zh. Eksp. Theor. Fyz., **62**(4), 1492 (1972)]
17. P. Cudazzo, G. Profeta, A. Sanna, A. Floris, A. Continenza, S. Massidda, E.K.U. Gross, Phys. Rev. Lett. **100**, 257001 (2008). DOI 10.1103/PhysRevLett.100.257001
18. M.I. Eremets, I.A. Troyan, Nat. Mater. **10**(12), 927 (2011). DOI 10.1038/nmat3175
19. R.T. Howie, C.L. Guillaume, T. Scheler, A.F. Goncharov, E. Gregoryanz, Phys. Rev. Lett. **108**, 125501 (2012). DOI 10.1103/PhysRevLett.108.125501
20. C.S. Zha, Z. Liu, R.J. Hemley, Phys. Rev. Lett. **108**, 146402 (2012). DOI 10.1103/PhysRevLett. 108.146402
21. J. Hama, Y. Shiomi, K. Suito, J. Phys.: Cond. Matter **2**(40), 8107 (1990). DOI 10.1088/0953-8984/2/40/011
22. M. Benoit, M. Bernasconi, P. Focher, M. Parrinello, Phys. Rev. Lett. **76**, 2934 (1996). DOI 10.1103/PhysRevLett.76.2934
23. G. Onida, L. Reining, A. Rubio, Rev. Mod. Phys. **74**, 601 (2002). DOI 10.1103/RevModPhys.74.601
24. L. Hedin, Phys. Rev. **139**, A796 (1965). DOI 10.1103/PhysRev.139.A796
25. M.S. Hybertsen, S.G. Louie, Phys. Rev. Lett. **55**, 1418 (1985). DOI 10.1103/PhysRevLett.55.1418
26. M.S. Hybertsen, S.G. Louie, Phys. Rev. B **34**, 5390 (1986). DOI 10.1103/PhysRevB.34.5390
27. E.E. Salpeter, H.A. Bethe, Phys. Rev. **84**(6), 1232 (1951). DOI 10.1103/PhysRev.84.1232
28. P.H. Hahn, W.G. Schmidt, K. Seino, M. Preuss, F. Bechstedt, J. Bernholc, Phys. Rev. Lett. **94**, 037404 (2005). DOI 10.1103/PhysRevLett.94.037404
29. V. Garbuio, M. Cascella, L. Reining, R.D. Sole, O. Pulci, Phys. Rev. Lett. **97**, 137402 (2006). DOI 10.1103/PhysRevLett.97.137402
30. G. Kresse, J. Hafner, Phys. Rev. B **48**, 13115 (1993). DOI 10.1103/PhysRevB.48.13115
31. G. Kresse, J. Furthmüller, Comput. Mater. Sci. **6**(1), 15 (1996). DOI 10.1016/0927-0256(96)00008-0
32. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A.D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, J. Phys.: Cond. Matter **21**(39), 395502 (2009). DOI 10.1088/0953-8984/21/39/395502
33. G. Kresse, D. Joubert, Phys. Rev. B **59**, 1758 (1999). DOI 10.1103/PhysRevB.59.1758
34. J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996). DOI 10.1103/PhysRevLett.77.3865. [Erratum Phys. Rev. Lett. **78**, 1396 (1997)]
35. N. Troullier, J.L. Martins, Phys. Rev. B **43**, 1993 (1991). DOI 10.1103/PhysRevB.43.1993
36. ELK is an all-electron full-potential linearized augmented-plane wave (FP-LAPW) code. See http://elk.sourceforge.net.
37. A. Marini, C. Hogan, M. Grüning, D. Varsano, Comp. Phys. Comm. **180**(8), 1392 (2009). DOI 10.1016/j.cpc.2009.02.003
38. Z.H. Levine, D.C. Allan, Phys. Rev. Lett. **63**, 1719 (1989). DOI 10.1103/PhysRevLett.63.1719

39. G.J. Schulz, Rev. Mod. Phys. **45**, 423 (1973). DOI 10.1103/RevModPhys.45.423
40. J.L. Kuo, M.L. Klein, J. Phys. Chem. B **108**(51), 19634 (2004). DOI 10.1021/jp0482363
41. K. Laasonen, M. Sprik, M. Parrinello, R. Car, J. Chem. Phys. **99**(11), 9080 (1993). DOI 10.1063/1.465574
42. M. Hanfland, R.J. Hemley, H.K. Mao, G.P. Williams, Phys. Rev. Lett. **69**, 1129 (1992). DOI 10.1103/PhysRevLett.69.1129
43. R.J. Hemley, Z.G. Soos, M. Hanfland, H.K. Mao, Nature **369**(6479), 384 (1994). DOI 10.1038/369384a0
44. C.J. Pickard, R.J. Needs, Nat. Phys. **3**(7), 473 (2007). DOI 10.1038/nphys625
45. J.S. Tse, D.D. Klug, Nature **378**(6557), 595 (1995). DOI 10.1038/378595a0
46. J.S. Tse, D.D. Klug, Y. Yao, Y. Le Page, J.R. Rodgers, Solid State Commun. **145**(1–2), 5 (2008). DOI 10.1016/j.ssc.2007.10.018
47. A.F. Goncharov, R.J. Hemley, H.K. Mao, J. Shu, Phys. Rev. Lett. **80**, 101 (1998). DOI 10.1103/PhysRevLett.80.101
48. K. Inoue, H. Kanzaki, S. Suga, Solid State Comm. **30**(10), 627 (1979). DOI 10.1016/0038-1098(79)90110-8
49. Y.A. Freiman, A.F. Goncharov, S.M. Tretyak, A. Grechnev, J.S. Tse, D. Errandonea, H.K. Mao, R.J. Hemley, Phys. Rev. B **78**, 014301 (2008). DOI 10.1103/PhysRevB.78.014301
50. J.C. Slater, Phys. Rev. **81**, 385 (1951). DOI 10.1103/PhysRev.81.385
51. Z.G. Soos, J.H. Eggert, R.J. Hemley, M. Hanfland, H.K. Mao, Chem. Phys. **200**(1-2), 23 (1995). DOI 10.1016/0301-0104(95)00166-L
52. A.F. Goncharov, J.S. Tse, H. Wang, J. Yang, V.V. Struzhkin, R.T. Howie, E. Gregoryanz, Phys. Rev. B **87**, 024101 (2013). DOI 10.1103/PhysRevB.87.024101
53. H. Hayashi, N. Watanabe, Y. Udagawa, C.C. Kao, Proc. Natl. Acad. Sci. **97**(12), 6264 (2000). DOI 10.1073/pnas.110572097
54. K. Kobayashi, J. Phys. Chem. **87**(21), 4317 (1983). DOI 10.1021/j100244a065
55. G.D. Kerr, R.N. Hamm, M.W. Williams, R.D. Birkhoff, L.R. Painter, Phys. Rev. A **5**, 2523 (1972). DOI 10.1103/PhysRevA.5.2523
56. T. Shibaguchi, H. Onuki, R. Onaka, J. Phys. Soc. Jpn. **42**(1), 152 (1977). DOI 10.1143/JPSJ.42.152
57. A. Hermann, W.G. Schmidt, P. Schwerdtfeger, Phys. Rev. Lett. **100**, 207403 (2008). DOI 10.1103/PhysRevLett.100.207403
58. A. Bernas, C. Ferradini, J.P. Jay-Gerin, Chem. Phys. **222**(2.3), 151..160 (1997). DOI 10.1016/S0301-0104(97)00213-9
59. D. Lu, F. Gygi, G. Galli, Phys. Rev. Lett. **100**, 147601 (2008). DOI 10.1103/PhysRevLett.100.147601
60. K. Vedam, P. Limsuwan, J. Chem. Phys. **69**(11), 4762 (1978). DOI 10.1063/1.436529
61. V.F. Petrenko, R.W. Whitworth, *Physics of Ice* (Oxford University Press, Oxford, 1999)
62. D. Pan, Q. Wan, G. Galli, Nat. Commun. **5** (2014). DOI 10.1038/ncomms4919
63. O.H. Nielsen, Phys. Rev. B **34**, 5808 (1986). DOI 10.1103/PhysRevB.34.5808
64. H.K. Mao, R.J. Hemley, Nature **351**(6329), 721 (1991). DOI 10.1038/351721a0

Chapter 10
Nonequilibrium Steady States and Electron Transport in Molecular Systems

I. Deretzis, S.F. Lombardo, G.G.N. Angilella, R. Pucci and A. La Magna

Abstract We discuss theoretically the correlation between atomic configurations and coherent stationary electronic transport in molecular systems contacted with metallic electrodes in stationary conditions. The nonequilibrium Green's function formalism is discussed as preliminary theoretical framework which allows, by means of a suitable mapping of electron-only Hamiltonians, the bridging between the atomic structure of the molecule+contacts system and the electron dynamics. Contrarily to the equilibrium case, the lack of a variational principle in nonequilibrium conditions, even when the stationary state is achieved, does not permit a correct formal approach to the correlation between the electron state in a biased molecular device and the molecule's configuration. A conjecture is proposed to extend the free energy Mermin's functional, and the related variational procedure, for a system in contact with two electron reservoirs ruled by two different electrochemical potentials. Results of the proposed variational procedure are presented for realistic molecular systems, and

I. Deretzis (✉) · S.F. Lombardo · A. La Magna
Istituto per la Microelettronica e Microsistemi (CNR-IMM), Z.I. VIII Strada 5,
95121 Catania, Italy
e-mail: ioannis.deretzis@imm.cnr.it

A. La Magna
e-mail: antonino.lamagna@imm.cnr.it

S.F. Lombardo · G.G.N. Angilella · R. Pucci
Dipartimento di Fisica e Astronomia, Università di Catania, Via S. Sofia, 64,
95123 Catania, Italy
e-mail: giuseppe.angilella@ct.infn.it

R. Pucci
e-mail: renato.pucci@ct.infn.it

G.G.N. Angilella · R. Pucci
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella · R. Pucci
IMM-CNR, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
INFN, Sez. Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
Scuola Superiore di Catania, Università di Catania, Via Valdisavoia, 9, 95123 Catania, Italy

the predicted anomalous effects of the biasing on both the atomic configurations and the transport features are discussed.

Introduction

Fundamental investigations of the nonequilibrium electron kinetics in nano- and molecular- systems in contact with macroscopic electrodes have been strongly solicited by the recent advancements in the fields of nano- and molecular- electronics [1]. Moreover, these studies could also significantly support the electronic characterization techniques with atomistic resolution, like tunneling spectroscopy [2]. Contrary to the case either of conventional or of mesoscopic electron devices, these systems manifest transport features which strongly depend on their internal atomic configurations and on the atomic structure in the contact regions. Moreover, also spin and charge instabilities of the electron system have a crucial role on the electrical characteristic of these "few atom" devices [3].

The electron phonon interaction or, more generically, the effects of nuclear dynamics on the electron could impact the electron transport in molecular components in a way which cannot be quantitatively addressed by any perturbative theory. In particular, molecular devices manifest peculiar nonlinear characteristics (e.g., negative differential resistance (NDR) or switching) [4–6], which can only be properly modeled by releasing the common assumptions of the electron transport theory: i.e., (a) the derivation of the electron structure "freezing" the nuclei at the equilibrium position, and (b) the kinetic pertubative approach to evaluate the scattering rates. Recent alternative transport theories have indicated that the strong coupling between a "nonstatic" nuclear configuration in molecular systems and the "nonequilibrium" electron population is a possible explanation of these effects [4, 6, 7]. As a consequence, considering model Hamiltonians, anomalous electron dynamics have been obtained in the framework of nonequilibrium polaronic-like (i.e., variational) solutions, derived in the adiabatic [5, 7], anti-adiabatic [8], and intermediate [4] regimes of the interactions between electron and molecule deformations.

In this chapter these issues will be reviewed discussing the status of the methodologies and the unsolved issues which merit further in-depth analyses. This chapter is organized as follows. In section "Quantum electron transport formalism," the nonequilibrium Green's function formalism is presented as the reference approach to quantify the coherent electron transport in molecular systems. The importance of a correct determination of the atomic configuration for a contacted molecule is discussed in the beginning of section "Atomic configurations and space distribution of mean-fields in nonequilibrium." In the same section the problem of the variational principle in nonequilibrium conditions induced by the biasing is discussed, and a suitable approach to solve this question is proposed and applied to polyyne-based devices. In section "Generalized trial states in nonequilibrium variational theories" the importance of the choice of the correct variational state for the whole range of

the coupling strength between molecules and contacts to study the electron transport is analyzed. Finally, section "Conclusions" contains some concluding remarks.

Quantum Electron Transport Formalism

Electron transport in molecular systems cannot be correctly addressed by any semi-classical approach generalizing the well-known drift diffusion model extensively applied to the study of the behavior of electronic devices. Indeed, a full quantum transport formalism is needed when the system size is well beyond the scattering and coherence characteristic lengths as in the case of a molecular system [9]. A reliable and convenient approach relies on the combination of the effective single-particle scattering theory using Green's function techniques with the Landauer derivation for the transmission probability and the current [10]. This purely quantum theory is very versatile since it allows immediate generalizations and further advancements compatible to the underlying mean-field approximation. A brief overview of this formalism, prerequisite of the subsequent theoretical derivations, is presented in this Section.

Nonequilibrium Green's Function

The Green's function technique was introduced to isolate, and approximatively compute, the effects of the part of the Hamiltonian (often the many-body component) which does not allow a direct, possibly analytical, solution of the quantum problem under study. In its nonequilibrium version this isolated interaction-like part includes also the alteration to the quantum state of the system under study due to the presence of contacts between a finite and two or more infinite systems [10–14].

Here, we consider the transport problem for a two-terminal geometry, where a finite device part is embedded between semi-infinite left and right contacts, which do not interact directly; however, the formalism can be extended to geometries with more than two terminals. Assuming that, using a suitable basis, a matrix representation of the full Hamiltonian can be obtained as [9]

$$H = \begin{pmatrix} H_L & -\tau_L & 0 \\ -\tau_L^\dagger & H_0 & -\tau_R^\dagger \\ 0 & -\tau_R & H_R \end{pmatrix}, \tag{10.1}$$

where H_0 is the finite Hamiltonian of the molecular device, and $H_{L,R}$ are the (semi-infinite) bulk contact Hamiltonians, while $\tau_{L,R}$ describe the coupling between contacts and the device. In the case of an orthonormal basis set, the Green function $\mathscr{G}$ can be defined as the solutions to the following formal equation

$$(EI - H)\mathscr{G}(E) = I, \tag{10.2}$$

where E is the energy and I is the unitary matrix, which has to be replaced by the overlap matrix between basis functions for a nonorthonormal basis. $\mathscr{G}$ is the response in one state (e.g., orbital + position) of a system within a multiterminal geometry from an excitation that takes places in a different state of the same system. It could be formally related to the scattering matrix (S-matrix) [10], albeit its applicability is more direct and general. Due the formal pole structure of Eq. 10.2, two equivalent solutions, which are called the advanced ($\mathscr{G}^A$) and the retarded ($\mathscr{G}^R$) Green's functions, can be derived, and they physically represent an incoming and an outgoing wave that are formed due to the same excitation. In the continuum counterpart of Eq. 10.2 we can obtain mathematically one of the two solutions by imposing suitable boundary conditions, which in our case correspond to adding or subtracting an infinitesimal imaginary part to the energy. It is also straightforward to show that $(\mathscr{G}^R)^\dagger = \mathscr{G}^A$ [9].

We can separate the Green's function that corresponds to the molecular system ($\mathscr{G}_0$) and the contacts ($\mathscr{G}_{L,R}$), and by substituting in Eq. (10.2) we have

$$\begin{pmatrix} EI - H_L & -\tau_L & 0 \\ -\tau_L^\dagger & EI - H_0 & -\tau_R^\dagger \\ 0 & -\tau_R & EI - H_R \end{pmatrix} \begin{pmatrix} \mathscr{G}_L & \mathscr{G}_{L0} & \mathscr{G}_{LR} \\ \mathscr{G}_{0L} & \mathscr{G}_0 & \mathscr{G}_{0R} \\ \mathscr{G}_{RL} & \mathscr{G}_{R0} & \mathscr{G}_R \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{pmatrix}, \tag{10.3}$$

which is equivalent to the following set of coupled matrix equations:

$$(EI - H_L)\mathscr{G}_{L0} - \tau_L \mathscr{G}_0 = 0, \tag{10.4a}$$

$$-\tau_L^\dagger \mathscr{G}_{L0} + (EI - H_0)\mathscr{G}_0 - \tau_R^\dagger \mathscr{G}_{R0} = I, \tag{10.4b}$$

$$(EI - H_R)\mathscr{G}_{R0} - \tau_R \mathscr{G}_0 = 0. \tag{10.4c}$$

In this coupled system Eqs. (10.4a) and (10.4c) give matrices with (∞, N) dimensions, whereas Eq. (10.4b) has (N, N) dimensions, where N is the dimension of the finite space where the molecular device Hamiltonian makes its action. If we solve for $\mathscr{G}_{L0}$ and $\mathscr{G}_{R0}$, we have

$$\mathscr{G}_{L0} = (EI - H_L)^{-1}\tau_L \mathscr{G}_0 = g_L \tau_L \mathscr{G}_0 \tag{10.5a}$$

$$\mathscr{G}_{R0} = (EI - H_R)^{-1}\tau_R \mathscr{G}_0 = g_R \tau_R \mathscr{G}_0, \tag{10.5b}$$

where $g_{L,R}$ are the Green functions of the left and right contact, respectively. These two (∞, N) matrices (in the real-space) can be projected onto (N, N) ones by substitution back into Eq. (10.4b) in order to obtain a finite expression. Such a procedure usually takes advantage of the semi-infinite lattice symmetry of the two leads allowing a **k**-space transformation of the real-space matrices $g_{L,R}(\mathbf{r})$ into $g_{L,R}(\mathbf{k})$ followed by their backward substitution in real coordinates. For a specific case, this can be done either analytically [10], while iterative procedures allow the numerical approach to this problem in the generic case [15]. The obvious advantage of such an approach is

the use of (N, N) matrices incorporating the effect of infinite ones. From Eq. (10.4b) we now obtain

$$- \tau_L^\dagger g_L \tau_L \mathscr{G}_0 + (EI - H_0)\mathscr{G}_0 - \tau_R^\dagger g_R \tau_R \mathscr{G}_0 = I. \tag{10.6}$$

If we now define the *self-energies* of the two contacts as

$$\Sigma_L = \tau_L^\dagger g_L \tau_L \tag{10.7a}$$
$$\Sigma_R = \tau_R^\dagger g_R \tau_R, \tag{10.7b}$$

we can obtain the final version of the nonequilibrium Green's function matrix of the system that takes also account for the interaction with the two contacts via the two self-energy terms:

$$\mathscr{G} = (EI - H_0 - \Sigma_L - \Sigma_R)^{-1}. \tag{10.8}$$

Conceptually, one can think of $\Sigma_{L,R}$ as matrices that account also for the exact effect of the leads in terms of effective scattering terms. As a consequence, they are formally considered as other types of interactions that take place in the molecular device during conduction (e.g., electron–phonon interactions [10, 16]), although such extensions, to be consistent with the one particle description here presented, turn out to be only approximate.

In the following subsections, we can therefore focus on the quantities relevant to transport that can be derived once the Green function is computed.

Spectral Functions

The spectral functions A and Γ of the device and the contact respectively are defined by the anti-Hermitian part of the Green's function and the self-energy's function respectively:

$$A = i(\mathscr{G} - \mathscr{G}^\dagger) \tag{10.9a}$$
$$\Gamma = i(\Sigma - \Sigma^\dagger). \tag{10.9b}$$

Usually, only A is referred to as the spectral function, and is related to the density of states of the system, while the Γ matrix is also called broadening matrix and physically represents the strength with which the contacts are bound to the device.

For the derivation of the relationship between the spectral function and the density of states we need to expand Green's function in the set of eigenfunctions. Without lack of generality we start from an orthonormal set:

$$\langle \psi_\beta, \psi_\alpha \rangle = \delta_{\beta,\alpha}, \tag{10.10}$$

where $\delta_{\beta,\alpha}$ is a Kronecker delta. From Eq. (10.2) and considering the retarded Green's function by adding an infinitesimal imaginary part to the energy, we obtain:

$$\mathscr{G}^R = \sum_\alpha \frac{|\psi_\alpha\rangle\langle\psi_\alpha|}{E + i\eta - \varepsilon_\alpha}, \tag{10.11}$$

where η is an infinitesimal positive number, and ε_α is the eigenvalue corresponding to the eigenfunction $\psi_\alpha(\mathbf{r})$. From Eqs. (10.9a) and (10.11), we have:

$$\begin{aligned} A &= i \sum_a \left(\frac{1}{E + i\eta - \varepsilon_\alpha} - \frac{1}{E - i\eta - \varepsilon_\alpha} \right) |\psi_\alpha\rangle\langle\psi_\alpha| \\ &= \sum_a \frac{2\eta}{(E - \varepsilon_\alpha)^2 + \eta^2} |\psi_\alpha\rangle\langle\psi_\alpha|. \end{aligned} \tag{10.12}$$

Since $\eta \to 0^+$, by integrating over E with a test function, we obtain [10]:

$$A = 2\pi \sum_a \delta(E - \varepsilon_\alpha)|\psi_\alpha\rangle\langle\psi_\alpha|. \tag{10.13}$$

We know that the expression for the density of states of the system is [10]:

$$D = \sum_a |\psi_\alpha|^2 \delta(E - \varepsilon_\alpha). \tag{10.14}$$

From the last two equations we obtain:

$$2\pi D = \mathrm{Tr}\, A. \tag{10.15}$$

Similarly, the broadening matrix can also be related to the electronic structure of the system. As its name suggests, it gives rise to a broadening of the energy levels in the channel, proportional to the strength of the coupling between the device and the contacts. Furthermore, it introduces a finite lifetime for the electronic states, suggesting that an electron introduced into a state after some time should escape to the contact. Reference [10] proposes a detailed mathematical and conceptual derivation of these physical aspects starting from simple toy models all the way up to second quantization arguments. Finally, a useful identity between A and Γ that can be found easily is [16]

$$A = \mathscr{G}\Gamma\mathscr{G}^\dagger = \mathscr{G}^\dagger\Gamma\mathscr{G}. \tag{10.16}$$

Nonequilibrium Electron Density

The electron density distribution in nonequilibrium conditions differs from the equilibrium one, since, when the device is in contact with reservoirs at different Fermi energies, each one of them tries to align the electrochemical potential device to its own Fermi level by injecting or subtracting electrons. The density distribution can be computed from the density (matrix) operator, which, if we consider the system in equilibrium (i.e., the molecular device connected to just one contact), is given by

$$\rho = \sum_{\alpha} f_0(\varepsilon_\alpha - \mu)|\psi_\alpha\rangle\langle\psi_\alpha|, \tag{10.17}$$

where μ is the electrochemical potential of the contact, and

$$f_0(\varepsilon_\alpha - \mu) = \frac{1}{1 + e^{(\varepsilon_\alpha - \mu)/k_B T}} \tag{10.18}$$

is the statistical Fermi–Dirac distribution of electrons. Equation (10.17) can be rewritten within the single-particle scattering theory notation as [10]

$$\begin{aligned}
\rho &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dE \, f_0(E - \mu) A \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dE \, f_0(E - \mu)[i(\mathscr{G} - \mathscr{G}^\dagger)] \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} dE \, f_0(E - \mu)\mathscr{G}\Gamma\mathscr{G}^\dagger.
\end{aligned} \tag{10.19}$$

If we generalize now the previous equation to the case of two contacts, we obtain the final form for the equation of the density operator as

$$\rho = \frac{1}{2\pi} \int_{-\infty}^{\infty} dE \, (f_L \mathscr{G}\Gamma_L \mathscr{G}^\dagger + f_R \mathscr{G}\Gamma_R \mathscr{G}^\dagger), \tag{10.20}$$

where $f_{L,(R)}$ is the Fermi–Dirac function of the left (right) contact.

Transmission

The transmission formalism was developed by Landauer and afterwards expanded by Büttiker for the calculation of current in point-like contacts/devices on the basis of the summation of the *transmission probability* of each *transverse mode* of the device channel [10]. This theory leads to the well-known Landauer's formula for the conductance:

$$G = \frac{2e^2}{h} M T. \tag{10.21}$$

Here, G is the conductance, M the number of transverse (or propagating) modes of the device, and T the average probability that an electron injected from the one lead will transmit to the other. By defining $T(E) = MT$ as the total transmission probability of the channel for energy E, it is possible to demonstrate [10] that in the case of coherent transport the total current can be derived from the relation:

$$I = \frac{2e}{h} \int_{-\infty}^{\infty} dE \, T(E) [f_L - f_R]. \tag{10.22}$$

Here, the contact Fermi functions represent the driving force for the propagation of current in the sense that only carriers with energies in between the electrochemical potentials of the two leads participate in the conduction process. The transmission here is a general formal concept that can be computed within the single-particle scattering formalism.

Out of equilibrium, the current that passes from a contact to the device is the difference of an in-flux and an out-flux current [10]. We can write the out-flux from the device to the left lead as:

$$I_{\text{out},L} = \frac{e}{h} \int_{-\infty}^{\infty} dE \, \text{Tr}(\Gamma_L \rho), \tag{10.23a}$$

and the out-flux from the device to the right lead as:

$$I_{\text{out},R} = \frac{e}{h} \int_{-\infty}^{\infty} dE \, \text{Tr}(\Gamma_R \rho), \tag{10.23b}$$

where $\Gamma_{L,(R)}/h$ represent the *rate* at which electrons escape from the device to the left (right) contact (considering that Γ matrices have the dimensions of energy), and ρ is the density matrix given by Eq. (10.20). At equilibrium, the in-flux current from the contacts to the device is equal to the out-flux current:

$$I_{\text{in},L} = \frac{e}{h} \int_{-\infty}^{\infty} dE \, \text{Tr}(\Gamma_L \rho_{eq}) \tag{10.24a}$$

$$I_{\text{in},R} = \frac{e}{h} \int_{-\infty}^{\infty} dE \, \text{Tr}(\Gamma_R \rho_{eq}), \tag{10.24b}$$

where

$$\begin{aligned} \rho_{eq} &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dE \, (f_L [\mathscr{G}\Gamma_L\mathscr{G}^\dagger + \mathscr{G}\Gamma_R\mathscr{G}^\dagger]) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dE \, (f_R [\mathscr{G}\Gamma_L\mathscr{G}^\dagger + \mathscr{G}\Gamma_R\mathscr{G}^\dagger]), \end{aligned} \tag{10.25}$$

since $\mu_L = \mu_R$, at equilibrium. In the nonequilibrium case the net current for contact L is

$$I = I_{\text{in},L} - I_{\text{out},L} = \frac{e}{h} \int_{-\infty}^{\infty} dE \ \text{Tr}(\Gamma_L \mathscr{G} \Gamma_R \mathscr{G}^{\dagger})(f_L - f_R). \quad (10.26\text{a})$$

With a similar procedure, for the second contact we can derive that

$$I = I_{\text{in},R} - I_{\text{out},R} = \frac{e}{h} \int_{-\infty}^{\infty} dE \ \text{Tr}(\Gamma_R \mathscr{G} \Gamma_L \mathscr{G}^{\dagger})(f_L - f_R). \quad (10.26\text{b})$$

If we multiply Eqs. (10.26a) and (10.26b) by 2, in order to account for spin degeneracy, and compare them with Eq. (10.22), we can write the transmission probability as

$$\begin{aligned} T(E) &= \text{Tr}(\Gamma_L \mathscr{G} \Gamma_R \mathscr{G}^{\dagger}) \\ &= \text{Tr}(\Gamma_R \mathscr{G} \Gamma_L \mathscr{G}^{\dagger}) \\ &= \text{Tr}(\Gamma_L A_R) \\ &= \text{Tr}(\Gamma_R A_L). \end{aligned} \quad (10.27)$$

Electronic Hamiltonians

The theoretical evaluation of quantum transport properties follows the preliminary determination of the electronic structure of the system under study by determining the equilibrium atomic configuration and using the proper electronic Hamiltonian. To this aim, ab initio and semiempirical approaches can be applied with an accuracy/efficiency mismatch between them. Ab initio (or *first principles*) methods use only fundamental physical constants and calculate self-consistently the wavefunction of the studied system on the basis of the (equilibrium) variational principle of the energy minimization. *Semiemprirical* methods on the other hand use parametrized Hamiltonians that have to be calibrated *ad hoc* for the system under study on the basis of experimental data or first-principles calculations.

It is noteworthy that the concept of self-consistency, which is on the basis of the ab initio approach, should be partially extended also to the case of a transport study applying *semiemprirical* models due to the effect of nonequilibrium charging of the molecular device (i.e., the deviation of the electron density distribution with respect to the equilibrium case). This additional electron interaction can be included in the single-particle scattering theory, here discussed by means of a self-consistent potential U_{sc} that is a functional of the electron density [17]

$$H = H_0 + q U_{sc}, \quad (10.28)$$

where H_0 is the bare device Hamiltonian without the presence of nonequilibrium electron–electron interactions and q is the electronic charge. The calculation of U_{sc} takes places by numerically solving the Poisson equation

$$\nabla^2 U_{sc} = -\rho_f / \varepsilon, \qquad (10.29)$$

where ρ_f is the nonequilibrium density of mobile charges that can be easily calculated within the Green function method [17], and ε is the dielectric constant of the material in which the device part is embedded.

Summary

The nonequilibrium Green function is a quantum mechanical tool, in the single-particle or mean field approximation, for the theoretical study of transport in molecular devices that can incorporate efficiently all the scattering mechanisms (contacts, electron–electron, electron–phonon). Derivation of all quantities relevant for transport can take place with the appropriate definition of the Green matrix. This involves the electronic Hamiltonian written at an appropriate basis set along with information on electron–electron interactions (self-consistent potential U_{sc}), scattering by the contacts (self-energy matrices $\Sigma_{L,R}$), and other scattering mechanisms (self-energy matrices Σ_{ext}). From the Green matrix derivation of the density of states, transmission probability, charge carrier density, and current becomes straightforward. The theory can also be extended to calculate conduction for systems where symmetry-breaking mechanisms (e.g., spin or charge density wave) could occur. On the other hand, criticism to the Landauer–Büttiker approach derives from the intrinsic single-electron picture of the transmission probability that does not take into account for many-body effects, which are enhanced in the nanoscale. Moreover, the variational principle based on the energy functionals is strictly valid in thermodynamic equilibrium conditions (i.e., in the presence of a single lead/reservoir), and eventual alterations of the atomic configuration or order parameters due to nonequilibrium effects cannot be properly computed, not even in stationary condition, being usually assumed frozen to the ones estimated in equilibrium conditions.

Atomic Configurations and Space Distribution of Mean-Fields in Nonequilibrium

Electronic structure and transport of molecular systems are strongly related to the atomic configurations of the contacts + device systems. As example of the relevance of such effects, in Figs. 10.1 and 10.2 we show the modification of the equilibrium transmission spectra for a carbon nanotube resistor contacted to two gold leads when
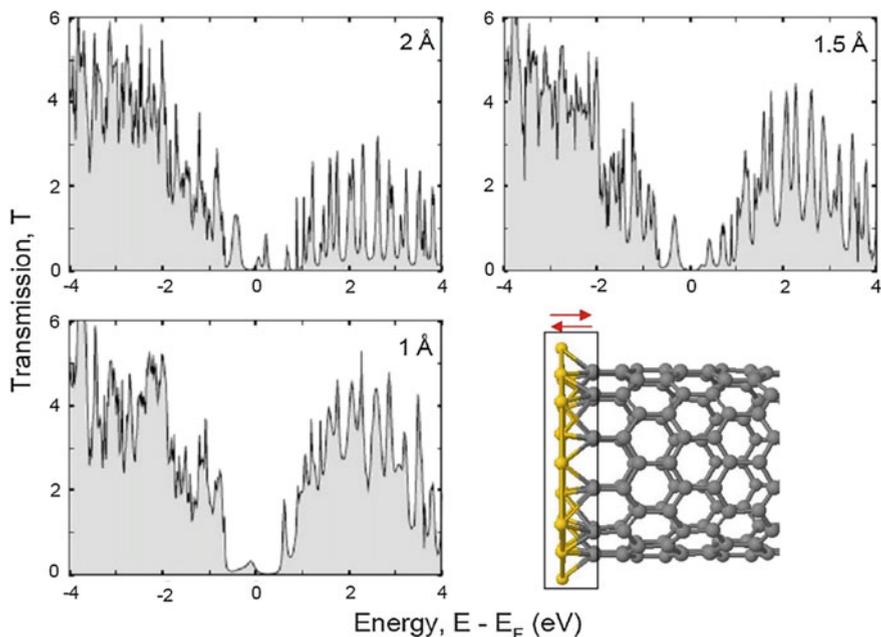
Fig. 10.1 Transmission as a function of energy for four snapshots of a $(10, 0)$ 8-unit-cell system with a varying contact distance of 2, 1.5, 1 Å

the atomic configurations in the contact region only are modified by changing the average distance (Fig. 10.1), or the bond symmetry at fixed distance (Fig. 10.2; see Ref. [17] for a complete discussion on this issue).

This interconnection between electron transport and atomic configuration is enhanced in the presence of a symmetry-breaking induced by electron–electron and electron-deformation (e.g., phonon) interactions. For example, carbon based chains, both pure (i.e., polyyne) or conjugated (i.e. polyacetylene) could manifest dimerized (stable) and symmetric (metastable) states with semiconductor and metallic behavior respectively, whereas the instability of the symmetric configuration is caused by the electron–phonon interaction. In general, instabilities of the electron system could lead to a formation of charge or spin density waves (e.g., an antiferromagnetic state, as in the case of graphene ribbons with zig–zag geometry [18–20]) with the consistent gap opening in the electron structure at the Fermi level, where for a symmetric configuration the electron structure should have a nonzero density of states. In some cases these density waves could be also related to oscillating bond length variations.

The study of the instabilities of many electron systems is a wide and well studied topic of modern solid-state physics. However, thermodynamic equilibrium is usually assumed for the theoretical approaches applied to study this issue; including the powerful variational methodologies often used to calculate self-consistently optimal atomic configurations and/or the space distribution of the mean value of the order
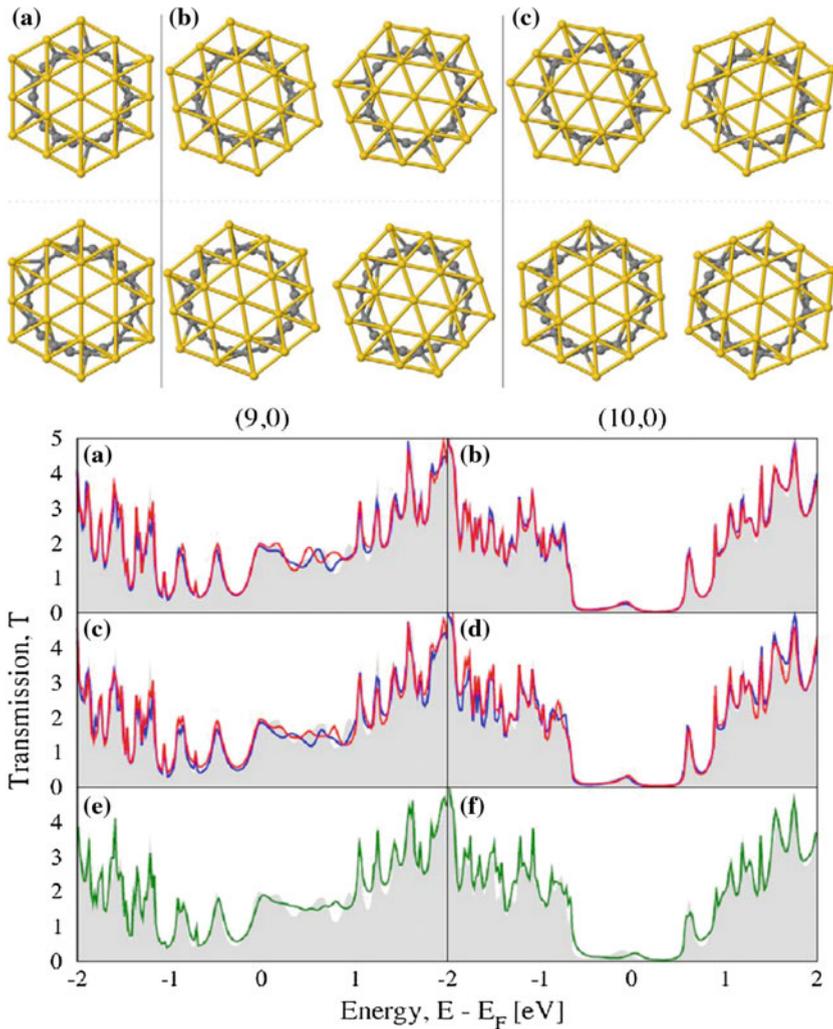
Fig. 10.2 (*Upper panel*) Contact-nanotube interface configurations used to calculate transmission spectra shown in the *lower panel* for a (9, 0) (*upper line*) and a (10, 0) (*lower line*) CNT: **a** 'standard' configuration used throughout this paper; **b** anticlockwise rotation of the contact with respect to the CNT by 10° and 20° for the (9, 0) CNT, and by 9° and 18° for the (10, 0) CNT; **c** an 'arbitrary' rotation and movement of the contacts with respect to the center of the CNTs (the two figures per CNT represent left and *right* contacts). (*Lower Panel*) Transmission as a function of energy for a (9, 0) and a (10, 0) CNT with altered contact-tube geometrical configurations with respect to previous calculations (*gray fillings*): **a** (9, 0) CNT with one contact rotated in an anticlockwise manner by 10° (*blue line*) and 20° (*red line*); **b** (10, 0) CNT with one contact rotated in an anticlockwise manner by 9° (*blue line*) and 18° (*red line*); **c** (9, 0) CNT with both contacts rotated by 10°, −10° (*blue line*) and 10°, −20° (*red line*), respectively; **d** (10, 0) CNT with both contacts rotated by 9°, −9° (*blue line*) and 9°, −18° (*red line*), respectively; **e** (9, 0) CNT with both contacts rotated and moved according to **c** in the *upper panel*; **f** (10, 0) CNT with both contacts rotated and moved according to **c** in the upper panel (color figure online)

parameters (e.g., magnetization or other fields which break the symmetry of the ideal state). Moreover, in absence of further information, the procedure for the transport calculation for molecular devices is preceded by determination of the optimal configuration at *equilibrium*: this configuration is then frozen and NEGF is applied using the correspondent electron Hamiltonian.

The assumption of the stable configuration in equilibrium conditions could be too stringent for several molecular systems. Indeed, a series of experiments have demonstrated, in a wide class of structures, that nonlinear transport behavior (e.g., NDR or switching) occurs in molecular devices also caused by atomic configuration changes provoked by a different (nonequilibrium) state of the electron system induced by the biasing [4, 6]. Anyhow, the correct theoretical framework where such effect could be studied is still under debate.

In the following, we present a theory for the understanding of anomalous electron dynamics based on a procedure which extends to biased systems the variational principle well established in the thermodynamic equilibrium. The method can be generally applied to any single-particle formulation (including mean field approximation); moreover, it is consistent with Laundauer's approach to the electronic transport (and its extension in the NEGF formulation) since it relies on the same general framework.

The Nonequilibrium Variational Theory

Variational theories have the merit to overcome the limits of pertubative methods reshuffling the zero-point state of the system. In general, they are based on the trial states, which are closer in terms of symmetry to the exact state and depend on additional variables/fields to be suitably optimized. As a consequence, in solid state applications, the procedure results often in the reduction of the original Hamiltonian to an effective pure electron one $H\{u_i\}$, where the many-body effects or the role of additional quantum variables (e.g., nuclear positions) are taken into account by means of a given set $\{u_i\}$ of variational variables. This procedure is rather simple in the adiabatic approximation for the nuclear motion where the set $\{u_i\}$ practically coincides with nuclear (quasi classical) coordinates, while the theory requires a more complex ansatz in other cases [4].

Here, we assume that the reduced Hamiltonian $H\{u_i\}$ is a single-particle approximation of the complete one. In the equilibrium conditions, i.e., assuming the system in contact with a single-particle's reservoir at temperature T and chemical potential μ, we can determine the optimal values of the variational variables minimizing the free energy functional [21]

$$\Omega_\mu\{u_i\} = -k_B T \int \ln\left[1 + \exp((\mu - \varepsilon)/k_B T)\right] n(\varepsilon, \{u_i\}) \, d\varepsilon + U_\Omega\{u_i\} + \mu N_0.$$

$$(10.30)$$

Here, N_0 is the total number of electrons ruled by μ, $n(\varepsilon, \{u_i\})$ is the density of states at energy ε, and $U_\Omega\{u_i\}$ an eventual term independent on the electron occupancy. As a consequence, the self-consistent estimates of $\{u_i\}$ can be obtained solving the electron problem within the constraints imposed by the extremum conditions $\delta\Omega_\mu\{u_i\}/\delta\{u_i\}$, which give the stable (and metastable) states of the system.

Parametric nonequilibrium theories have been proposed [6] and applied to the electron transport substituting the *nonequilibrium* electron occupancy in the formal expression for $\{u_i\}$ derived in *equilibrium* conditions. This procedure has not a general validity, since it is not based on a variational principle, moreover the variational equations are explicitly functions of the electron occupancy only in particular cases. An extension of the variational principle, allowing a direct numerical minimization procedure, has been introduced for the first time in Ref. [5] using a modification of the functional in Eq. (10.30), which is consistent with Landauer's scheme of the electron transport for the case of stationary electron currents (cf. Ref. [10]).

Landauer's theory, for a system which is in contact with two leads (i.e., with two independent particle reservoirs) at chemical potentials μ_L and μ_R, establishes that the device states are populated by electrons ($+k$ states) with distribution $n_L(\varepsilon, \{u_i\}) = \mathrm{Tr}(\mathscr{G}\Gamma_L\mathscr{G}^\dagger)/2\pi$ in equilibrium with the left contact and by electrons ($-k$ states) with distribution $n_R(\varepsilon, \{u_i\}) = \mathrm{Tr}(\mathscr{G}\Gamma_R\mathscr{G}^\dagger)/2\pi$ in equilibrium with the right contact.

According to the Gibbs prescription, in the equilibrium case, the chemical potential μ rules the charge fluctuation when the system is in contact with a single-particle reservoir. As a consequence, we may assume that μ_L and μ_R play the same role in the device configuration for $+k$ and $-k$ electron states, respectively. The conjecture reported in Ref. [5] is that the extended free energy functional for the composition of $+k$ ($-k$) electron states should have a Mermin-like expression [21] ruled by the two chemical potentials of the leads:

$$
\begin{aligned}
\Omega_{\mu_L,\mu_R}\{u_i\} = \; & U_\Omega\{u_i\} + \mu_L N_L + \mu_R N_R \\
& - k_B T \int \ln\left[1 + \exp((\mu_L - \varepsilon)/k_B T)\right] n_L(\varepsilon, \{u_i\})\, d\varepsilon \\
& - k_B T \int \ln\left[1 + \exp((\mu_R - \varepsilon)/k_B T)\right] n_R(\varepsilon, \{u_i\})\, d\varepsilon, \quad (10.31)
\end{aligned}
$$

where for symmetry considerations $\mu_L N_L = \mu_R N_R = N_0/2$. Moreover, taking into account the identity $\mathscr{G}(\Gamma_L + \Gamma_R)\mathscr{G}^\dagger = i(\mathscr{G} - \mathscr{G}^\dagger)$, the functional in Eq. (10.31) correctly reduces to Eq. (10.30) when $\mu_L = \mu_R = \mu$. After the optimization procedure with respect to the $\{u_i\}$ variables, the current can be calculated by means of the Landauer's expression as

$$
I = \frac{2e}{h} \int_{-\infty}^{\infty} T(\varepsilon, \{u_i\})(f_L - f_R)\, d\varepsilon \tag{10.32}
$$

for the coherent stationary case, where the transmission is formally similar to Eq. (10.22).

Electron Transport in Carbon Chains

As an application example of the method presented in the previous section, we discuss the behavior of a simple quantum resistor based on a pure carbon chain, i.e., a linear arrangement of N bonded carbon atoms coupled by means of the first ($n = 1$) and last ($n = N$) ones to two semi-infinite metallic electrodes. We can assume that a reliable model for this system is the well studied Su–Schrieffer–Heeger (SSH) [22] Hamiltonian $H[u_n]$, with the addition of the contacts and the device contact coupling:

$$H[u_n] + \sum_{k \in \{L,R\},\sigma} \varepsilon(k) c_{k,\sigma}^{\dagger} c_{k,\sigma}$$

$$+ \sum_{k \in \{L\},\sigma} V_k c_{k,\sigma}^{\dagger} c_{1,\ell,\sigma} + \text{H.c.}$$

$$+ \sum_{k \in \{R\},\sigma} V_k c_{k,\sigma}^{\dagger} c_{N,\ell,\sigma} + \text{H.c.}, \qquad (10.33)$$

where the SSH model reads

$$H[u_n] = -t_0 \sum_{n,\ell,\sigma} c_{n+1,\ell,\sigma}^{\dagger} c_{n,\ell,\sigma} + c_{n,\ell,\sigma}^{\dagger} c_{n+1,\ell,\sigma}$$

$$+ \alpha \sum_{n,\ell,\sigma} (u_{n+1} - u_n) \left[c_{n+1,\ell,\sigma}^{\dagger} c_{n,\ell,\sigma} + c_{n,\ell,\sigma}^{\dagger} c_{n+1,\ell,\sigma} \right]$$

$$+ \sum_n \left[\frac{p_n^2}{2M} + (u_{n+1} - u_n)^2 \right]. \qquad (10.34)$$

Here, $c_{n,\ell,\sigma}^{\dagger}$ ($c_{n,\ell,\sigma}$) is the creation (annihilation) operator of an electron with spin σ, u_n is the dimerization coordinate, p_n the conjugated momentum, $t_0 = 2.7\,\text{eV}$ is the hopping integral between carbons, and α the electron–phonon coupling energy. The momentum k runs over the band states of the left (L) and right (R) contacts, and V_k are the device-leads coupling parameters. A negligible mass parameter $1/M$ is assumed (adiabatic approximation).

In the case of an infinite chain (with no contacts), the equilibrium ground state ($T = 0$) energy as a functional of $[u_n]$ can be analytically evaluated, assuming $u_n = (-1)^{n-1} u_0$ and diagonalizing the resulting Hamiltonian in momentum space (assuming two atoms per Aunit cell) by means of a Bogoliubov transformation as

$$E(k) = \varepsilon \pm \sqrt{t^2 + t'^2 + 2tt' \cos(kL)}, \qquad (10.35)$$

where $t = t_0 + 2\alpha u_0$, $t' = t_0 - 2\alpha u_0$, and L is the lattice periodicity. As a consequence, the electronic structure of these systems has a gap whose value is related to the equilibrium distance between the neighboring atoms to the bond length alternation (BLA) value, BLA $= 4u_0$.
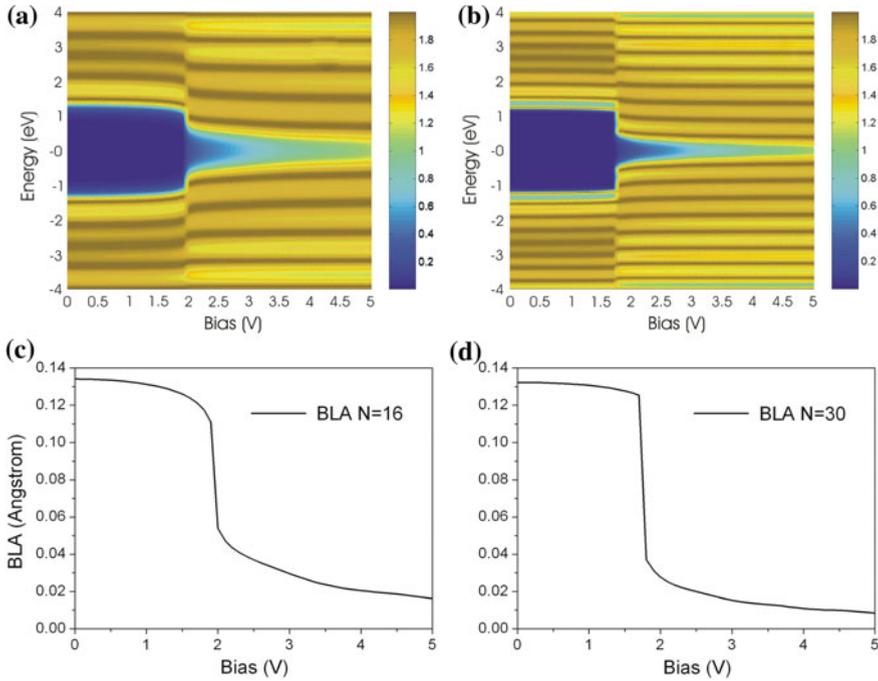
Fig. 10.3 Transmission (*color* scale in G_0 units) as a function of the applied voltage V and energy E for $N = 16$ (**a**) and $N = 30$ (**b**) carbon chains contacted with $< 111 >$ gold leads. **c**, **d** BLA as a function of V for the same systems

In Fig. 10.3a, b we show the transmission $T(E, V)$, as a function of the applied bias V and the energy E, calculated by means of the variational procedure based on the Eqs. (10.31) and (10.32). Two systems are considered: the $N = 16$ and $N = 30$ carbon chains contacted with $< 111 >$ gold leads. As expected, close to equilibrium ($V = 0$) the transmission is practically zero in the energy gap while it is characterized by a sequence of peaks reaching a value $\sim 2G_0$ when the energy value is larger (smaller) than the LUMO (HOMO) level. As we could expect, the equilibrium BLA related gap $\Delta(N = 16) = 2.75$ eV, $\Delta(N = 30) = 2.42$ eV does not change significantly for small values of the potential between the two electrodes. However, in the $V = 1.7 - 1.9$ V region, the gap decreases discontinuously and gradually disappears for larger biases and the system manifests a 'metallic-like' transmission value. The gap reduction is consistently related to a BLA reduction (i.e., to a polyyne-cumulene transition), and both BLA and the gap tend to zero for large V (Fig. 10.3c, d).

In Fig. 10.4 the I-V curves (solid lines), calculated with Eq. (10.32) after the optimization of $\{u_i\}$, for the $N = 16$ (Fig. 10.4a) and $N = 30$ (Fig. 10.4b) systems are shown and compared to those calculated with the equilibrium values of $\{u_i^{eq}\}$. The nonequilibrium variational procedure, for the determination of the atomic configurations $\{u_i\}$ of the chains, predicts a current–voltage dependence for this 1D

Fig. 10.4 Self-consistently calculated current voltage characteristic (*solid-line*) for carbon chains with $N = 16$ (**a**) and $N = 30$ (**b**) atoms contacted with $< 111 >$ gold leads. I-V characteristics calculated without self-consistency are also shown (*dashes*)
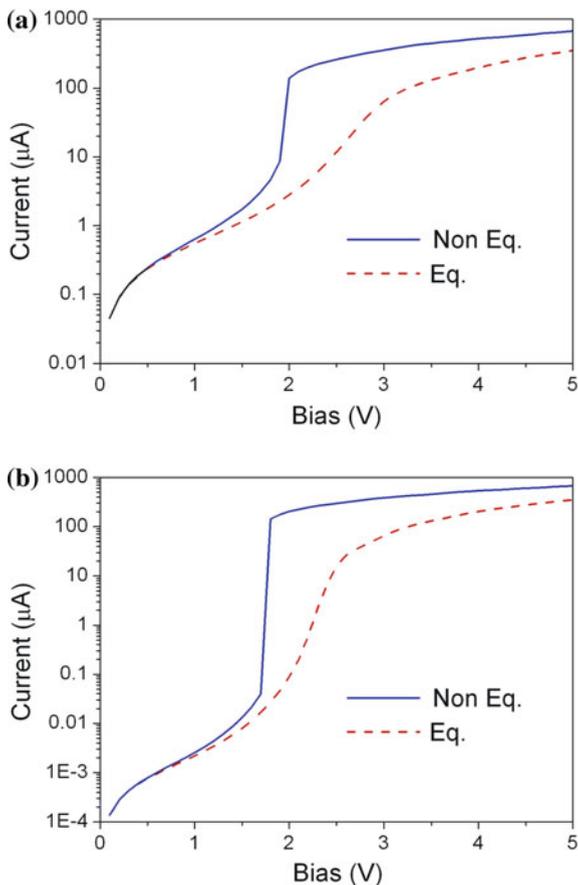
resistor characteristic of a insulator-metal transition; moreover, several diode-like I-Vs are calculated, when the equilibrium estimate $\{u_i^{eq}\}$ is considered (dashed lines in Fig. 10.4).

The change of atomic configuration of the C-chains, which is consistently related to nonequilibrium charging, could similarly emerge in other molecular devices, since the atomic configuration of the molecule is often strongly related to the electron distribution in the molecular orbitals. For the particular case of the carbon chains the biasing competes with the Peierls instability: the nonequilibrium electron distribution in the C orbitals makes the dimerized state progressively less favored until both BLA and the gap are strongly reduced.

Generalized Trial States in Nonequilibrium Variational Theories

In the previous section, a theory which extends the variational procedure in nonequilibrium stationary condition was discussed and applied to a standard case, where the variational variables coincide with the nuclear coordinates (i.e., assuming the validity of the adiabatic limit for the nuclear kinetics). However, this Born–Oppenheimer (BO) approximation, related to the smallness of the phonon energy scales ω with respect to all the other relevant electron energy scales, in general is not valid in molecular devices. Indeed, weak electron coupling Δ between lead states and molecular states could arise in the realization of a molecular device which makes questionable the validity of the BO approximation. Extremely weak coupling could be modeled applying the opposite anti-adiabatic small polaron (SP) approximation [23]. However, variational states valid irrespectively of the strength of the electron, phonon, and electron–phonon interactions can be formulated [23–25]. Here, we discuss the results obtained when this generalized variational approach is applied to the transport through molecular orbitals.

Prototype Model for the Electronic Transport in Molecular Orbitals: Equilibrium Solutions

In order to focus our problem to the essential element, we consider a minimal model of the molecular device where the Hamiltonian, describing the system, consists of a single spinless electron level coupled to a single vibrational mode and to two identical leads:

$$
\begin{aligned}
H = {} & \varepsilon_0 c_0^\dagger c_0 + \hbar\omega a^\dagger a + \chi \left(a^\dagger + a\right) c_0^\dagger c_0 \\
& + \sum_{k\in\{L,R\},\sigma} \varepsilon(k) c_k^\dagger c_k \\
& + \sum_{k\in\{L\}} \left(V_k c_k^\dagger c_0 + \text{H.c.}\right) \\
& + \sum_{k\in\{R\}} \left(V_k c_k^\dagger c_0 + \text{H.c.}\right),
\end{aligned}
\tag{10.36}
$$

where ε_0 is the electron energy of the molecular state, ω the vibrational frequency, χ the electron–phonon coupling energy, $\varepsilon(k)$ the energy of the electron states in the two leads L and R, and V_k the device-leads coupling. The BO approximation can be obtained by means of the translation transformation

$$T_1 = \exp\left[-\sqrt{\frac{\lambda}{\hbar\omega}}\left(a^\dagger - a\right)\tilde{x}_0\right],\tag{10.37}$$

which introduces the related classical nuclear coordinate variable $\tilde{x}_0$, where $\lambda = \chi^2/\hbar\omega$. The opposite anti-adiabatic limit and the quantum state of the vibration for the intermediate conditions $\Delta \sim \hbar\omega$, characterized by non-Gaussian fluctuations, can be considered by the transformation

$$T_2 = \exp\left[-\sqrt{\frac{\lambda}{\hbar\omega}}\vartheta\left(a^\dagger - a\right)\left(c_0^\dagger c_0 - \tilde{x}_0\right)\right],\tag{10.38}$$

and by the squeezed two-phonon state

$$|ph\rangle = \exp\left[-\alpha\left(aa - a^\dagger a^\dagger\right)\right]|0_{ph}\rangle,\tag{10.39}$$

respectively. Here, ϑ and α are two additional variational variables improving the description of the vibrational quantum state away form the adiabatic limit.

Following the variational procedure, the phonon operators are eliminated by means of Eqs. (10.37)–(10.39), and the resulting electron Hamiltonian is

$$\begin{aligned}
H = {} & \tilde{\varepsilon}_0 c_0^\dagger c_0 \\
& + \sum_{k\in\{L,R\}c,\sigma} \varepsilon(k)c_k^\dagger c_k \\
& + \sum_{k\in\{L\}} \left(\tilde{V}_k c_k^\dagger c_0 + \text{H.c.}\right) \\
& + \sum_{k\in\{R\}} \left(\tilde{V}_k c_k^\dagger c_0 + \text{H.c.}\right) \\
& + \lambda\tilde{x}_0^2\left(1 - \vartheta^2\right) + \hbar\omega\left(\tau^2 + \tau^{-2}\right),
\end{aligned}\tag{10.40}$$

where

$$\tilde{V}_k = V_k \times \exp\left[-\frac{\lambda}{2\hbar\omega}\tau^2\vartheta^2\right],\tag{10.41a}$$

$$\tilde{\varepsilon}_0 = \varepsilon_0 - \lambda + \lambda\left(1 - \vartheta\right)^2\left(1 - 2\tilde{x}_0\right),\tag{10.41b}$$

$$\tau = \exp\left(-2\alpha\right).\tag{10.41c}$$

As can be easily argued from the analysis of the various expressions in the effective electron model Eq. (10.40), the coupling with the vibrational mode produces a renormalization of the electron energy level $\tilde{\varepsilon}_0$ and of the coupling strength between the contact and the single level molecular device:

Fig. 10.5 Ground state
energy estimate as a function
of the variation parameters n,
y (*upper panel*) for a fixed
value of $\tau = 0.7814$. The
model parameters are
$W = 20$, $\Delta = 0.5$, $\lambda = 2$,
$\varepsilon_0 = 2$
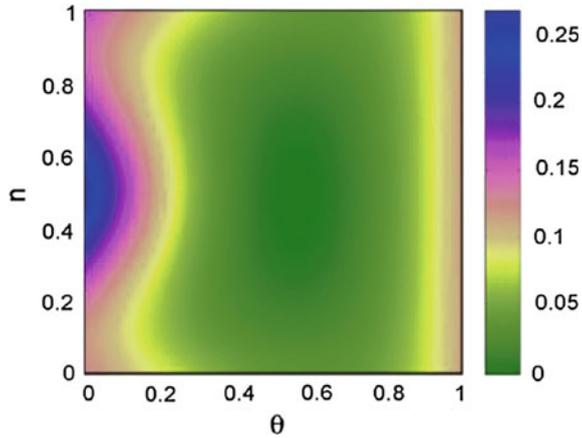
Fig. 10.6 Region of the
$\Delta - \lambda$ plane where a
bi-stable solution occurs for
$W = 100$
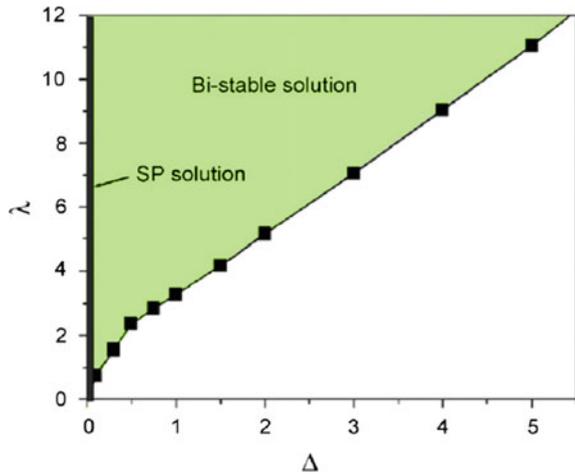
$$\tilde{\Delta}_E = \Delta_E \times \exp\left[-\frac{\lambda}{2\hbar\omega}\tau^2\vartheta^2\right]$$

$$= \left[2\pi \sum_{\substack{k\in\{L\}\\k\in\{R\}}} |V_k|^2 \delta_{\text{Dirac}}(E - \varepsilon(k))\right] \times \exp\left[-\frac{\lambda}{2\hbar\omega}\tau^2\vartheta^2\right]. \quad (10.42)$$

Assuming that the bandwidth W of the contacts is very large, we can neglect the
energy dependence of the coupling strength $\Delta_E \simeq \Delta$, and at $T = 0$ the model can
be solved analytically [4, 8].

In Fig. 10.5, we show the analysis of the equilibrium ($V_{bias} = \mu_R - \mu_L = 0$) solu-
tion for the following set of parameters: $W = 20$, $\Delta = 0.5$, $\lambda = 2$, $\varepsilon_0 = 2$. (Here we
use the vibration energy $\hbar\omega$ as unit for the energy). The global minimum for the

energy functional has been obtained for $\tilde{x}_0 = 0.5$ (for the units under usage, the adimensional displacement is equal to the average electron density in the level n), $\vartheta = 0.5666$, $\tau = 0.7841$. Contrary to the adiabatic solution ($\vartheta = 0$, $\tau = 1$), the optimal one does not show either the bi-stable character (as a function of the electron occupancy level) of the adiabatic solution [7], or the lack of dependence on electron filling typical of the anti-adiabatic solution ($\vartheta = 1$, $\tau = 1$). Moreover, bi-stable solutions, which could give rise to hysteresis in the I-V characteristics of this prototype molecular device, are globally stable in only a limited portion of the model parameters space (cf. Fig. 10.6). This fact has a crucial impact in the calculation of the electron transport properties of the system.

Prototype Model for the Electronic Transport in Molecular Orbitals: Nonequilibrium Solutions

Nonequilibrium solutions can be derived by the optimization of the proposed non-equilibrium Mermin-like functional. An example of the electron filling analysis in the single level of this prototype model as a function of the applied bias is shown in Fig. 10.7 for the following set of the model parameters: $W = 20$, $\Delta_L = 1.5$, $\Delta_R = 0.5$, $\lambda = 6.5$, $\varepsilon_0 = 6.4$, $T = 50$ K.

The optimal generalized solution has a bi-stable character at the thermodynamic equilibrium ($V_{bias} = 0$); however, the behavior of the electron density as a function of V_{bias} shows significant differences with respect to the solutions obtained within the BO approximation. Indeed, the hysteresis cycles occur in different potential ranges and, especially, have different amplitudes. We note that the asymmetric values of the electron filling n for negative and positive voltage reflect the contact asymmetry. Moreover, the inner root of the generalized solution shows two distinct particular trends in the positive and negative bias regions.
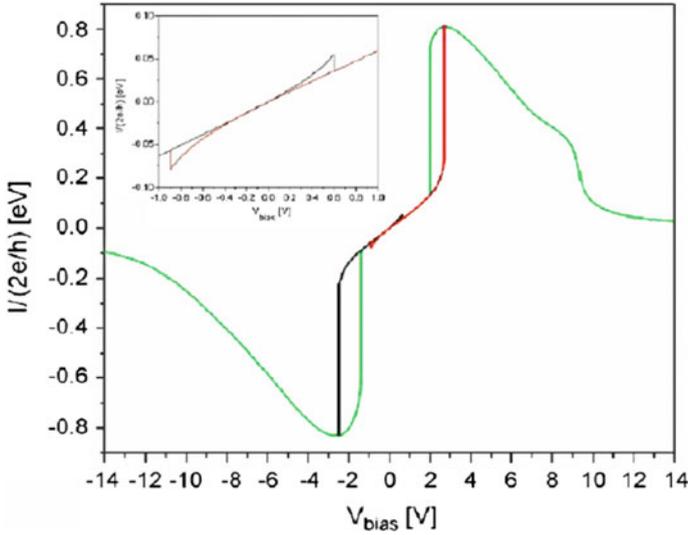
Fig. 10.8 Current–voltage curve of the different extremal solutions. The parameters set are $W = 20$, $\Delta_L = 1.5$, $\Delta_R = 0.5$, $\lambda = 6.5$, $\varepsilon_0 = 6.4$, $T = 50$ K. Outer roots are plotted in *black* and *red*, inner roots in green. In the *inset* we show a magnification of the low-bias region (color figure online)

The current–voltage curves are tightly related to the density dependence on V_{bias} and, in general, hysteresis cycles occur in the $I(V)$ curves corresponding to hysteresis cycles in the $n(V)$ curves. In Fig. 10.8, the $I(V)$ estimated using the optimal solutions for the same set of parameters as in Fig. 10.7 is shown and the correspondence between $I(V)$ and $n(V)$ can be inferred.

NDR is predicted by the optimal solution for the larger value of the voltage outside the bi-stability region and it is related to the behavior of the inner stable solution. The nature of the NDR has been discussed in detail in Ref. [4], and it is due to the decrease with V_{bias} of the effective coupling parameters $\tilde{\Delta}_L$ and $\tilde{\Delta}_R$ (i.e., the increase of the expression which depends on the variational parameters in the exponent). In addition, due to contact asymmetry the $I(V)$ trend in the potential regions where the NDR occurs is strongly asymmetric for positive and negative values of the bias. We note that the NDR related to potential-dependent polaronic-type corrections [23, 25] of the device-leads coupling parameters $\tilde{\Delta}_L$ and $\tilde{\Delta}_R$ cannot be evidenced in the adiabatic solutions, since in this case the coupling parameters depend neither on the variational parameters nor on the electron–phonon interactions.

Conclusions

The variational approach is an important methodology in theoretical physics which complements the pertubative methods in the case of the lack of small parameters ruling the interaction part of the model Hamiltonian. The method is well established in the thermodynamic equilibrium when generic energy functionals can be properly defined and utilized for the minimization procedure. Here, we have discussed the question of the lack of energy functionals in stationary nonequilibrium conditions when, for the particular case of the molecular device, intuitively a self-consistently relationship between atomic coordinate and electronic transport should occur [26, 27]. An extension of the Mermin's functional compatible with Landauer's scheme for the quantum coherent transport is proposed as a basis of a variational procedure in nonequilibrium stationary conditions. This variational method can be extended also to the phonon variables. It can be easily applied in any calculation based on the BO approximation (e.g., DFT) but also for any other generic variational set of variables [27]. Considering the results discussed above, the estimate of transport characteristics could be crucially modified by including these nonequilibrium corrections to the vatiational estimates. Moreover, since the method is not limited by the typology of the variational variables, in future research works it could be extended to the study of instabilities of the electron system caused by electron–electron interactions.

References

1. B. Xu, Y. Dubi, J. Phys.: Condens. Matter **27**(26), 263202 (2015). DOI 10.1088/0953-8984/27/26/263202
2. F. Demir, G. Kirczenow, J. Chem. Phys. **136**(1), 014703 (2012). DOI 10.1063/1.3671455
3. C.L. Kane, E.J. Mele, Phys. Rev. Lett. **95**, 226801 (2005). DOI 10.1103/PhysRevLett.95.226801
4. A. La Magna, I. Deretzis, Phys. Rev. Lett. **99**, 136404 (2007). DOI 10.1103/PhysRevLett.99.136404
5. A. La Magna, I. Deretzis, V. Privitera, Eur. Phys. J. B **70**(3), 311 (2009). DOI 10.1140/epjb/e2009-00220-6
6. M. Galperin, M.A. Ratner, A. Nitzan, A. Troisi, Science **319**(5866), 1056 (2008). DOI 10.1126/science.1146556
7. M. Galperin, M.A. Ratner, A. Nitzan, Nano Lett. **5**(1), 125 (2005). DOI 10.1021/nl048216c
8. A. La Magna, I. Deretzis, Physica E **40**(7), 2289 (2008). DOI 10.1016/j.physe.2007.09.088. Proceedings of the E-MRS 2007 Symposia L and M: Electron Transport in Low-Dimensional Carbon Structures and Science and Technology of Nanotubes and Nanowires
9. I. Deretzis, A. La Magna, Eur. Phys. J. B **81**(1), 15 (2011). DOI 10.1140/epjb/e2011-20134-x
10. S. Datta, *Electronic transport in mesoscopic systems* (Cambridge University Press, Cambridge, 1997)
11. A. La Magna, I. Deretzis, G. Forte, R. Pucci, Phys. Rev. B **80**, 195413 (2009). DOI 10.1103/PhysRevB.80.195413
12. I. Deretzis, G. Fiori, G. Iannaccone, A. La Magna, Phys. Rev. B **81**, 085427 (2010). DOI 10.1103/PhysRevB.81.085427

13. I. Deretzis, G. Fiori, G. Iannaccone, A. La Magna, Phys. Rev. B **82**, 161413 (2010). DOI 10.1103/PhysRevB.82.161413

14. F. Giannazzo, I. Deretzis, A. La Magna, F. Roccaforte, R. Yakimova, Phys. Rev. B **86**, 235422 (2012). DOI 10.1103/PhysRevB.86.235422

15. S. Datta, Superlattices and Microstructures **28**(4), 253 (2000). DOI 10.1006/spmi.2000.0920

16. S. Datta, *Quantum transport: atom to transistor* (Cambridge University Press, Cambridge, 2005)

17. I. Deretzis, A. La Magna, J. Chem. Phys. **128**(16), 164706 (2008). DOI 10.1063/1.2905216

18. Y.M. Lin, V. Perebeinos, Z. Chen, P. Avouris, Phys. Rev. B **78**, 161409 (2008). DOI 10.1103/PhysRevB.78.161409

19. A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, Rev. Mod. Phys. **81**(1), 109 (2009). DOI 10.1103/RevModPhys.81.109

20. R.Y. Oeiras, F.M. Araújo-Moreira, E.Z. da Silva, Phys. Rev. B **80**, 073405 (2009). DOI 10.1103/PhysRevB.80.073405

21. N.D. Mermin, Phys. Rev. **137**, A1441 (1965). DOI 10.1103/PhysRev.137.A1441

22. A.J. Heeger, S. Kivelson, J.R. Schrieffer, W.P. Su, Rev. Mod. Phys. **60**, 781 (1988). DOI 10.1103/RevModPhys.60.781

23. A. La Magna, R. Pucci, Phys. Rev. B **53**, 8449 (1996). DOI 10.1103/PhysRevB.53.8449

24. A. La Magna, R. Pucci, Phys. Rev. B **55**, 6296 (1997). DOI 10.1103/PhysRevB.55.6296

25. A. La Magna, R. Pucci, Phys. Rev. B **55**, 14886 (1997). DOI 10.1103/PhysRevB.55.14886

26. G. Vignale, M. Di Ventra, Phys. Rev. B **79**, 014201 (2009). DOI 10.1103/PhysRevB.79.014201

27. M. Brandbyge, J.L. Mozos, P. Ordejón, J. Taylor, K. Stokbro, Phys. Rev. B **65**, 165401 (2002). DOI 10.1103/PhysRevB.65.165401

Chapter 11
Peierls and Spin-Density Instability: From Polyacetylene to Graphene

M. Baldo

Abstract The sp^2 hydrocarbon conjugated molecules have remarkable properties mainly due to their delocalized π-electron system. They are at the basis of novel materials that have been produced and studied in the last few decades and which have found an extensive number of applications and are expected to be more and more involved in the development of nanostructures. The properties of these molecules are determined not only by external factors, like doping, but also by their intrinsic structure that can be affected by the presence of the Peierls and spin-density instabilities. The paper presents a short overview on the effects of these types of structural rearrangements on the molecular properties and how they develop from strictly one-dimensional to two-dimensional cases. Specifically, we will consider the one-dimensional polyacetylene polymer, the two-dimensional graphene molecule and, as a case intermediate between one and two dimensions, the quasi-one dimensional polyacenes molecules.

Introduction

Low-dimensional sp^2-based nanocarbons exhibit many peculiar properties and for this reason they have been at the basis of an impressive development in materials science. At the same time they are characterized by novel physical phenomena, that are specific of this class of materials, but of great interest for physics as a whole [1].

Besides the peculiarities of the carbon atom and its hybridized bond orbitals, which are at the origin of the mechanical properties of graphene and derived materials, the main feature of the electronic structure of these molecules is the presence of an extended system of delocalized π-electrons. In normal conditions the π-orbitals are half-filled and this determines the electrical properties of the material. The essential parameter is the value of the gap between the occupied and non-occupied orbitals, that eventually can collapse to zero. The gap for an undoped molecule is affected by

M. Baldo (✉)
Istituto Nazionale di Fisica Nucleare, Sezione di Catania,
Via S. Sofia 64, 95123 Catania, Italy
e-mail: ab_baldo@ct.infn.it

two types of structural instability, the Peierls distortion and the spin-density wave formation. Both of them introduce a gap between occupied and unoccupied orbitals, even if the undistorted molecule has zero gap. It is therefore essential, from the theoretical point of view, to establish to what extent these types of instability are present. It can be expected that their appearance depends on the dimensionality of the molecules, i.e., if it is strictly one-dimensional, quasi one-dimensional, or two-dimensional. Strictly, three-dimensional molecules of this type cannot exist, since the π-electron system would be absent, like in diamond.

In this paper, we will consider the development of Peierls and spin-density instabilities as we proceed from one-dimensional molecules to two-dimensional ones. Throughout the paper only elementary methods will be used, which however shed light on the main mechanisms that occur in fixing the electron structure of the π-electron system, at least qualitatively.

We will start in section "Polyacetylene" with the one-dimensional polyacetylene molecule, that has been studied extensively in the past since it is characterized by peculiar phenomena. In section "Polyacenes" we will go to the quasi one-dimensional polyacene molecule, where carbon rings are fused one after the other to form a linear chain. In the polymer limit the molecule can be viewed as two polyacetylene chains bonded together. We will consider also the corresponding finite size molecules, which will be studied as function of their length. In section "Graphene" we will consider the graphene carbon system. Associated with this system we will consider also finite molecule of different compactness, to see its influence on the electronic structure. Summary and conclusions will be presented in section "Summary".

Polyacetylene

The π-electron system in the polymer polyacetylene can be considered one-dimensional, at least in the sense that the electrons can hop from one carbon p-orbital to the other as along a linear chain. In a simple one-body picture, if all the sp^2 bonds are equal, the corresponding molecular orbits (MO) are half-filled because there is one single electron for each carbon sites. This is the undistorted model for polyacetylene. The molecular orbits can be readily obtained in the Hückel (tight binding) approximation, where hopping can occur only between next-neighbor atomic sites. The hopping hamiltonian V is therefore such that

$$\langle i|V|j\rangle = \beta \left(\delta_{j,i+1} + \delta_{j,i-1} \right) \tag{11.1}$$

where the indexes i, j number the carbon p atomic orbitals along the chain and β is the hopping strength parameter (the well known resonant integral). The molecular orbits are linear combination of atomic orbits (LCAO) and, taking into account the periodicity of the system, they can be written as

Fig. 11.1 Energies of the occupied (*lower curves*) and unoccupied (*upper curves*) molecular orbitals in polyacetylene. The *full lines* correspond to the undistorted molecule, the *dashed lines* to the Peierls distorted one
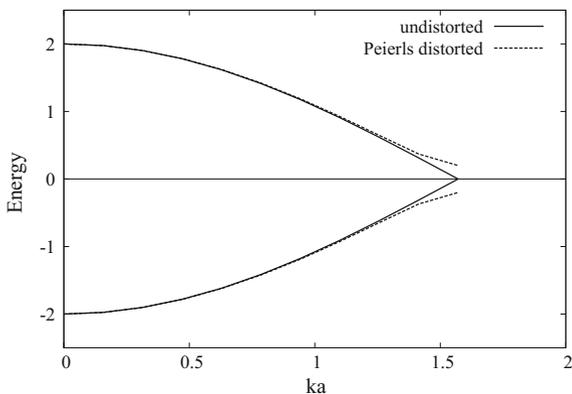
$$|k\rangle = \sum_j e^{ikaj} |j\rangle \tag{11.2}$$

where

$$k = \frac{n\pi}{Na}, \quad n = \pm 1, \pm 2, \ldots, \tag{11.3}$$

where N is the total number of carbon sites and we have introduced for convenience the (fictitious) step size a, which actually disappears in the exponential of Eq. (11.2), but it gives a form to the MO that resembles the solid-state notation. The hopping interaction V is diagonal in the MO of Eq. (11.2), and the corresponding energy E_k is given by

$$E_k = \varepsilon - 2\beta \cos(ka), \tag{11.4}$$

where ε is the energy of the π-electron atomic orbit. The MO spectrum (solid line), reported in the first Brillouin zone, is illustrated in Fig. 11.1, where the zero is at the atomic energy ε, and energies are in unit of β. The energy of the occupied MO, lower curve, matches with the one of the unoccupied MO (upper curve) at the Fermi quasi-momentum $ka = \pi/2$. Since the density of state is nonzero at the Fermi energy, the molecule would be assimilated to a metallic character.

However, it is well known that polyacetylene has a dimerized structure, i.e., an alternation of longer and shorter bond, corresponding to the alternation of the values of the resonant integral. This is depicted in Fig. 11.2. As noticed in Ref. [2, 3] the ground state must be in the dimerized form. To see this, let us calculate the MO in the case that the resonant integral alternates between $\beta - \delta$ and $\beta + \delta$, with $\delta > 0$ a parameter to be determined. In this case the unit cell contains two atoms, or equivalently the polymer can be seen as two sublattices that are shifted with respect to each other. One sublattice is formed by the carbon sites for which the stronger hopping parameter is on the left, the other sublattice by the carbon sites where the stronger hopping is on the right. To be definite let us enumerate the first sublattice sites by even numbers $2j$, and the ones of the second sublattice by the
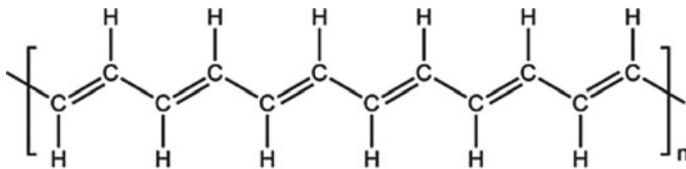
Fig. 11.2 Scheme of the Peierls distorted polyacetylene molecule. A segment of n sites is indicated

odd numbers $2j + 1$, with $j = 0, \pm 1, \pm 2, \ldots \pm N/4$. The matrix elements of the hopping interaction are then

$$\langle 2j|V|2j' + 1\rangle = (\beta - \delta)\delta_{j,j'} + (\beta + \delta)\delta_{j',j-1} \tag{11.5}$$

and the complex conjugate thereof.

The quasi-momentum k is still diagonal, but the MO contain two amplitudes, c_1, c_2, for the even and the odd sites, respectively, which are determined by a two-dimensional linear system

$$- E_k e^{ika(2j)}c_1 + \left[(\beta + \delta)e^{ika(2j-1)} + (\beta - \delta)e^{ika(2j+1)}\right]c_2 = 0, \quad (11.6a)$$
$$- E_k e^{ika(2j+1)}c_2 + \left[(\beta - \delta)e^{ika(2j)} + (\beta + \delta)e^{ika(2j+2)}\right]c_1 = 0. \quad (11.6b)$$

From the secular equation for the energy E_k one gets

$$E_k = \pm \left[2(\beta^2 + \delta^2) + 2(\beta^2 - \delta^2)\cos(2ka)\right]^{1/2}. \tag{11.7}$$

The spectrum, modified by the Peierls distorsion, is depicted in Fig. 11.1 (dashed line), for $\delta/\beta = 0.2$. A gap has opened at the Fermi energy, of width 2δ. Accordingly the material can be associated with a semiconductor or insulator character, depending on the value of δ. The main point is that, as noticed in Ref. [2, 3], the derivative with respect to δ^2 at $\delta = 0$ is infinitely negative and therefore it dominates the minimization toward a nonzero value of δ. This can be seen by calculating the ground state energy, integrating over the occupied MO. After some algebra, one gets for the energy per site E

$$E = -\frac{4}{\pi} \int_0^{\pi/2} dy \left[\beta^2 - (\beta^2 - \delta^2)\sin^2 y\right]^{\frac{1}{2}}. \tag{11.8}$$

It is readily verified that the derivative of the integral with respect to δ^2 is diverging at $\delta = 0$. This means that the Peierls distorsion and the consequent opening of the gap lower the ground state energy, at least in this simple Hückel approximation. The Peierls distortion is confirmed by extensive *ab initio* calculations of the polyacetylene structure [4], where one can determine the actual values of the bond lengths, which turn out to be in close agreement with the experimental data. The semiconductor character of polyacetylene is confirmed by the measurements on the electric

Fig. 11.3 Conductivity of polyacetylene material as a function of doping. In the *inset*, the trans- and cis- form of the polyacetylene molecule is reported for illustration (Redrawn after Ref. [5].)
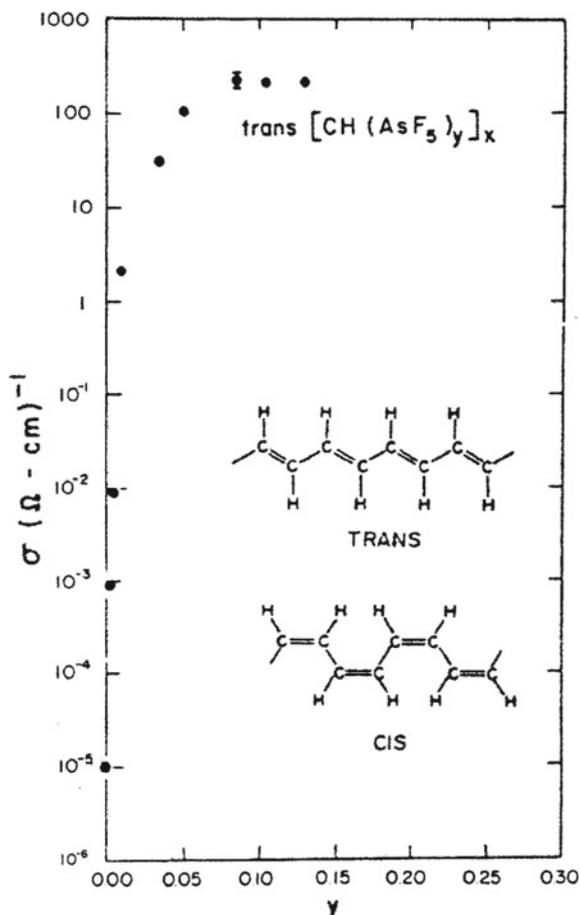
conductivity, as shown in Fig. 11.3 [5]. At vanishing doping, the residual conductivity can be attributed to the thermally excited electrons.

It has to be noticed that the presence of the Peierls distorsion induces a density ocillation along the chain. Indeed in the bond of shorter length (not to be confused with the fictitious step size a introduced above) the two neighboring p orbitals have a larger overlap, which produces a pile up of electron density. In the elementary theory we have used this effect cannot be described.

Polyacenes

The next class of molecules is the series of polyacenes, which are obtained by fusing a certain number of benzene rings along a linear chain. Let us first consider the polymer limit of the series, even if it has not been synthesized. The dimerized picture of the polymer is reported in Fig. 11.4. It can be viewed also as two polyacetylene polymers bonded together. The number of π-electron is still equal to the number of sites. An additional resonance integral β_t fixes the hopping amplitude between the two polyacetylene chains. In order to analyze the possible Peierls distorsion we calculate the total energy in the Hückel approximation as a function of δ. To this purpose we need to find the MO and their energies. The unit cell contains four sites, and the corresponding coefficients will be indicated as $c^{(l,a)}$, with $l = 1, 2$ and $a = u, d$ (for 'upper' and 'down'), see Fig. 11.4. With the same convention as for polyacetylene, the nonzero hopping matrix elements are

$$\langle 2j, a|V|2j'+1, a\rangle = (\beta + \delta)\delta_{j',j+1} + (\beta - \delta)\delta_{j',j-1} \qquad (11.9a)$$

$$\langle 2j'+1, u|V|2j''+1, d\rangle = \langle 2j'+1, d|V|2j''+1, u\rangle = \beta_t \delta_{j,j'} \quad (11.9b)$$

The four coefficients $c^{l,a}$ are now determined by a four dimensional linear system, where the two equations of the polyacetylene case (one for each chain) are coupled together. The four-dimensional matrix of the system can be written

$$\hat{M} = \begin{pmatrix} D & \hat{\beta}_t \\ \hat{\beta}_t & D \end{pmatrix},$$

where the two-dimensional submatrix D coincides with the one for polyacetylene and

$$\hat{\beta}_t = \begin{pmatrix} 0 & \beta_t \\ 0 & 0 \end{pmatrix},$$

with a suitable choice of the indexes ordering. The secular equation is then
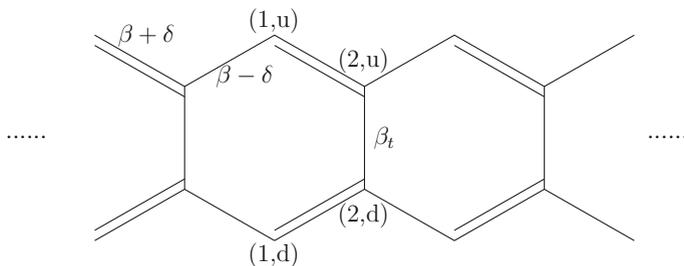
Fig. 11.4 Schematic representation of the polyacene polymer. Indicated are the three basic resonance integrals for the three distinct types of bond. The labels for each basic site are also indicated

$$\parallel D \parallel^2 - \beta_t^2 E_k^2 = 0, \tag{11.10}$$

which gives for the energies E_k

$$E_k = \pm \frac{1}{2}\beta_t \pm \frac{1}{2}\sqrt{\beta_t^2 + 16\beta^2 - 16(\beta^2 - \delta^2)\sin^2(ka)}. \tag{11.11}$$

In total we have four bands, two occupied $(\pm, -)$ and two unoccupied $(\pm, +)$. They reduce to the two of polyacetylene, Eq. (11.7), for $\beta_t \rightarrow 0$. The corresponding total energy per site is

$$E = -\frac{4}{\pi}\int_0^{\pi/2} dy\sqrt{(\beta_t/4)^2 + \beta^2 - (\beta^2 - \delta^2)\sin y^2}. \tag{11.12}$$

For $\beta_t \neq 0$ the derivative of E with respect to δ^2 at $\delta = 0$ is limited. This brings a drastic difference with respect to the one-dimensional polyacetylene chain. It is likely that this result will hold true for any number of connected chain, i.e., even for graphite ribbons or fully two-dimensional graphene. Of course the derivative is still negative, and therefore dimerization tends to lower the energy, but it can be contrasted by other electron correlations that stabilize the molecule. Only a more refined treatment that includes electron–electron correlations and ion–ion repulsion can give a definite answer.

Let us assume for the moment that the polyacene molecule is undimerized and look for other types of structural instabilities of the simple mean field approximation. In Ref. [6] the excitation spectrum of the polyacenes series was studied in the Random Phase Approximation (RPA). It was found that the energy of the triplet excitation in the B_{2u} symmetry decreases at increasing number of rings and collapses for naphtacene. As it is well known, the instability of the RPA solution signals the instability of the mean field and the appearance of a structural change. In this case, since the instability is in the triplet channel, it corresponds to the formation of spin-density oscillations along the molecule. The calculations thus predict the formation of spin-density "waves" for a number of rings equal or larger than 4. The RPA calculations were performed on the basis of a semi-empirical approximation for the repulsion integrals, and therefore they have to be taken at a qualitative level. The indication of the formation of a correlated ground state was confirmed in Ref. [7] within the alternant molecular orbital theory (AMO) [8], which corresponds to the unrestricted Hartree–Fock approximation where different orbitals are allowed for spin up and spin down. The authors report the ratio between the chemical potential and the energy per electron as a function of the number of rings in the Hückel approximation and in the AMO scheme. The two predictions essentially coincide for a number of rings n less or equal to 6, but they differ for $n > 6$, where the correlations play a relevant role. However in the AMO method, at least in its standard formulation, the ground state has not a definite total spin, and therefore is not suitable to fix the possible onset of a triplet ground state. The most extensive microscopic theoretical study on the onset of spin-density wave along the polyacenes series is probably the one of Ref. [9].
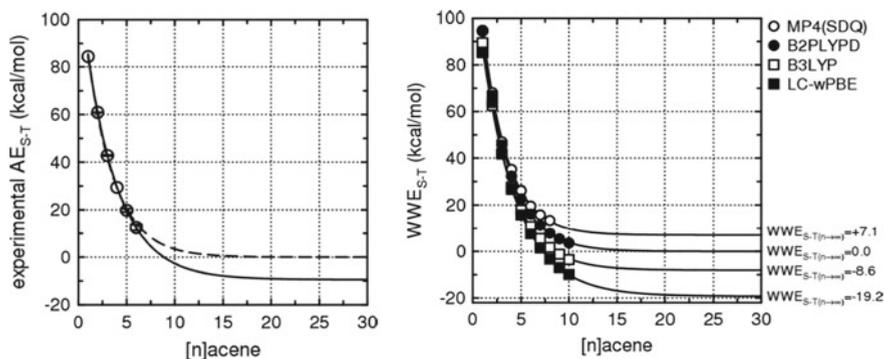
Fig. 11.5 Experimental triplet-singlet gap for the polyacene series. On the *left panel* are reported the well-established experimental values, together with two possible extrapolations to the polymer limit. On the *right panel* are reported the corresponding calculated values up to decacene for a representative set of energy density functionals among the ones of a much wider set. Both panels are redrawn from Ref. [9]

The author presents a set of calculations of the gap between the lowest singlet and triplet states based on a very wide choice of energy density functionals (EDF). They found that for most of the EDF the gap reduces as the number of rings increases and it changes sign for the nonacene molecule, where the triplet state become lower than the singlet one. In Fig. 11.5, left panel, is reported the experimental position of the lowest triplet state with respect to the singlet ground state. It is systematically decreasing as the number of rings increases. As shown in the figure, the possible extrapolations (full and dashed lines) to the polymer limit can give either a vanishing small gap or a negative one. On the right panel of the same figure are reported the results for the same quantity obtained with some of the considered EDF up to decacene and again extrapolated to the polymer limit. They are representative of the over 40 EDF considered in the analysis. Unfortunately, up to now it was not possible to synthesize in a stable way polyacenes higher than hexacene. It has to be noticed that the authors find that the Hartree–Fock solutions give the transition to a negative gap for pentacene, in line with the results of Ref. [6].

The instability toward a triplet ground state can be understood in simple terms by considering the triplet excitation as the number of rings increases. To this purpose let us first consider the MO and the corresponding energies of the polymer. The energies of the π-electron MO in the Hückel approximation are given by Eq. (11.11) (for $\delta = 0$) and are reported in Fig. 11.6, in units of the resonant integral β, and $\beta_t = 0.8\beta$ is assumed for definiteness. Because of the half-filling of the MO, the higher occupied MO and the lower unoccupied MO touches at $ka = \pi/2$, and they are both tangent to the x-axis. Each MO has a definite point group symmetry. If we take as point group D_{2h}, in the Schoenflies notation [10], each MO energy band is twofold degenerate, corresponding to two of the four irreducible representation A_1, A_2, B_1, and B_2. The assignments are summarized in Table 11.1.
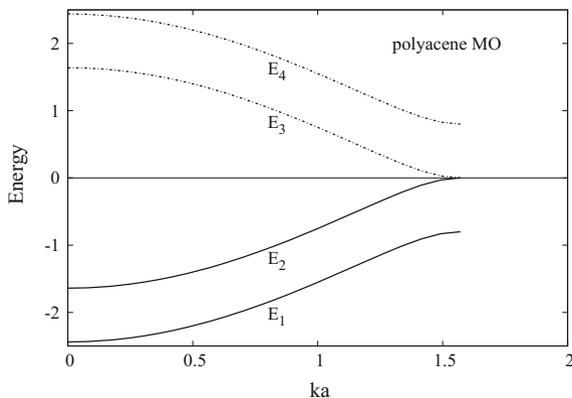
Fig. 11.6 Energies of the occupied (*solid lower lines*) and unoccupied (*dash-dotted upper lines*) molecular orbitals in polyacene. The zero of the energy is taken at the Fermi energy. Energies in unit of the resonant integral β, and $\beta_t = 0.8\beta$

Table 11.1 Irreducible representations of the MO included in each energy band of polyacene

E_1	E_2	E_3	E_4
A_1, B_2	A_2, B_1	A_1, B_2	A_2, B_1

The degeneracy between E_2 and E_3 at $ka = \pi/2$ is removed if an attractive interaction is present, however small it can be. The situation is similar to the one in superconducting metals, where any attractive interaction at the Fermi momentum will produce an electron–electron bound state [11]. In this case the system is mainly one-dimensional, i.e., the "momentum" k has only one direction along the chain. Notice anyhow that the density of single-particle states is in principle diverging at $ka = \pi/2$. Furthermore the possible bound state is here between a particle and a hole, rather than between two particles. Breaking such a bound state requires some energy, and this produces a gap between occupied and unoccupied single-particle states. That correlations can produce a gap is apparent in the AMO method, where it has been shown [12] that the unrestricted HF orbitals have an energy

$$e_i = \pm\sqrt{E_i^2 + \Delta^2},\tag{11.13}$$

which indeed resembles the one for a superconductor. The parameter Δ is related to the electron interaction (in the Hubbard approximation), and the gap is 2Δ. However, this result does not fix the spin ground state, which remains undefined, and it is unclear if the ground state has indeed a spin-density wave structure.

To see the mechanism that can create a particle-hole bound state in the triplet channel, let us consider the scattering matrix of a particle and a hole. The MO can be characterized by the quasi-momentum k and the band index i. The particle-hole pair

can be labeled by its total momentum P, which is conserved, and relative momentum q and their band indexes. Neglecting for the moment the band indexes, the scattering matrix T satisfies the integral equation

$$T(q, q'; P) = v(q, q'; P) + a \int dq'' v(q, q''; P) \frac{1}{\omega - E(q'', P)} T(q'', q'; P),$$
(11.14)

where ω is the entry energy, $v(q, q''; P)$ is the particle-hole interaction and $E(q'', P)$ the particle-hole energy

$$E(q'', P) = E_p(P/2 + q'') - E_h(P/2 - q'').$$
(11.15)

Here p indicates an unoccupied orbital and h an occupied one. The integration is properly limited in such a way that all momenta are inside the Brillouin zone. We can take the particle orbital as E_3 and hole one as E_2, see Fig. 11.6, since it is for these orbitals that the interaction is critical. Expanding around $ka = \pm\pi/2$, and assuming $Pa \ll \pi/2$, one gets

$$E(q'', P) \approx 16\beta^2/\beta_t(q''a - \pi/2)^2.$$
(11.16)

If the interaction v is attractive around the Fermi momenta $\pm\pi/2a$, then the homogeneous part of Eq. (11.14) has a solution for a definite negative value of the energy ω, which signals the presence of a bound state in the correlated ground state. To see this, let us take a constant interaction around the Fermi momentum. This is qualitatively correct because the interaction is expected to be a smooth function of the momenta and the density of state has a sharp peak at the Fermi momenta $\pm\pi/2a$. Then the scattering matrix is also independent of momenta, and putting $v(q, q'') \approx -|v_0|$, one gets the implicit equation for the energy ω_B of the bound state

$$1 = -|v_0| \frac{\sqrt{\beta_t}}{8\beta} \int_0^{\varepsilon_0} \frac{d\varepsilon}{\sqrt{\varepsilon}} \frac{1}{\omega_B - \varepsilon},$$
(11.17)

where ε_0 is a suitable cutoff. For values of ω approaching zero from the negative side, the integral will to diverge to $-\infty$, and therefore there is necessarily a solution for a finite negative value of ω_B. This way of proceeding is closely parallel to the one in Ref. [11] for a superconductor. The main difference is in the sharp peaking of the density of state at the Fermi momenta, due to the quasi-one-dimensional character of the molecule. Furthermore, the searched bound state is between a particle and a hole and not between two particles. Despite that, the appearance of a gap in the spectrum has a similar origin.

It remains to demonstrate that the interaction is indeed attractive in the triplet channel. The first thing to be noticed is that the particle-hole interaction for the triplet channel v^T contains only the exchange term

$$(k_1 i, k_2 j | v^T | k_1' l, k_2' m) = -(k_1 i, k_2 j | v | k_2' m, k_1' l),$$
(11.18)

where v is the (effective) electron-electron interaction, which is mainly determined by the repulsion integrals. The exchange interaction is usually negative (attractive) and this is indeed the case in some particle-hole channels involving the highest occupied MO (E_2) and the first non-occupied MO (E_3). It can be verified that, in the zero overlapping approximation, the matrix elements of Eq. (11.18) for v^T are indeed attractive near the Fermi momenta, provided the particle-hole configurations are in the A_2 irreducible representation. The symmetry can be obtained combining A_1 particle symmetry and A_2 hole symmetry, or B_2 particle symmetry and B_1 hole symmetry. Equation (11.14) is therefore more correctly a 2×2 system, but the considerations developed above can be easily extended to such a situation.

It can be concluded that spin-density wave should appear for a long enough polyacenes molecules. Which one is the critical number of rings can be decided only by microscopic many-body calculations, as described above.

Graphene

The two-dimensional version of a polycyclic carbon molecule is the graphene, which was produced and isolated for the first time only in 2004 [13]. It is not possible to make justice in this short note to the countless works that have been devoted to the study of graphene and to all the spectacular applications based on graphene and its variants. Extensive reports on all that can be found in the excellent review articles existing in the literature [1, 14]. We limit here to a discussion on the *absence* of distorsions and charge fluctuations, which is one of the reason for its peculiar properties.

Graphene can be viewed as a single sheet of graphite. Figure 11.7, redrawn from Ref. [1], shows a portion of the honeycomb structure of graphene, together with the unit cell in coordinates and in the momentum reciprocal lattice. For no bond distorsions the band structure was first calculated by Wallace already in 1947 [15]. The no distorsion assumption is supported by the considerations developed in the previous section "Polyacenes", and we take it as established.

We sketch here the calculation of the standard results for the graphene MO. We first observe that the honeycomb structure can be divided into two equivalent sublattices. This is best illustrated if one performs a 30 degrees rotation with respect to the Fig. 11.7. The result can be depicted as in Fig. 11.8, where the two equivalent sublattice sites are marked by full and open circles, respectively. One notices that each site of one sublattice has three bonded neighboring sites of the other sublattice (and no other neighbor). The picture shows also the analogy with the polyacene polymer, see Fig. 11.4. The symmetry here is even higher, since the three bonds are equivalent.

One can then follow the same tight binding procedure and find the MO. In Fig. 11.7 the basis vectors of the Bravais lattice is shown. Taking in this configuration the horizontal and vertical axes as x- and y-axis, respectively, we report here for definiteness the standard result for the MO energies
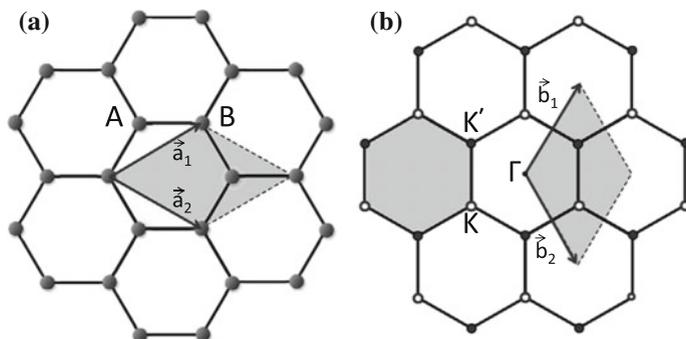
Fig. 11.7 Graphene unit cell in configuration space (*panel a*) and in the momentum space of the reciprocal lattice (*panel b*). The unit vectors of the Bravais lattice are also shown. Redrawn with permission from Ref. [1]
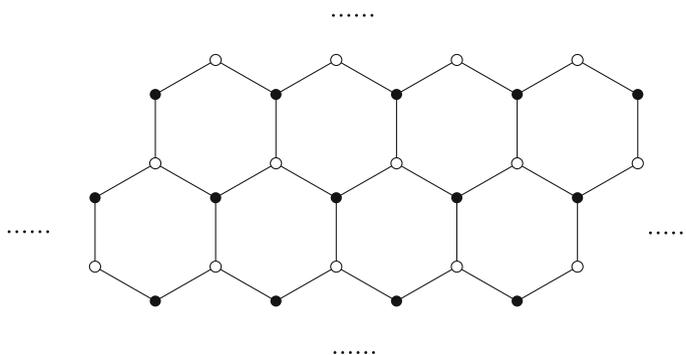
Fig. 11.8 The honeycomb structure of graphene with the sites of the two sublattices indicated by full and open *circles*, respectively

$$E_k = \pm\beta \left[1 + 4\cos\left(\frac{3}{2}k_x a\right)\cos(\frac{\sqrt{3}}{2}k_y a) + 4\cos^2\left(\frac{\sqrt{3}}{2}k_y a\right) \right]^{\frac{1}{2}}, \quad (11.19)$$

where the minus sign corresponds to the occupied orbitals and the plus sign to the unoccupied ones. The orbitals are therefore composed by two symmetrical energy surfaces (the Fermi energy is taken again equal to zero). They are depicted in Fig. 11.9. One can see that occupied and unoccupied MO are separated, except for particular points where they touch. They correspond to the points K, K' indicated in the panel (b) of Fig. 11.7. The peculiarities of these points can be summarized as follows.

1. They are isolated in the two-dimensional k-space.
2. The energies around these points show a linear dependence on the deviation of the momentum from the ones of the points.
3. They coincide with the Fermi level.
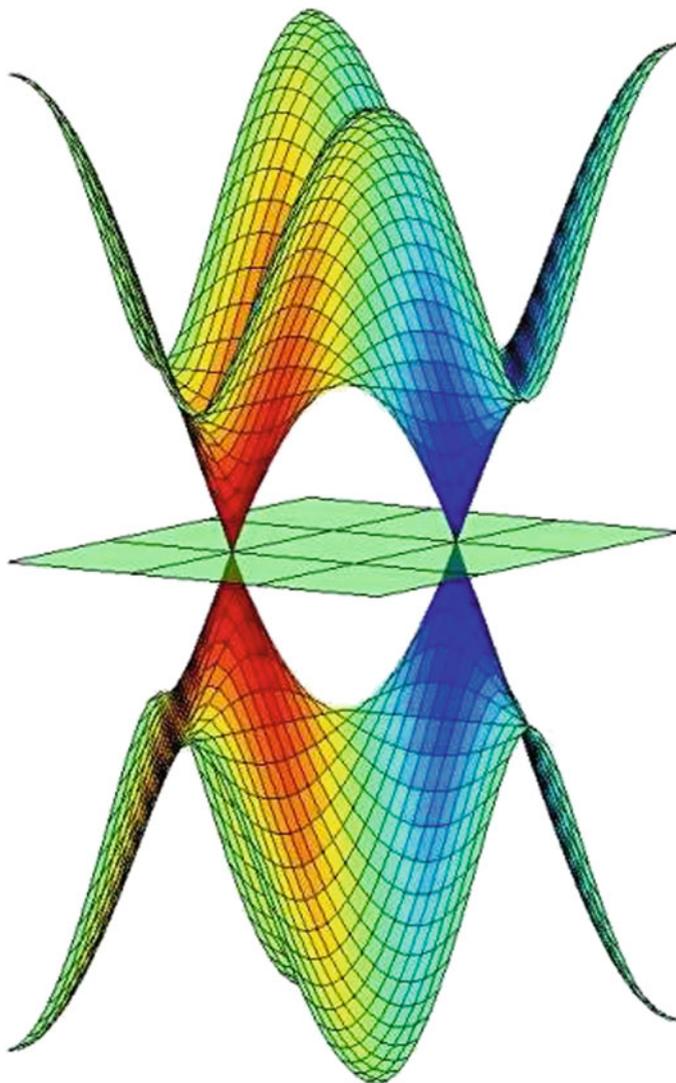
Fig. 11.9 Graphene MO energies as a function of momentum. *Upper* surface corresponds to the unoccupied MO, the *lower* one to the occupied MO. Credits to Naval Research Laboratory (http://www.eurekalert.org/multimedia/pub/31816php)

For the property (2) these points are called "Dirac points," since the energy-momentum relationship is the same as for a massless relativistic particle.

All that give rise to many exceptional properties of graphene [1]. However, it has to be checked that the picture of the MO of Fig. 11.9 is stable with respect to possible distortions, in particular to spin-density oscillations, as the ones that seem to

characterize the long polyacenes, as we have seen in the previous section "Polyacenes". Indeed one could argue that the effective particle-hole interaction could be attractive in some particular channel. In this case a correlation gap could be formed, and the electrical properties would be drastically modified. An equation similar to Eq. (11.17) can be devised to show if there is a solution for a bound state. However, it turns out that the density of state at the Dirac points is zero, and the argument based on the divergence of the integral kernel fails. In fact the density of state per spin N at the energy E near the Dirac points can be written as

$$N = 2\pi \int q dq \delta(E_q - E) = 2\pi \overline{q} / \left| \frac{dE_q}{dq} \right|, \qquad (11.20)$$

where the momentum q and the energy E_q are measured from the Dirac point. Here, $E_{\overline{q}} = E$. Notice that the phase space in the integral is two-dimensional because the Dirac point is isolated. The derivative in the denominator is a constant, because of the linear behavior of the energy, and therefore N vanishes at $\overline{q} = 0$. For this reason graphene is considered a *semi-metal*. Of course these considerations are only semi-quantitative and more refined calculations are desirable to establish on a firmer basis these conclusions. One must say that the electron structure of graphene has not yet been clarified in a satisfactory way. More work has still to be done in this direction.

In order to gain further insight on the difference between one-dimensional and two-dimensional structure for polycyclic molecules, one can follow an approach similar to the one described for polyacene, by considering molecules of finite size. This method was followed in Ref. [16], where molecules with the same symmetry of graphene (coronene and hexabenzocoronene) where studied in comparison with more "elongated" molecules (ovalene and pyrene). Their bond structures are depicted in Figs. 11.10 and 11.11, respectively. The lowest RPA excitation, which turns out to be a triplet one in all cases, was calculated employing the same semi-empirical interaction used for polyacene [6].

This eigenvalue is at 0.7 eV for coronene and hexabenzocoronene, which indicates a certain stability with respect to the size of the molecule. With an extreme extrap-

Fig. 11.10 Skeletons of hexabenzocoronene (*left*) and coronene (*right*) molecules. They have the same symmetry of graphene
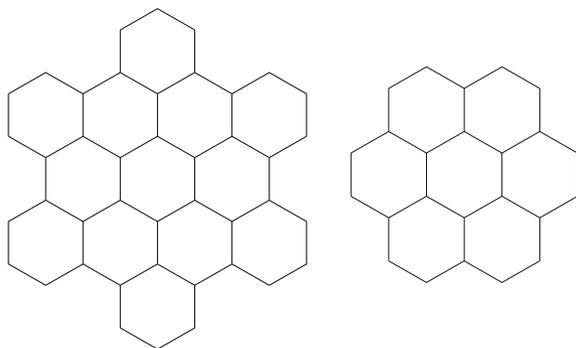
Fig. 11.11 Skeletons of
ovalene (*left*) and pyrene
(*right*) molecules. They have
an "elongated" structure
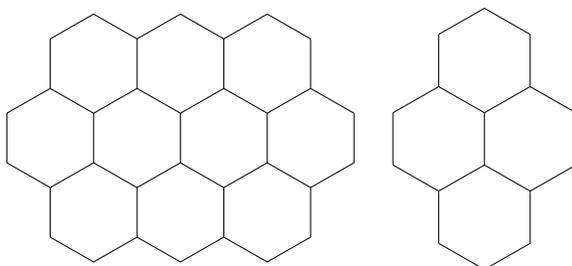
olation one could argue that the graphene ground state indeed does not show any
spin-density instability. On the contrary in the more "elongated" pyrene and ovalene
molecules the eigenvalue decreases from 0.43 to 0.07 eV. This is suggestive of the
effect that the geometrical shape can have on the possible appearance of spin-density
instability, which is in line (roughly speaking!) with what happens going from the
quasi-one dimensional polyacene to the two-dimensional graphene.

Summary

We have shown how the possible instability of the π-electron molecules can develop
going from the one-dimensional polyacetylene to the two-dimensional graphene,
and considering the polyacene as a molecule of character intermediate between one
and two dimensions. As it is well known, polyacetylene presents a Peierls distortion,
which affects his properties and gives rise to many chemical and physical phenomena.
The polyacene polymer is not expected to have Peierls distorsion, while the series of
polyacene molecules display a transition to a triplet ground sate at increasing number
of rings, probably in octacene. This is confirmed by the study of the polyacene
polymer, where one can show that a particle-hole bound state is present, which gives
rise to a correlation gap, and consequently to a semiconductor or insulator behavior.
Unfortunately, very long polyacenes have not been synthesized. The two-dimensional
graphene molecule has many peculiar properties. It can be shown that many of them,
in particular the electric properties, are a consequence of the *absence* of distorsions
like the spin-density oscillations, and in turn this is a consequence of the particular
honeycomb structure of the molecule.

References

1. V. Meunier, A.G. Souza Filho, E.B. Barros, M.S. Dresselhaus, Rev. Mod. Phys. **88**, 025005 (2016). DOI 10.1103/RevModPhys.88.025005
2. W. Su, J. Schrieffer, A. Heeger, Phys. Rev. B **22**(4), 2099 (1980). DOI 10.1103/PhysRevB.22.2099
3. W. Su, J. Schrieffer, A. Heeger, Phys. Rev. B **28**(2), 1138 (1983). DOI 10.1103/PhysRevB.28.1138
4. M. Barborini, L. Guidoni, J. Chem. Theory and Comp. **11**(9), 4109 (2015). DOI 10.1021/acs.jctc.5b00427
5. A.J. Heeger, Rev. Mod. Phys. **73**, 681 (2001). DOI 10.1103/RevModPhys.73.681
6. M. Baldo, A. Grassi, R. Pucci, P. Tomasello, J. Chem. Phys. **77**(5), 2438 (1982). DOI 10.1063/1.444112
7. R. Pucci, N.H. March, Phys. Lett. A **96**(2), 105 (1983). DOI 10.1016/0375-9601(83)90602-3
8. P.O. Löwdin, Phys. Rev. **97**, 1509 (1955). DOI 10.1103/PhysRev.97.1509
9. S. Rayne, K. Forest, Comp. Theor. Chem. **976**(1–3), 105 (2011). DOI 10.1016/j.comptc.2011.08.010
10. V. Heine, *Group Theory in Quantum Mechanics* (Pergamon Press, London, 1960)
11. L.N. Cooper, Phys. Rev. **104**, 1189 (1956). DOI 10.1103/PhysRev.104.1189
12. N.N. Tyutyulkov, Int J. Quantum Chem. **9**(4), 683 (1975). DOI 10.1002/qua.560090410
13. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, A.A. Firsov, Science **306**, 666 (2004). DOI 10.1126/science.1102896
14. A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, Rev. Mod. Phys. **81**(1), 109 (2009). DOI 10.1103/RevModPhys.81.109
15. P.R. Wallace, Phys. Rev. **71**, 622 (1947). DOI 10.1103/PhysRev.71.622
16. R. Pucci, M. Baldo, A. Martín-Rodero, G. Piccitto, P. Tomasello, Int. J. Quantum Chem. **26**(5), 783 (1984). DOI 10.1002/qua.560260518

Chapter 12
Generalized Dicke Model of Graphene Cavity QED

F.M.D. Pellegrino

Abstract We present a theory of the cavity quantum electrodynamics of graphene cyclotron resonance. By employing a canonical transformation, we derive an effective Hamiltonian for the system comprised of two neighboring Landau levels dressed by the cavity electromagnetic field (integer quantum Hall polaritons). This generalized Dicke Hamiltonian, which contains terms that are quadratic in the electromagnetic field and respects gauge invariance, is then used to verify the impossibility of superradiant instability.

Introduction

Light-matter interactions in graphene, a two-dimensional (2D) honeycomb crystal of carbon atoms [1–3], have been intensively explored in the past decade for both fundamental and applied purposes [4–7]. Recent experimental advances have made it possible to monolithically integrate graphene with optical microcavities [8, 9], paving the way for fundamental studies of cavity quantum electrodynamics (QED) [10] at the nanometer scale with graphene as an active medium. Another approach, which has been successful [11] in achieving the so-called strong-coupling regime of cavity QED [10] in conventional 2D electron systems in semiconductor quantum wells, consists in coupling graphene carriers with the photonic modes of an array of split-ring resonators [12].

Graphene-based cavity QED offers, at least in principle, a number of unique advantages. First, graphene is a highly tunable active medium since its electrical and heat transport properties can be easily controlled by employing gates [1–3]. Second, graphene offers many pathways to achieve the strong-coupling regime: these include (i) the exploitation of intrinsic Dirac plasmons [6, 7] and (ii) the combination of graphene with other plasmonic nanostructures [13]. Third, the active medium can be enriched by employing 2D vertical heterostructures [14–17] comprising graphene as

F.M.D. Pellegrino (✉)
NEST and Scuola Normale Superiore and Istituto Nanoscienze-CNR,
Piazza Dei Cavalieri, 7, Pisa, Italy
e-mail: francesco.pellegrino@sns.it

167

well as other 2D crystals/systems such as hexagonal boron nitride [18–20], transition metal dichalcogenides [21, 22] (e.g., MoS_2, WS_2, WSe_2), gallium arsenide quantum wells [23, 24], etc.

A central role in cavity QED is played by the Dicke model [25], which describes a non-dissipative closed system of identical two-level subsystems interacting with a single-mode radiation field. For a sufficiently strong light-matter coupling constant, the thermodynamic limit of the Dicke model exhibits a second-order quantum phase transition to a super-radiant state [26] with macroscopic photon occupation and coherent atomic polarization.

When an external magnetic field is applied to a 2D electron system, transitions between states in full and empty Landau levels (LLs) are dispersionless [27–29], mimicking atomic transitions and enabling [30] a condensed matter realization of the Dicke model. The light-matter interaction in the Dicke Hamiltonian is linear in the vector potential $\mathbf{A}_{em}$ of the cavity. For condensed matter states described by parabolic band models, a quadratic $\mathbf{A}_{em}^2$ term whose strength is related to the system's Drude weight and f-sum rule [29, 31], also emerges naturally from minimal coupling. It has long been understood [32–34] that the Dicke model's super-radiant phase transition is suppressed when the quadratic terms are retained. Demonstrations of this property are often referred to as *no-go* theorems.

The problem is more subtle in graphene, where electronic states near the charge neutrality point are described in a wide range of energies by a 2D massless Dirac fermion (MDF) Hamiltonian [2, 3]. The MDF Hamiltonian contains only one power of momentum $\mathbf{p}$: minimal coupling applied to this Hamiltonian does *not* generate a term proportional to $\mathbf{A}_{em}^2$.

The main scope is to lay down a formal theory of the cavity QED of the graphene cyclotron resonance. The key point is that one must derive a low-energy effective Hamiltonian by taking into account the coupling of the two-level systems which are resonant with the cavity photon field to all nonresonant states. This coupling is crucially important in the strong-coupling regime, where all the terms that are proportional to $\mathbf{A}_{em}^2$, which are generated by our renormalization procedure, must be taken into account. Indeed, these guarantee gauge invariance as well as a *no-go* theorem for the occurrence of a super-radiant phase transition.

This chapter is partially based on Ref. [35].

Generalized Dicke Hamiltonian

In this section, we derive an effective low-energy Hamiltonian for the cavity QED of the graphene cyclotron resonance.

Landau Levels in Graphene

At low energies, charge carriers in graphene are modeled by the usual single-channel massless Dirac fermion Hamiltonian [2, 3]

$$\mathscr{H}_{\mathrm{D}} = v_{\mathrm{D}} \sigma \cdot \mathbf{p} \, , \tag{12.1}$$

where $v_{\mathrm{D}} \approx 10^6$ m/s is the Dirac velocity. Here, $\sigma = (\sigma_x, \sigma_y)$ is a 2D vector of Pauli matrices acting on sublattice degrees-of-freedom, and $\mathbf{p} = -i\hbar\nabla_{\mathbf{r}}$ is the 2D momentum measured from one of the two corners (valleys) of the Brillouin zone.

A quantizing magnetic field $\mathbf{B} = B\hat{\mathbf{z}}$ perpendicular to the graphene sheet is coupled to the electronic degrees-of-freedom by replacing the canonical momentum $\mathbf{p}$ in Eq. (12.1) with the kinetic momentum $\boldsymbol{\Pi} = \mathbf{p} + e\mathbf{A}_0/c$, where $\mathbf{A}_0$ is the vector potential that describes the static magnetic field $\mathbf{B}$. The corresponding Hamiltonian is

$$\mathscr{H}_0 = v_{\mathrm{D}} \sigma \cdot \boldsymbol{\Pi} \, . \tag{12.2}$$

We work in the Landau gauge $\mathbf{A}_0 = -By\hat{\mathbf{x}}$. In this gauge the canonical momentum along the $\hat{\mathbf{x}}$ direction, p_x, coincides with magnetic translation operator [29] along the same direction and it commutes with the Hamiltonian $\mathscr{H}_0$. Thus, the eigenvalues of p_x are good quantum numbers. A complete set of eigenfunctions of the Hamiltonian $\mathscr{H}_0$ in Eq. (12.2) is provided by the two component pseudospinors [36]

$$\langle \mathbf{r} | \lambda, n, k \rangle = \frac{e^{ikx}}{\sqrt{2L}} \begin{pmatrix} w_{-,n}\phi_{n-1}(y - \ell_B^2 k) \\ \lambda w_{+,n}\phi_n(y - \ell_B^2 k) \end{pmatrix} \, , \tag{12.3}$$

where $\lambda = +1 \, (-1)$ denotes conduction (valence) band levels, $n \in \mathbb{N}$ is the Landau level (LL) index, and k is the eigenvalue of the magnetic translation operator in the $\hat{\mathbf{x}}$ direction. In Eq. (12.3)

$$w_{\pm,n} = \sqrt{1 \pm \delta_{n,0}} \tag{12.4}$$

guarantees that the pseudospinor corresponding to the $n = 0$ LL has weight only on one sublattice. Furthermore, $\phi_n(y)$ with $n = 0, 1, 2, \ldots$ are the normalized eigenfunctions of a 1D harmonic oscillator with frequency equal to the MDF cyclotron frequency $\omega_{\mathrm{c}} = \sqrt{2}v_{\mathrm{D}}/\ell_B$. Here, $\ell_B = \sqrt{\hbar c/(eB)} \simeq 25$ nm$/\sqrt{B[\text{Tesla}]}$ is the magnetic length.

The spectrum of the Hamiltonian (12.2) has the well-known form [36]

$$\varepsilon_{\lambda,n} = \lambda\hbar\omega_{\mathrm{c}}\sqrt{n} \, . \tag{12.5}$$

Each LL has a degeneracy $\mathscr{N} = N_{\mathrm{f}}S/(2\pi\ell_B^2)$, where $N_{\mathrm{f}} = 4$ is the spin-valley degeneracy and $S = L^2$ is the sample area.

Total Hamiltonian

We now couple the 2D electron system described by the Hamiltonian (12.2) to a single photon mode in a cavity. We denote by the symbol $\mathbf{A}_{\mathrm{em}}$ the vector potential that describes the cavity photon mode. Carriers in graphene are coupled to the cavity electromagnetic field via the minimal substitution:

$$\mathbf{\Pi} \rightarrow \mathbf{\Pi}' = \mathbf{\Pi} + \frac{e}{c}\mathbf{A}_{\mathrm{em}} \, . \tag{12.6}$$

The cavity vector potential $\mathbf{A}_{\mathrm{em}}$ will be treated within the dipole approximation. We can neglect the spatial dependence of the electromagnetic field in the cavity because the photon wavelength is much larger than any other length scale of the system.

Introducing photon annihilation a and creation operators $a^\dagger$ we can write

$$\mathbf{A}_{\mathrm{em}} = \sqrt{\frac{2\pi\hbar c^2}{\varepsilon\omega V}}\mathbf{e}_{\mathrm{em}}(a + a^\dagger), \tag{12.7}$$

where $\mathbf{e}_{\mathrm{em}}$ is a unit vector describing the polarization of the electromagnetic field, ω is the photon frequency, ε is the cavity dielectric constant, and $V = L_z L^2$ is the volume of the cavity. Here, $L_z \ll L$ is the length of the cavity in the $\hat{\mathbf{z}}$ direction.

The total Hamiltonian reads

$$\mathcal{H} = \mathcal{H}_{\mathrm{em}} + \mathcal{H}_0 + \mathcal{H}_{\mathrm{int}} \, , \tag{12.8}$$

where the first term is the cavity photon Hamiltonian, the second term is the MDF Hamiltonian in the presence of a quantizing magnetic field, i.e., Eq. (12.2), and the third term describes the coupling between MDFs and the cavity photon mode. More explicitly,

$$\mathcal{H}_{\mathrm{em}} = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) \, , \tag{12.9}$$

$$\mathcal{H}_0 = \sum_{\lambda,n,k} \varepsilon_{\lambda,n} c_{\lambda,n,k}^\dagger c_{\lambda,n,k} \, , \tag{12.10}$$

and

$$\mathcal{H}_{\mathrm{int}} = \frac{g}{\sqrt{\mathcal{N}}} \sum_{\lambda,\lambda',n,n',k} \left(\lambda w_{\lambda,n} e_{\mathrm{em}}^- \delta_{n',n+1} + \lambda' w_{\lambda',n'} e_{\mathrm{em}}^+ \delta_{n',n-1}\right) \left(a + a^\dagger\right) c_{\lambda',n',k}^\dagger c_{\lambda,n,k} \, . \tag{12.11}$$

In Eqs. (12.10) and (12.11) $c_{\lambda nk}^\dagger$ ($c_{\lambda,n,k}$) creates (annihilates) an electron with band index λ, LL index n, and wave number k. Finally,

$$g \equiv \hbar\omega_{\mathrm{c}}\sqrt{\frac{e^2}{2\varepsilon L_z \hbar\omega}} \, , \tag{12.12}$$

and $e_{\text{em}}^{\pm} = e_{\text{em}}^x \pm i e_{\text{em}}^y$, e_{em}^x and e_{em}^y being the components of the polarization vector $\mathbf{e}_{\text{em}}$.

We consider the integer quantum Hall regime in which a given number of LLs are fully occupied and the remaining ones are empty. Since the MDF Hamiltonian is particle-hole symmetric, we can consider, without loss of generality, the situation in which graphene is n-doped and the Fermi energy lies in conduction band ($\lambda = +$). We denote by $n = M$ the highest occupied LL. The lowest empty LL is therefore $n = M + 1$ and the Fermi energy lies in the middle between $n = M$ and $n = M + 1$, i.e.,

$$\mathscr{E}_M \equiv \frac{1}{2}\hbar\omega_{\text{c}}(\sqrt{M+1} + \sqrt{M}) . \tag{12.13}$$

Canonical Transformation

The aim of this section is to present a systematic procedure that allows us to derive an *effective* low-energy Hamiltonian for the LL doublet $n = M, M + 1$ as dressed by light-matter interactions. We are interested in the case in which the cavity photon is nearly resonant with the transition between the two conduction band LLs $n = M, M + 1$:

$$\hbar\omega \approx \Omega_M \equiv \hbar\omega_{\text{c}}(\sqrt{M+1} - \sqrt{M}) . \tag{12.14}$$

We anticipate [37] that the effective Hamiltonian will be different from the *bare* Dicke Hamiltonian that one obtains from Eqs. (12.9), (12.10), and (12.11) by selecting $\lambda = +1$ and $n = M, M + 1$, i.e.,

$$\mathscr{H}_{\text{Dicke}} = \mathscr{H}_{\text{em}} + \sum_{k=1}^{\mathscr{N}} \left[\mathscr{E}_M \mathbb{I}_k + \frac{\Omega_M}{2}\tau_k^z + \frac{g}{\sqrt{\mathscr{N}}}(a + a^\dagger)(e_{\text{em}}^- \tau_k^+ + e_{\text{em}}^+ \tau_k^-) \right] . \tag{12.15}$$

By direct comparison of Eqs. (12.9)–(12.11) with Eq. (12.15), we see that the Pauli matrices we have introduced are a shorthand for the following combinations of creation and destruction operators:

$$\mathbb{I}_k = c_{+,M+1,k}^\dagger c_{+,M+1,k} + c_{+,M,k}^\dagger c_{+,M,k} , \tag{12.16a}$$

$$\tau_k^z = c_{+,M+1,k}^\dagger c_{+,M+1,k} - c_{+,M,k}^\dagger c_{+,M,k} , \tag{12.16b}$$

$$\tau_k^+ = c_{+,M+1,k}^\dagger c_{+,M,k} , \tag{12.16c}$$

$$\tau_k^- = c_{+,M,k}^\dagger c_{+,M+1,k} . \tag{12.16d}$$

Here, $\mathbb{I}_k, \tau_k^z, \tau_k^\pm$ with $k = 1 \ldots \mathscr{N}$ is a set of Pauli matrices that act in the $2^{\mathscr{N}}$-fold degenerate subspace of the LL doublet $n = M, M + 1$, $\mathbb{I}_k$ being the 2×2 identity and $\tau_k^\pm \equiv (\tau_k^x \pm i\tau_k^y)/2$. More precisely, the final result of the canonical transformation yields a generalized Dicke Hamiltonian of the form—see Eq. (12.49):

$$\mathscr{H}_{\mathrm{GDH}} = \mathscr{H}_{\mathrm{Dicke}} + \Delta_M (a + a^\dagger)^2 + \sum_{k=1}^{\mathscr{N}} \left[\frac{\kappa}{\mathscr{N}} (a + a^\dagger)^2 \mathbb{I}_k - \frac{\kappa^z}{\mathscr{N}} (a + a^\dagger)^2 \tau_k^z \right] . \tag{12.17}$$

We notice that $\mathscr{H}_{\mathrm{GDH}}$ differs from the bare Dicke Hamiltonian (12.15) because of the presence of three terms that are quadratic in the operator $a + a^\dagger$, and that renormalize both $\mathscr{H}_{\mathrm{em}}$ and the light-matter interaction Hamiltonian. Microscopic expressions for the parameters Δ_M, κ, and κ^z are derived below.

We denote by the symbol $\mathscr{S}_M$ the subspace of the fermionic Hilbert space spanned by the two LLs which are resonantly coupled to the cavity field, i.e., $n = M, M + 1$, and lay on opposite sides of the Fermi energy. The symbol $\mathscr{S}_N$, on the other hand, denotes the subspace of the fermionic Hilbert space, which is comprised of all LLs *but* $n = M, M + 1$. We employ a canonical transformation with the aim of decoupling the LL doublet $n = M, M + 1$ from the $\mathscr{S}_N$ sector (see Refs. [38–40], and Chap. 8 in Ref. [29]).

Before proceeding further, it is convenient to rewrite the Hamiltonian (12.8) in the following manner:

$$\mathscr{H} = \mathscr{H}_{\mathrm{em}} + \mathscr{H}_0 + V_{\mathrm{D}} + V_{\mathrm{O}} , \tag{12.18}$$

where $\mathscr{H}_{\mathrm{em}}$ and $\mathscr{H}_0$ have been introduced in Eqs. (12.9) and (12.10), respectively, whereas the light-matter interaction Hamiltonian $\mathscr{H}_{\mathrm{int}}$ has been written as the sum of two terms: (i) V_{D}, which connects states either belonging to the subspace $\mathscr{S}_M$ or to the subspace $\mathscr{S}_N$ and (ii) V_{O}, which connects states belonging to different subspaces. Therefore V_{D} is a block-diagonal operator with one block referring to the $\mathscr{S}_M$ subspace and the other one to the $\mathscr{S}_N$ subspace. In the same representation, $\mathscr{H}_0$ is trivially a block-diagonal operator since it is a diagonal operator and $\mathscr{H}_{\mathrm{em}}$ is also a block-diagonal operator since it contains only photonic creation and annihilation operators and therefore acts as the identity operator with respect to fermionic labels. On the other hand, V_{O} is a block-off-diagonal operator in the same representation.

We now introduce an unitary transformation

$$U = e^S , \tag{12.19}$$

where S is its anti-Hermitian generator. The transformed Hamiltonian reads

$$\mathscr{H}' = e^S \mathscr{H} e^{-S} . \tag{12.20}$$

The spirit of the canonical transformation [38–40] is to transform the original Hamiltonian $\mathscr{H}$ onto an Hamiltonian $\mathscr{H}'$ that has no block-off-diagonal terms. A necessary condition to achieve this, is that the generator S be a block-off-diagonal operator.

The operator S can be found with the desired level of accuracy by following a perturbative approach. We use the Baker–Campbell–Hausdorff formula to rewrite Eq. (12.20):

$$\mathscr{H}' = \mathscr{H} + [S, \mathscr{H}] + \frac{1}{2!}[S, [S, \mathscr{H}]] + \dots , \tag{12.21}$$

where $[A, B]$ denotes the commutator between the two operators A and B.

We now expand the generator S in a power series:

$$S = \sum_{j=1}^{\infty} S^{(j)}, \tag{12.22}$$

where $S^{(j)}$ is proportional to $(g)^j$, i.e., the jth power of a suitable dimensionless coupling constant that is controlled by the strength g of light-matter interactions.

After inserting Eq. (12.22) in (12.21), we require that each term of the expansion cancels the corresponding block-off-diagonal term, order by order in the perturbative expansion in powers of g. This approach leads to a hierarchy of equations, one for each order in perturbation theory.

For example, the equation for the generator $S^{(1)}$ up to *first* order in g reads as follows:

$$[S^{(1)}, \mathscr{H}_0 + \mathscr{H}_{\mathrm{em}}] + V_{\mathrm{O}} = 0 . \tag{12.23}$$

The transformed Hamiltonian is given by the following expression:

$$\mathscr{H}' = \mathscr{H}_{\mathrm{em}} + \mathscr{H}_0 + V_{\mathrm{D}} + \frac{1}{2}[S^{(1)}, V_{\mathrm{O}}] + \mathscr{O}(g^3) . \tag{12.24}$$

In sections "Explicit Form of the Canonical Transformation up to Order g" and "Elimination of the Off-Diagonal Terms in $\mathscr{H}_N$ and Pauli Blocking" we derive the desired low-energy effective Hamiltonian by using the canonical transformation approach described in this section. The procedure is carried out in *three* steps: (i) we first decouple the subspace $\mathscr{S}_N$ from the subspace $\mathscr{S}_M$ by applying the canonical transformation S up to first order in the small parameter g—Eq. (12.23); (ii) we then use a different canonical transformation to take care of inter-band transitions between LLs belonging to the subspace $\mathscr{S}_N$; (iii) finally, we take into account Pauli blocking.

Explicit Form of the Canonical Transformation up to Order g

Following the notation of section "Canonical Transformation", we start from the original Hamiltonian in Eq. (12.18). Here, $\mathscr{H}_0$, which has been introduced in Eq. (12.10), refers to bare electrons in the presence of a quantizing magnetic field and it is diagonal with respect to spin projection, valley index, and the eigenvalue of the magnetic translation operator in the $\hat{\mathbf{x}}$ direction. It does not couple states belonging to the subspace $\mathscr{S}_M$ with states belonging to the subspace $\mathscr{S}_N$:

$$\mathscr{H}_0 = \sum_{m \in \mathscr{S}_M} \varepsilon_m c_m^\dagger c_m + \sum_{n \in \mathscr{S}_N} \varepsilon_n c_n^\dagger c_n \; . \qquad (12.25)$$

Here, $c_m^\dagger$ and $c_n^\dagger$ (c_m and c_n) are fermionic creation (annihilation) operators for a bare electron. We emphasize that, in this section, the indices m and n are *collective* labels for the spin projection along the $\hat{\mathbf{z}}$ axis, the valley index, the eigenvalue of the magnetic translation operator in the $\hat{\mathbf{x}}$ direction, the intra-band LL integer label, and the conduction/valence band label.

The Hamiltonian that couples electronic degrees-of-freedom with the electromagnetic field is written as a sum of a block-diagonal term V_D and a block-off-diagonal term V_O:

$$V_D = \sum_{m,m' \in \mathscr{S}_M} \frac{g_{mm'}}{\sqrt{\mathscr{N}}} \left(a + a^\dagger \right) c_m^\dagger c_{m'} + \sum_{n,n' \in \mathscr{S}_N} \frac{g_{nn'}}{\sqrt{\mathscr{N}}} \left(a + a^\dagger \right) c_n^\dagger c_{n'} \; , \qquad (12.26)$$

and

$$V_O = \sum_{m \in \mathscr{S}_M, n \in \mathscr{S}_N} \left[\frac{g_{mn}}{\sqrt{\mathscr{N}}} \left(a + a^\dagger \right) c_m^\dagger c_n + \frac{g_{nm}}{\sqrt{\mathscr{N}}} \left(a + a^\dagger \right) c_n^\dagger c_m \right] . \qquad (12.27)$$

In Eqs. (12.26) and (12.27) we have introduced

$$g_{mn} = \delta_{k,k'} \left(\lambda w_{\lambda,\bar{n}} e_{\text{em}}^- \delta_{\bar{m},\bar{n}+1} + \lambda' w_{\lambda',\bar{m}} e_{\text{em}}^+ \delta_{\bar{m},\bar{n}-1} \right) \; , \qquad (12.28)$$

where n (m) is the collective label $\bar{n}, \lambda, k$ ($\bar{m}, \lambda', k'$). Each of these three numbers represents an intra-band LL label ($\bar{n}, \bar{m}$), a band index (λ, λ'), and a collective label (k, k') comprising the eigenvalue of the magnetic translation operator in the $\hat{\mathbf{x}}$ direction, together with the spin projection along the along the $\hat{\mathbf{z}}$ axis and the valley index.

By solving Eq. (12.23) we obtain an explicit expression for the anti-Hermitian generator S up to first order in g:

$$S^{(1)} = \sum_{m \in \mathscr{S}_M, n \in \mathscr{S}_N} \left(\frac{g_{mn}}{\sqrt{\mathscr{N}}} \mathscr{A}_\omega c_m^\dagger c_n - \frac{g_{nm}}{\sqrt{\mathscr{N}}} c_n^\dagger c_m \mathscr{A}_\omega^\dagger \right) \; , \qquad (12.29)$$

where we have introduced the operator

$$\mathscr{A}_\omega \equiv \frac{a}{\varepsilon_{mn} - \hbar\omega} + \frac{a^\dagger}{\varepsilon_{mn} + \hbar\omega} \; . \qquad (12.30)$$

Given the first-order generator $S^{(1)}$, the commutator $[S^{(1)}, V_O]$ generates a new block-diagonal term. Using the dipole selection rules, the commutator reads

$$[S^{(1)}, V_O] = 2 \left(a + a^\dagger \right)^2 \sum_{m \in \mathscr{S}_M, n \in \mathscr{S}_N} \frac{\varepsilon_{mn}}{\varepsilon_{mn}^2 - (\hbar\omega)^2} \frac{g_{mn} g_{nm}}{\mathscr{N}} \left(c_m^\dagger c_m - c_n^\dagger c_n \right) + \mathscr{B}_\omega \; , \qquad (12.31)$$

where

$$\mathcal{B}_\omega = \frac{2\hbar\omega}{\mathcal{N}} \left[a, a^\dagger \right] \Bigg\{ \sum_{m \in \mathscr{S}_M, n \in \mathscr{S}_N} \frac{g_{mn} g_{nm}}{\varepsilon_{mn}^2 - \hbar^2\omega^2} \left(c_m^\dagger c_m + c_n^\dagger c_n \right)$$

$$+ \sum_{m,m' \in \mathscr{S}_M} \sum_{n,n' \in \mathscr{S}_N} \left[\frac{g_{m'n'} g_{mn}}{\varepsilon_{mn}^2 - \hbar^2\omega^2} c_{m'}^\dagger c_{n'} c_m^\dagger c_n + \frac{g_{n'm'} g_{mn}}{\varepsilon_{mn}^2 - \hbar^2\omega^2} c_{n'}^\dagger c_{m'} c_m^\dagger c_n \right.$$

$$\left. + \frac{g_{m'n'} g_{nm}}{\varepsilon_{mn}^2 - \hbar^2\omega^2} c_{m'}^\dagger c_{n'} c_n^\dagger c_m + \frac{g_{n'm'} g_{nm}}{\varepsilon_{mn}^2 - \hbar^2\omega^2} c_{n'}^\dagger c_{m'} c_n^\dagger c_m \right] \Bigg\} . \tag{12.32}$$

Note that $\mathcal{B}_0 = 0$. More importantly, the operator $\mathcal{B}_\omega$, being proportional to $\mathcal{N}^{-1}$, is negligible in the limit of a macroscopic LL degeneracy ($\mathcal{N} \gg 1$).

Using Eq. (12.24) and neglecting terms that are $\mathcal{O}(g^3)$, we finally find the effective Hamiltonian $\mathcal{H}'$, which is correct up to order g^2:

$$\mathcal{H}' = \mathcal{H}_{\mathrm{em}} + \mathcal{H}_M + \mathcal{H}_N . \tag{12.33}$$

Here, $\mathcal{H}_M$ is the sum of $\mathcal{N}$ independent contributions, one for each value of $k = 1 \ldots \mathcal{N}$, i.e., $\mathcal{H}_M = \sum_{k=1}^{\mathcal{N}} \mathcal{H}_k$ with

$$\mathcal{H}_k = \mathscr{E}_M \mathbb{I}_k + \frac{\Omega_M}{2} \tau_k^z + \frac{g}{\sqrt{\mathcal{N}}} (a + a^\dagger)(e_{\mathrm{em}}^- \tau_k^+ + e_{\mathrm{em}}^+ \tau_k^-)$$

$$- \frac{\kappa^z}{\mathcal{N}} (a + a^\dagger)^2 \tau_k^z + \frac{\kappa}{\mathcal{N}} (a + a^\dagger)^2 \mathbb{I}_k , \tag{12.34}$$

where $\mathscr{E}_M$ and Ω_M have been introduced earlier in Eqs. (12.13) and (12.14), respectively.

The quadratic terms in the electromagnetic field, i.e., the terms in the second line of Eq. (12.34), stem from the canonical transformation. In Eq. (12.34) we have introduced

$$\kappa^z \equiv \kappa_s^z - \kappa_d^z , \tag{12.35}$$

where the first term is independent of the cavity photon frequency while the second term, that we define "dynamical," explicitly depends on the cavity photon frequency:

$$\kappa_s^z = \frac{g^2}{\Omega_M} \tag{12.36}$$

and

$$\kappa_d^z = \frac{\omega^2}{\omega_c} \frac{g^2}{\hbar} \left\{ \frac{\sqrt{M+1}[\omega^2 - (4M+5)\omega_c^2]}{[(2M+3)\omega_c^2 - \omega^2]^2 - 4(M+1)(M+2)\omega_c^4} \right.$$

$$\left. + \frac{\sqrt{M}[\omega^2 - (4M-3)\omega_c^2]}{[(2M-1)\omega_c^2 - \omega^2]^2 - 4M(M-1)\omega_c^4} \right\} . \tag{12.37}$$

Note that $\kappa_{\mathrm{d}}^z = 0$ for $\omega = 0$. Finally,

$$\kappa = \frac{\omega^2}{\omega_{\mathrm{c}}} \frac{g^2}{\hbar} \left\{ \frac{\sqrt{M+1}[\omega^2 - (4M+5)\omega_{\mathrm{c}}^2]}{[(2M+3)\omega_{\mathrm{c}}^2 - \omega^2]^2 - 4(M+1)(M+2)\omega_{\mathrm{c}}^4} \right.$$
$$- \frac{\sqrt{M}[\omega^2 - (4M-3)\omega_{\mathrm{c}}^2]}{[(2M-1)\omega_{\mathrm{c}}^2 - \omega^2]^2 - 4M(M-1)\omega_{\mathrm{c}}^4}$$
$$\left. + \frac{\sqrt{M+1} - \sqrt{M}}{(\sqrt{M+1} + \sqrt{M})^2 \omega_{\mathrm{c}}^2 - \omega^2} \right\} . \tag{12.38}$$

The second term in Eq. (12.33) reads as follows:

$$\mathscr{H}_N = \sum_{n \in \mathscr{S}_N} \left[\varepsilon_n + \sum_{m \in \mathscr{S}_M} \frac{\varepsilon_{nm}}{\varepsilon_{nm}^2 - (\hbar\omega)^2} \left(a + a^\dagger \right)^2 \frac{g_{mn} g_{nm}}{\mathscr{N}} \right] c_n^\dagger c_n$$
$$+ \sum_{n,n' \in \mathscr{S}_N} \frac{g_{nn'}}{\sqrt{\mathscr{N}}} \left(a + a^\dagger \right) c_n^\dagger c_{n'} . \tag{12.39}$$

Elimination of the Off-Diagonal Terms in $\mathscr{H}_N$ and Pauli Blocking

The Hamiltonian (12.33) is not yet the desired result, i.e., an effective Hamiltonian for the $n = M$, $M + 1$ doublet. Indeed, $\mathscr{H}_N$ contains fermionic operators that act on the subspace $\mathscr{S}_N$.

In particular, we note that the last term in Eq. (12.39) is an off-diagonal contribution in the labels $n, n' \in \mathscr{S}_N$. We utilize a suitable canonical transformation that eliminates this term. For the sake of simplicity, we here report only the final result. We find a renormalized Hamiltonian operating on the subspace $\mathscr{S}_N$, which is diagonal in the labels $n, n' \in \mathscr{S}_N$:

$$\mathscr{H}_N' = \sum_{n \in \mathscr{S}_N} \varepsilon_n c_n^\dagger c_n + \left(a + a^\dagger \right)^2 \sum_{n \in \mathscr{S}_N} \sum_\ell \frac{\varepsilon_{n\ell} g_{n\ell} g_{\ell n}}{\varepsilon_{n\ell}^2 - \hbar^2 \omega^2} c_n^\dagger c_n , \tag{12.40}$$

where the index ℓ runs over *all* LLs.

Since the Dirac model applies over a large but *finite* energy region, we must regularize [37] Eq. (12.40) by employing a cut-off $\nu_{\max}$. Moreover, we treat the fermionic portion of the renormalized Hamiltonian (12.40) as a mean field for the photons. We therefore replace

$$c_n^\dagger c_n \rightarrow n_{\mathrm{F}}(\varepsilon_n) \equiv \frac{1}{\exp\left[(\varepsilon_n - \mu_{\mathrm{e}})/(k_{\mathrm{B}} T) \right] + 1} , \tag{12.41}$$

where μ_e is the chemical potential of the electronic system.

In the low-temperature limit,

$$k_B T \ll |\varepsilon_n - \mu_e| \,, \quad \forall n \in \mathscr{S}_N \,, \tag{12.42}$$

we can replace the Fermi–Dirac function in Eq. (12.41) with a Fermi step.

We are therefore led to define the prefactor of the $(a + a^\dagger)^2$ term in Eq. (12.40) as

$$\Delta_M(\nu_{max}) = \sum_{n \in \mathscr{S}_N} \sum_{\ell} \frac{\varepsilon_{n\ell} g_{n\ell} g_{\ell n}}{\varepsilon_{n\ell}^2 - \hbar^2 \omega^2} \Theta(\mu_e - \varepsilon_n) \,, \tag{12.43}$$

where the sums are regularized with the cut-off ν_{max}. More explicitly, for every $M \neq 0$, we have:

$$\Delta_M(\nu_{max}) = -2\varepsilon_{max} \frac{g^2}{\hbar^2 \omega_c^2} + \frac{g^2}{\hbar \omega_c} \mathscr{I}_{M-1}(\nu_{max}) - \frac{g^2}{\Omega_{M-1}} \frac{\hbar^2 \omega^2}{\Omega_{M-1}^2 - \hbar^2 \omega^2} \,, \tag{12.44}$$

where $\varepsilon_{max} \equiv \hbar \omega_c \sqrt{\nu_{max}}$ and

$$\mathscr{I}_{M-1}(\nu_{max}) = \sum_{\ell=M}^{\nu_{max}} \left[\frac{(\sqrt{\ell+1} - \sqrt{\ell})\omega^2}{\omega^2 - \omega_c^2(\sqrt{\ell} + \sqrt{\ell+1})^2} + \frac{(\sqrt{\ell} - \sqrt{\ell-1})\omega^2}{\omega^2 - \omega_c^2(\sqrt{\ell-1} + \sqrt{\ell})^2} \right]. \tag{12.45}$$

As explained in Refs. [37, 41], we must regularize the expression in Eq. (12.44) by subtracting the cut-off dependent term $-2\varepsilon_{max} \, g^2/(\hbar^2 \omega_c^2)$. After applying this regularization, one can take the limit $\nu_{max} \to \infty$, discovering that the quantity

$$\begin{aligned} \Delta_M &\equiv \lim_{\nu_{max} \to \infty} \left[\Delta_M(\nu_{max}) + 2\varepsilon_{max} \frac{g^2}{\hbar^2 \omega_c^2} \right] \\ &= \frac{g^2}{\hbar \omega_c} \mathscr{I}_{M-1}^\infty - \frac{g^2}{\Omega_{M-1}} \frac{\omega^2}{\Omega_{M-1}^2 - \omega^2} \,, \end{aligned} \tag{12.46}$$

with

$$\begin{aligned} \mathscr{I}_M^\infty &\equiv \lim_{\nu_{max} \to \infty} \mathscr{I}_M(\nu_{max}) \\ &= \sum_{\ell=M+1}^{\infty} \left[\frac{(\sqrt{\ell+1} - \sqrt{\ell})\omega^2}{\omega^2 - \omega_c^2(\sqrt{\ell} + \sqrt{\ell+1})^2} + \frac{(\sqrt{\ell} - \sqrt{\ell-1})\omega^2}{\omega^2 - \omega_c^2(\sqrt{\ell-1} + \sqrt{\ell})^2} \right], \end{aligned} \tag{12.47}$$

is well defined.

Discarding constant terms [29] (i.e., terms that do not contain the photon field operators a and $a^\dagger$), the renormalized Hamiltonian (12.40) becomes

$$\mathscr{H}'_N = \Delta_M \left(a + a^\dagger\right)^2 . \tag{12.48}$$

We stress that Δ_M as defined in Eq. (12.46) depends both on the LL label M and the photon frequency ω and that it vanishes in the static $\omega = 0$ limit.

Final Result for the Effective Hamiltonian

In summary, the correct low-energy Hamiltonian is given by $\mathscr{H}'$ as in Eq. (12.33) with $\mathscr{H}_N$ replaced by $\mathscr{H}'_N$ in Eq. (12.48), i.e.,

$$\mathscr{H}_{\text{GDH}} \equiv \mathscr{H}_{\text{em}} + \Delta_M(a + a^\dagger)^2 + \sum_{k=1}^{\mathscr{N}} \mathscr{H}_k , \tag{12.49}$$

where $\mathscr{H}_k$ has been defined in Eq. (12.34) and, without loss of generality, we have chosen a specific polarization of the electromagnetic field, i.e., $\mathbf{e}_{\text{em}} = \mathbf{u}_x$.

Equation (12.49) represents a low-energy effective Hamiltonian for the cavity QED of the graphene cyclotron resonance. It is evident that $\mathscr{H}_{\text{GDH}}$ differs from the bare Dicke Hamiltonian (12.15) since it contains terms that are quadratic in the electromagnetic field. We will therefore refer to the low-energy effective Hamiltonian (12.49) as to the generalized Dicke Hamiltonian (GDH).

For future purposes, it is useful to highlight the following identity,

$$\Delta_M = \frac{g^2}{\Omega_M} + \frac{g^2}{\hbar\omega_c}\mathscr{I}_M^\infty - \kappa^z - \kappa , \tag{12.50}$$

and the following inequality

$$F_M(\omega) \leq \mathscr{I}_M^\infty \leq F_{M+1}(\omega) , \tag{12.51}$$

which is valid $\omega \leq \omega_c\sqrt{M}$. Here,

$$F_M(\omega) \equiv \frac{\omega}{2\omega_c} \log\left(\frac{2\omega_c\sqrt{M} - \omega}{2\omega_c\sqrt{M} + \omega}\right). \tag{12.52}$$

For large M one therefore finds

$$\mathscr{I}_M^\infty \simeq \frac{\omega}{2\omega_c} \log\left(\frac{2\mathscr{E}_M - \hbar\omega}{2\mathscr{E}_M + \hbar\omega}\right) . \tag{12.53}$$

In the resonant $\hbar\omega = \Omega_M$ case, the quantities κ^z and κ defined earlier in Eqs. (12.35)–(12.38) reduce to:

$$\kappa^z = -\frac{g^2}{\hbar\omega_c}\frac{1}{2}\sqrt{M}\,, \tag{12.54}$$

and

$$\kappa = \frac{g^2}{\hbar\omega_c}\left[(M+1)\sqrt{M+1} + \left(M - \frac{1}{2}\right)\sqrt{M}\right.$$
$$\left. + \frac{1}{4\sqrt{M(M+1)}(\sqrt{M+1} + \sqrt{M})^3}\right]. \tag{12.55}$$

Linear-Response Theory Analysis

In this section we demonstrate that the GDH, Eq. (12.49), is gauge invariant. To this end, we treat the cavity electromagnetic field as a weak perturbation with respect to the MDF Hamiltonian in the presence of a quantizing magnetic field. The cavity electromagnetic field induces a matter current that can be calculated by the powerful means of linear-response theory [29, 31]. In particular, the physical matter current in response to the electromagnetic field is composed by paramagnetic and diamagnetic contributions [29, 31].

It is easy to demonstrate that the paramagnetic response function of a system described by the GDH (12.49) to the electromagnetic field is given by

$$\chi_P(\omega) = \frac{g^2}{\mathcal{N}}\langle\langle\tau_{tot}^x; \tau_{tot}^x\rangle\rangle_\omega = g^2\frac{2\Omega_M}{\hbar^2\omega^2 - \Omega_M^2}\tanh\left(\frac{\beta\Omega_M}{4}\right), \tag{12.56}$$

where $\tau_{tot}^x = \sum_{k=1}^{\mathcal{N}}\tau_k^x$ and $\beta = 1/(k_BT)$. In Eq. (12.56) we have introduced the Kubo product [29]

$$\langle\langle A; B\rangle\rangle_\omega \equiv -\frac{i}{\hbar}\int_0^\infty dt\, e^{i(\omega+i0^+)t}\langle[A(t), B]\rangle\,, \tag{12.57}$$

where $\langle\ldots\rangle$ denotes a thermal average and $A(t)$ is the operator A in the Heisenberg representation, i.e., $A(t) \equiv \exp(i\mathcal{H}_{GDH}t)A\exp(-i\mathcal{H}_{GDH}t)$.

Similarly, the diamagnetic response function is given by

$$\chi_D(\omega) = \frac{2}{\mathcal{N}}\langle\langle\kappa\mathbb{I}_{tot} - \kappa^z\tau_{tot}^z\rangle\rangle_\omega + 2\Delta_M = 2\kappa + 2\Delta_M + 2\kappa^z\tanh\left(\frac{\beta\Omega_M}{4}\right), \tag{12.58}$$

where $\tau_{tot}^z = \sum_{k=1}^{\mathcal{N}}\tau_k^z$ and $\mathbb{I}_{tot} = \sum_{k=1}^{\mathcal{N}}\mathbb{I}_k$.

The diamagnetic response function $\chi_D(\omega)$ can be rewritten in a compact form as

$$\chi_D(\omega) = 2\Omega_g\,, \tag{12.59}$$

where

$$\Omega_g = \Omega_g(\beta) \equiv \frac{g^2}{\Omega_M} + \frac{g^2}{\hbar\omega_c}\mathscr{I}_M^\infty - \kappa^z\left[1 - \tanh\left(\beta\Omega_M/4\right)\right] . \qquad (12.60)$$

In writing Eqs. (12.59) and (12.60) we have used the mathematical identity (12.50).

Therefore, the physical current–current response function is the sum of these two contributions:

$$\chi_J(\omega) = \chi_P(\omega) + \chi_D(\omega) = g^2\frac{2\Omega_M}{\hbar^2\omega^2 - \Omega_M^2}\tanh\left(\frac{\beta\Omega_M}{4}\right) + 2\Omega_g . \quad (12.61)$$

In the static $\omega = 0$ limit we have

$$\chi_P(\omega \to 0) = -\frac{2g^2}{\Omega_M}\tanh\left(\frac{\beta\Omega_M}{4}\right) \qquad (12.62)$$

and

$$\chi_D(\omega \to 0) = 2\kappa_s^z\tanh\left(\frac{\beta\Omega_M}{4}\right) = \frac{2g^2}{\Omega_M}\tanh\left(\frac{\beta\Omega_M}{4}\right) . \qquad (12.63)$$

Paramagnetic and diamagnetic contributions in Eqs. (12.62) and (12.63) are equal in magnitude and opposite in sign. Hence, a quasi-homogeneous vector potential does not induce any response in the static limit: in this limit the vector potential represents a pure gauge and cannot induce any physical effect unless gauge invariance is broken [29, 31].

Alert readers will note that the paramagnetic contribution to the physical current–current response function dominates over the diamagnetic contribution in the resonant limit $\hbar\omega \to \Omega_M$. Indeed, $\chi_P(\omega)$ has a pole at $\hbar\omega \to \Omega_M$, while $\chi_D(\omega)$ is finite at the same frequency. As we will see below in section "No-Go Theorem for the Super-Radiant Phase Instability", however, the quadratic terms in the photon field in Eq. (12.49), which yield a finite diamagnetic response, are absolutely crucial to ensure thermodynamic stability of the system.

No-Go Theorem for the Super-Radiant Phase Instability

In this section, we verify the nonexistence of the super-radiant phase instability of the GDH, Eq. (12.49). The simplest approximation to evaluate the ground state energy is a mean-field approach, where the bosonic fields are treated as classical fields, i.e., complex numbers, namely one replaces a and $a^\dagger$ with ϕ and ϕ^*. This approximation is a good one when the average photon number is macroscopic, i.e., when it is $\mathscr{O}(\mathscr{N})$. This is precisely what occurs in a super-radiant phase. Under this mean field approach, one finds a noninteracting two-level system coupled with an external classical field

$$\mathscr{H}_{\mathrm{MF}} = \mathscr{N}[\hbar\omega y^2 + (\hbar\omega + 4\Delta_M + 4\kappa)x^2] + \sum_k \left(\frac{\Omega_M}{2} - 4\kappa_z x^2\right)\tau_k^z + 2gx\tau_k^z,$$

(12.64)

where $\phi = \sqrt{\mathscr{N}}(x + iy)$. By solving the eigenvalue problem, the ground state energy up to $\mathscr{O}(g^2)$ is expressed as

$$\mathscr{E}_{\mathrm{MF}} = \mathscr{N}[\hbar\omega y^2 + (\hbar\omega + 4\Delta_M + 4\kappa)x^2] - \sqrt{(\Omega_M/2)^2 - 4[g^2 - \Omega_M\kappa_z]x^2},$$

(12.65)

and it is a function of classical fields x and y. Since $\omega > 0$, the ground state energy is minimized for $y = 0$, while further remarks need about x. Since $\mathscr{E}_{\mathrm{MF}}$ depends on x through x^2, then $x_0 = 0$ is always an extremum. Physically, the solution $x_0 = 0$ corresponds to the "normal phase" in which the number of photons vanishes in the thermodynamic limit. A finite extreme $x_0 \neq 0$ exists if and only if

$$\frac{(\hbar\omega + 4\Delta_M + 4\kappa)\Omega_M}{4(g^2 - \Omega_M\kappa_z)} < 1,$$

(12.66)

in this case the solution corresponds to the "super-radiant phase." By using Eq. (12.50) and $g/(\hbar\omega_c) < 1$, one finds that the previous condition is satisfied only if the dimensionless function

$$\bar{f}_M(\omega) \equiv -4\frac{\omega_c}{\omega}\mathscr{I}_M^\infty,$$

(12.67)

is *larger* than unity. Note that $\bar{f}_M(\omega)$ is *independent* of the cavity dielectric constant ε. Since we are interested in the resonant regime, we can set $\omega = \Omega_M/\hbar$ in Eq. (12.67). In this case, $\bar{f}_M$ becomes a function of the LL label M only. Figure 12.1 illustrates the dependence of $\bar{f}_M = \bar{f}_M(\omega = \Omega_M/\hbar)$ on M. We clearly see that $\bar{f}_M(\omega = \Omega_M/\hbar) < 1$ for every M. We can therefore conclude that $x_0 = 0$ is always the only minimum of $\mathscr{E}_{\mathrm{MF}}$, and one can exclude the occurrence of a super-radiant instability.

In this section, we have verified the absence of super-radiant instability at zero temperature within a mean field approach, in Ref. [35] within the path-integral formalism it is verified the impossibility to have a transition to a super-radiant phase

Fig. 12.1 Dependence of the function $\bar{f}_M$ defined in Eq. (12.67) on the LL index M
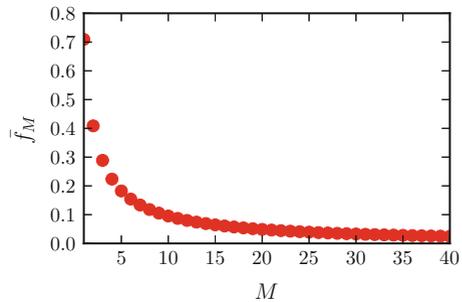
at finite temperature, and the robustness of normal phase with respect to quantum fluctuations of the electromagnetic field.

Finally we show that a super-radiant phase transition can occur when the quadratic terms in the photon field are neglected [26].

In this case, a maximum of $\mathscr{E}_{\mathrm{MF}}$ at $x_0 \neq 0$ can occur if [26, 42]

$$\frac{\hbar\omega\Omega_M}{4g^2} < 1 . \tag{12.68}$$

This implies that, choosing a suitable cavity dielectric constant for a given M or a value of the LL index M for a given ε, a super-radiant phase transition is possible. Consider, for instance, a half-wavelength cavity and set $\hbar\omega = \Omega_M$, where $\omega = \pi c/(L_z \sqrt{\varepsilon})$. In this case, $g = \hbar\omega_{\mathrm{c}}\sqrt{\alpha/(2\pi\sqrt{\varepsilon})}$ and the critical condition (12.68) becomes: $\sqrt{M+1} + \sqrt{M} > 2\pi\sqrt{\varepsilon}/\alpha$. A super-radiant phase transition is therefore possible [42] for large enough values of M. If the condition (12.68) is satisfied, the minimum of the ground state energy appears at

$$x_0 = \frac{g}{\hbar\omega}\left[1 - \left(\frac{\hbar\omega\Omega_M}{4g^2}\right)^2\right]^{1/2} \tag{12.69}$$

in the zero-temperature limit. Hence, one can gain energy when the photon occupation number becomes macroscopic, $n_{\mathrm{ph}} = x_0^2 \mathscr{N}$.

These are artifacts stemming from the neglect of quadratic terms in the photon field.

Summary and Conclusions

We have presented a theory of the cavity QED of the graphene cyclotron resonance. We have first employed a canonical transformation to derive an effective Hamiltonian for the system comprised of two neighboring Landau levels dressed by the cavity electromagnetic field (integer quantum Hall polaritons). The final result is contained in Eq. (12.49). This effective Hamiltonian, which we have termed "generalized Dicke Hamiltonian," respects gauge invariance and contains terms that are *quadratic* in the electromagnetic field. We have then used Eq. (12.49) and a mean field approximation to confirm that no super-radiant phase transitions are possible in the cavity QED of the graphene cyclotron resonance.

References

1. A.K. Geim, K.S. Novoselov, Nat. Mater. **6**(3), 183 (2007). DOI 10.1038/nmat1849
2. A.H. Castro Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, Rev. Mod. Phys. **81**(1), 109 (2009). DOI 10.1103/RevModPhys.81.109

3. M.I. Katsnelson, *Graphene: Carbon in Two Dimensions* (Cambridge University Press, Cambridge, 2012)
4. F. Bonaccorso, Z. Sun, T. Hasan, A.C. Ferrari, Nature Photonics **4**(9), 611 (2010). DOI 10.1038/nphoton.2010.186
5. N.M.R. Peres, Rev. Mod. Phys. **82**(3), 2673 (2010). DOI 10.1103/RevModPhys.82.2673
6. F.H.L. Koppens, D.E. Chang, F.J. García de Abajo, Nano Lett. **11**(8), 3370 (2011). DOI 10.1021/nl201771h
7. A.N. Grigorenko, M. Polini, K.S. Novoselov, Nature Photonics **6**(11), 749 (2012). DOI 10.1038/nphoton.2012.262
8. M. Engel, M. Steiner, A. Lombardo, A.C. Ferrari, H.v. Löhneysen, P. Avouris, R. Krupke, Nature Commun. **3**, 906 (2012). DOI 10.1038/ncomms1911
9. M. Furchi, A. Urich, A. Pospischil, G. Lilley, K. Unterrainer, H. Detz, P. Klang, A.M. Andrews, W. Schrenk, G. Strasser, T. Mueller, Nano Lett. **12**(6), 2773 (2012). DOI 10.1021/nl204512x
10. J.M. Raimond, M. Brune, S. Haroche, Rev. Mod. Phys. **73**, 565 (2001). DOI 10.1103/RevModPhys.73.565
11. G. Scalari, C. Maissen, D. Turčinková, D. Hagenmüller, S. De Liberato, C. Ciuti, C. Reichl, D. Schuh, W. Wegscheider, M. Beck, J. Faist, Science **335**(6074), 1323 (2012). DOI 10.1126/science.1216022
12. F. Valmorra, G. Scalari, C. Maissen, W. Fu, C. Schönenberger, J.W. Choi, H.G. Park, M. Beck, J. Faist, Nano Lett. **13**(7), 3193 (2013). DOI 10.1021/nl4012547
13. T.J. Echtermeyer, L. Britnell, P.K. Jasnos, A. Lombardo, R.V. Gorbachev, A.N. Grigorenko, A.K. Geim, A.C. Ferrari, K.S. Novoselov, Nature Commun. **2**, 458 (2011). DOI 10.1038/ncomms1464
14. K.S. Novoselov, Rev. Mod. Phys. **83**, 837 (2011). DOI 10.1103/RevModPhys.83.837
15. K.S. Novoselov, A.H. Castro Neto, Physica Scripta **2012**(T146), 014006 (2012). DOI 10.1088/0031-8949/2012/T146/014006
16. F. Bonaccorso, A. Lombardo, T. Hasan, Z. Sun, L. Colombo, A.C. Ferrari, Materials Today **15**(12), 564 (2012). DOI 10.1016/S1369-7021(13)70014-2
17. A.K. Geim, I.V. Grigorieva, Nature **499**(7459), 419 (2013). DOI 10.1038/nature12385
18. L.A. Ponomarenko, A.K. Geim, A.A. Zhukov, R. Jalil, S.V. Morozov, K.S. Novoselov, I.V. Grigorieva, E.H. Hill, V.V. Cheianov, V.I. Fal'ko, K. Watanabe, T. Taniguchi, R.V. Gorbachev, Nat. Phys. **7**(12), 958 (2011). DOI 10.1038/nphys2114
19. R.V. Gorbachev, A.K. Geim, M.I. Katsnelson, K.S. Novoselov, T. Tudorovskiy, I.V. Grigorieva, A.H. MacDonald, S.V. Morozov, K. Watanabe, T. Taniguchi, L.A. Ponomarenko, Nat Phys **8**(12), 896 (2012). DOI 10.1038/nphys2441
20. L. Britnell, R.V. Gorbachev, R. Jalil, B.D. Belle, F. Schedin, A. Mishchenko, T. Georgiou, M.I. Katsnelson, L. Eaves, S.V. Morozov, N.M.R. Peres, J. Leist, A.K. Geim, K.S. Novoselov, L.A. Ponomarenko, Science **335**(6071), 947 (2012). DOI 10.1126/science.1218461
21. Q.H. Wang, K. Kalantar-Zadeh, A. Kis, J.N. Coleman, M.S. Strano, Nature Nanotech. **7**(11), 699 (2012). DOI 10.1038/nnano.2012.193
22. L. Britnell, R.M. Ribeiro, A. Eckmann, R. Jalil, B.D. Belle, A. Mishchenko, Y.J. Kim, R.V. Gorbachev, T. Georgiou, S.V. Morozov, A.N. Grigorenko, A.K. Geim, C. Casiraghi, A.H. Castro Neto, K.S. Novoselov, Science **340**(6138), 1311 (2013). DOI 10.1126/science.1235547
23. A. Principi, M. Carrega, R. Asgari, V. Pellegrini, M. Polini, Phys. Rev. B **86**(8), 085421 (2012). DOI 10.1103/PhysRevB.86.085421
24. A. Gamucci, D. Spirito, M. Carrega, B. Karmakar, A. Lombardo, M. Bruna, L.N. Pfeiffer, K.W. West, A.C. Ferrari, M. Polini, V. Pellegrini, Nature Commun. **5**, 5824 (2014). DOI 10.1038/ncomms6824
25. R.H. Dicke, Phys. Rev. **93**, 99 (1954). DOI 10.1103/PhysRev.93.99
26. K. Hepp, E.H. Lieb, Ann. Phys. (N.Y.) **76**(2), 360 (1973). DOI 10.1016/0003-4916(73)90039-0
27. R.E. Prange, S.M. Girvin, *The Quantum Hall Effect* (Springer, New York, 1990)
28. A.H. MacDonald, in *Proceedings of the Les Houches Summer School on Mesoscopic Physics*, ed. by E. Akkermans, G. Montambeaux, J.L. Pichard (Elsevier, Amsterdam, 1995)

29. G. Giuliani, G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, 2005)
30. D. Hagenmüller, S. De Liberato, C. Ciuti, Phys. Rev. B **81**, 235303 (2010). DOI 10.1103/ PhysRevB.81.235303
31. D. Pines, P. Noziéres, *The Theory of Quantum Liquids* (W. A. Benjamin, New York, 1966)
32. K. Rzażewski, K. Wódkiewicz, W. Żakowicz, Phys. Rev. Lett. **35**, 432 (1975). DOI 10.1103/ PhysRevLett.35.432
33. I. Bialynicki-Birula, K. Rzazewski, Phys. Rev. A **19**, 301 (1979). DOI 10.1103/PhysRevA.19. 301
34. K. Gawedzki, K. Rzazewski, Phys. Rev. A **23**, 2134 (1981). DOI 10.1103/PhysRevA.23.2134
35. F.M.D. Pellegrino, L. Chirolli, R. Fazio, V. Giovannetti, M. Polini, Phys. Rev. B **89**, 165406 (2014). DOI 10.1103/PhysRevB.89.165406
36. M.O. Goerbig, Rev. Mod. Phys. **83**, 1193 (2011). DOI 10.1103/RevModPhys.83.1193
37. L. Chirolli, M. Polini, V. Giovannetti, A.H. MacDonald, Phys. Rev. Lett. **109**, 267404 (2012). DOI 10.1103/PhysRevLett.109.267404
38. D.R. Hamann, A.W. Overhauser, Phys. Rev. **143**, 183 (1966). DOI 10.1103/PhysRev.143.183
39. J.R. Schrieffer, P.A. Wolff, Phys. Rev. **149**(2), 491 (1966). DOI 10.1103/PhysRev.149.491
40. S. Bravyi, D.P. DiVincenzo, D. Loss, Ann. Phys. (N.Y.) **326**(10), 2793 (2011). DOI 10.1016/ j.aop.2011.06.004
41. A. Principi, M. Polini, G. Vignale, Phys. Rev. B **80**, 075418 (2009). DOI 10.1103/PhysRevB. 80.075418
42. D. Hagenmüller, C. Ciuti, Phys. Rev. Lett. **109**, 267403 (2012). DOI 10.1103/PhysRevLett. 109.267403

Chapter 13
A Comparison Between Quantum Transport and Band Structure Unfolding in Defected Graphene Nanoribbons

I. Deretzis and A. La Magna

Abstract This article considers a graphene-based quasi-one-dimensional system and explores the impact of structural perturbations on the electronic and transport properties of the material. Two phenomenologically different quantum mechanical approaches are used to describe the perturbation, namely the spectral weight of the unfolded band structure and the transmission coefficient of the propagated electrons. We show that these two descriptions present strong qualitative similarities and yield complementary information for the understanding of the induced electronic alterations.

Introduction

During the last decade, the isolation of a stable and truly two-dimensional (2D) system like graphene [1] has given the possibility to experimentally probe for the quantum mechanical nature of electronic transport in strongly confined systems, even at elevated temperatures. Characteristic examples are the measurement of ballistic transport for graphene systems on various substrates [2–5] and the manifestation of a half-integer quantum hall effect [6–10]. Moreover, when a 2D graphene sheet is further confined in a quasi one-dimensional (1D) structure through patterning and nanolithography (usually referred to as a graphene nanoribbon), integer plateaus on the conductance appear, corresponding to the presence of 1D sub-bands in the electronic structure [11]. Finally, the signature of single defects or adatoms can be resolved through scanning tunneling microscopy measurements [12].

Such fully quantum mechanical picture necessitates for an adequate theoretical description of charge transport with an atomic accuracy, as alterations of the structural symmetry even in single lattice sites can induce significant modifications in

I. Deretzis (✉) · A. La Magna
Istituto per la Microelettronica e Microsistemi (CNR-IMM), Z.I. VIII Strada 5,
95121 Catania, Italy
e-mail: ioannis.deretzis@imm.cnr.it

A. La Magna
e-mail: antonino.lamagna@imm.cnr.it

the calculated conductance spectra [13]. A widespread paradigm of such methodology is the non-equilibrium Green's function formalism [14], used for the calculation of transport in both ideal and defected/disordered graphene systems [13, 15–17]. Such formalism naturally incorporates the concept of electronic scattering either by structural defects or by the metallic contacts. It is by now well-consolidated that the presence of structural imperfections and the perturbations that these induce in the otherwise ideal electronic structure is one of the main origins for conductance degradation in graphene-based systems [15]. Hence, perturbation potentials can be considered as the cause for such scattering mechanisms, whereas conductance degradation (with respect to the ideal value) their immediate consequence. Within this picture, a key quantity for the determination of the current-carrying capacity of a single quantum channel is the transmission coefficient $T(\varepsilon)$ [14], which can obtain values between zero and one, corresponding to complete reflection and perfect transmission of the propagated electrons, respectively. The latter can be only retrieved for ideal systems, whereas a lower value is usually calculated for strongly disordered structures.

In the last years, a group of methodologies have been developed that are particularly suited for the calculation of the electronic structure in systems with defects. They consist in using big supercells that contain the defects, calculating the electronic properties for such supercells and consequently unfolding the supercell band structure [18–26]. The resulting unfolded band structure contains information not only of the original unit-cell bands, but also on the impact of the induced perturbation on the robustness of the total band structure. The key quantity here is the spectral weight $w(\varepsilon, \mathbf{k})$ of each unfolded band, whose value ranges form zero to one, and indicates the degree of maintenance of the original bands in the defected supercell [26]. The $0 < w(\varepsilon, \mathbf{k}) < 1$ range of the spectral weight with respect to the perturbation potential, unavoidably brings to mind the respective $0 < T(\varepsilon) < 1$ range of the transmission coefficient in quantum transport calculations of defected systems. In this article we attempt to explore the relationship between quantum transport and band structure unfolding using a graphene nanoribbon with a single-vacancy defect as a model system. We will show that strong qualitative similarities exist for the two methods, making them complementary for the study of the electronic and transport properties of systems with defects.

The article is organized as follows: In Sect. Methodology we will introduce the theoretical methodology for both quantum transport and band structure unfolding calculations. In Sect. Results we will confront the results obtained by the two formalisms. Finally, in Sect. Discussion we will discuss our results.

Methodology

The basis for the calculation of both electronic structure and quantum transport is an appropriate definition of the electronic Hamiltonian. For the case of graphene, the nearest-neighbor tight-binding (TB) model is sufficient for the description of

the low-energy spectrum of the material, accounting only for the π atomic orbitals (as σ states are too far away from the Fermi level). Hence, the TB Hamiltonian can be written as

$$H = -t \sum_{\langle i,j \rangle, s} c_{i,s}^{\dagger} c_{j,s} + H.c., \tag{13.1}$$

where c_i ($c_i^{\dagger}$) is the annihilation (creation) operator for an electron with spin s at site i, and t is the hopping integral with a typical value $t = 2.7$ eV. We point out that the nearest-neighbor model gives a symmetric description of the valence and conduction bands of ideal graphene that can be only lifted with higher accuracy Hamiltonians, (e.g., by considering a second neighbor in the tight-binding scheme). Considering that graphene has two atoms per unit cell, the eigenstates can be approximated by a linear combination of the two atomic wavefunctions multiplied by a single phase factor, which denotes translation of the entire supercell in space [27]:

$$\Psi(\mathbf{k}, \mathbf{r}) = c_A(\mathbf{k}) \Psi_A(\mathbf{k}, \mathbf{r}) + c_B(\mathbf{k}) \Psi_B(\mathbf{k}, \mathbf{r}) =$$
$$= \frac{1}{\sqrt{\mathcal{N}}} \sum_j e^{i\mathbf{k} \cdot \mathbf{R}_j} \left[c_A(K) \phi(\mathbf{r} - \mathbf{R}_j^A) + c_B(\mathbf{k}) \phi(\mathbf{r} - \mathbf{R}_j^B) \right] \tag{13.2}$$

Here c_A and c_B are expansion coefficients of the $\phi(\mathbf{r})$ wave function of atomic orbitals A and B, respectively, and $\mathcal{N}$ is the number of elementary cells. The $\mathbf{R}_j = n\mathbf{a}_1 + m\mathbf{a}_2$ vectors specify the position of the graphene unit-cell, with $\mathbf{a}_1$ and $\mathbf{a}_2$ being the lattice vectors. Equation (13.2) derives directly from Bloch's theorem [27].

For the calculation of quantum transport, two-terminal graphene devices are considered, where a single graphene nanoribbon is contacted by two semi-infinite leads. As the objective of the study is to comprehend the internal scattering mechanisms due to the presence of defects, here we consider ideal contacts, i.e., contacts made of graphene with the same lateral width as the channel material. According to the non-equilibrium Green's function formalism [14], the single-particle retarded Green's function matrix can be written as

$$\mathcal{G}^r(\varepsilon) = [\varepsilon S - H - \Sigma_L - \Sigma_R]^{-1}, \tag{13.3}$$

where ε is the energy, H the real-space Hamiltonian, and S the overlap matrix, which is identical with the unitary matrix I in the case of an orthonormal basis set. Moreover, $\Sigma_{L,R}$ are self-energies that account for the effect of the contacts, calculated from the expression:

$$\Sigma_{L(R)} = \tau_{L(R)}^{\dagger} g_{L(R)} \tau_{L(R)}. \tag{13.4}$$

Here $\tau_{L,R}$ are interaction Hamiltonians that describe the coupling between the contacts and the device, while $g_{L,R}$ are the contact surface Green functions. The transmission coefficient of an incident Bloch state with energy ε can be thereon computed as the trace of the matrix product:

$$T(\varepsilon) = Tr\{\Gamma_L \mathscr{G}^r \Gamma_R [\mathscr{G}^r]^\dagger\}, \tag{13.5}$$

where

$$\Gamma_{L(R)} = i\{\Sigma_{L(R)} - [\Sigma_{L(R)}]^\dagger\} \tag{13.6}$$

are the spectral functions of the two contacts. According to the Landauer–Buttiker theory [14], conductance can be calculated as:

$$G = \frac{2e^2}{h} T, \tag{13.7}$$

where $G_0 = 2e^2/h \approx 77.5\,\mu\mathrm{S}$ is the conductance quantum.

For the calculation of the unfolded band structure, we consider the same graphene nanorribon used in the quantum transport scheme. Such ribbon can be viewed as an integer multiple of the ribbon's unit cell along the transport direction. Based on the TB Hamiltonian described above, the spectral weight of the unfolded bands on the unit-cell Brillouin zone can be defined as [26]:

$$w(\varepsilon, \mathbf{k}) = \frac{1}{\mathscr{N}} \sum_{a \in PC} \left(\sum_{j}^{\mathscr{N}} \{(c^{a+j}(\mathbf{k}))^* \cdot e^{-i\mathbf{k}\cdot\mathbf{R}_j}\} \right) \times \left(\sum_{j}^{\mathscr{N}} \{c^{a+j}(\mathbf{k}) \cdot e^{-i\mathbf{k}\cdot\mathbf{R}_j}\} \right), \tag{13.8}$$

where a are the atoms that belong to the primitive cell. The important aspect of this formula is that the sum of the expansion coefficients c should involve only equivalent atoms within the elementary cells that comprise the supercell.

We consider single carbon vacancies as the origin of the structural/electronic perturbation for our graphene system. Vacancies can be modeled within the nearest-neighbor TB scheme by removing a π electron, either by switching to infinite the related on-site energy term ε_i in the Hamiltonian, or equivalently, by switching to zero the hopping t_{ij} terms between the vacancy and the neighboring sites. However, a more accurate treatment of the states introduced by such defect within the electronic spectrum has to take into account the structural reconstruction around the vacancy. This can be achieved through a multiscale approach, where first principles calculations are performed initially and the TB Hamiltonian is parameterized consequently, on the basis of the *ab initio* results. In our implementation, we calibrate the TB parameters from density functional theory computations of defected graphene quantum dots [28]. The optimized values for the on-site energy of the vacancy and the hopping integral between this and its neighboring sites are $\varepsilon_i = 10\,\mathrm{eV}$ and $t_{ij} = 1.9\,\mathrm{eV}$, respectively.

Results

We consider a graphene nanoribbon with an armchair-type lateral confinement, having a width of $N_a = 38$ carbon dimer lines (≈ 4.5 nm) and a length of $N_z = 64$ zigzag chains ($l \approx 6.7$ nm), as seen in Fig. 13.1. By attaching two ideal semi-infinite contacts along the zigzag confinement (i.e., by considering the ribbon infinite along its length) we can calculate the quantum transport properties of this system using the non-equilibrium Green's function formalism. According to the TB scheme (where ribbons with $N_a = 3p + 2$ dimer lines are semi-metallic $\forall p \in \mathbb{N}$ and the rest are semiconducting), the $N_a = 38$ ribbon presents a semi-metallic character with a secondary band-gap of few meV. Figure 13.2 shows the ideal transmission coefficient and the respective density of states as a function of the energy (black lines). The conductance is characterized by the presence of integer plateaus that correspond to the number of available conduction channels (i.e., the number of 1D subbands) at a given energy. Conductance steps take place at energies where van Hove singularities appear in the density of states spectrum (see Fig. 13.2). If we now consider a single vacancy defect, there are two plausible configurations that give rise to different conduction characteristics. In the first case, if the vacancy is introduced at a $N_a = 3q$ position ($\forall q \in \mathbb{N}, q \leq p$), the defect states do not perturb the first conductance plateau around the Fermi level [13] and the transmission coefficient is equal with the ideal value (see results for V_1 in Fig. 13.2c). When the vacancy is introduced at a $N_a \neq 3q$ site (as V_2), the defect perturbation influences the first conductance plateau (Fig. 13.2b) and a dip in the transmission coefficient appears at the resonance of the defect state,
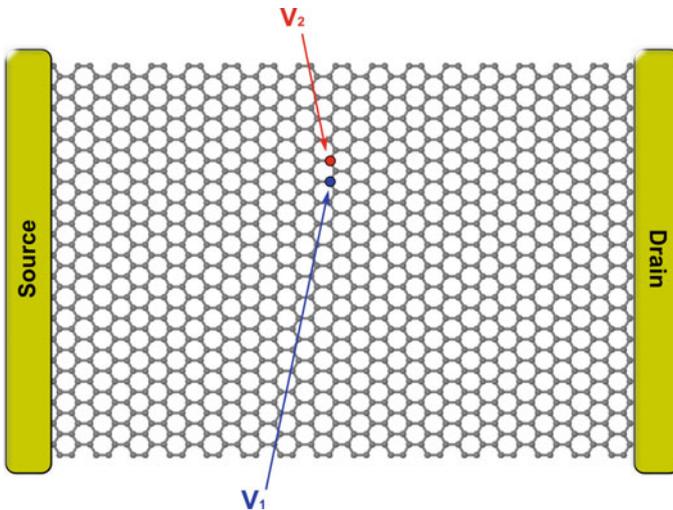
Fig. 13.1 Scheme of the armchair graphene nanoribbon with a width of 38 dimer lines and a length of 64 zigzag chains, used for the quantum transport and band structure calculations. A single vacancy is introduced in either the 27th (V_1) or the 29th (V_2) *dimer line*
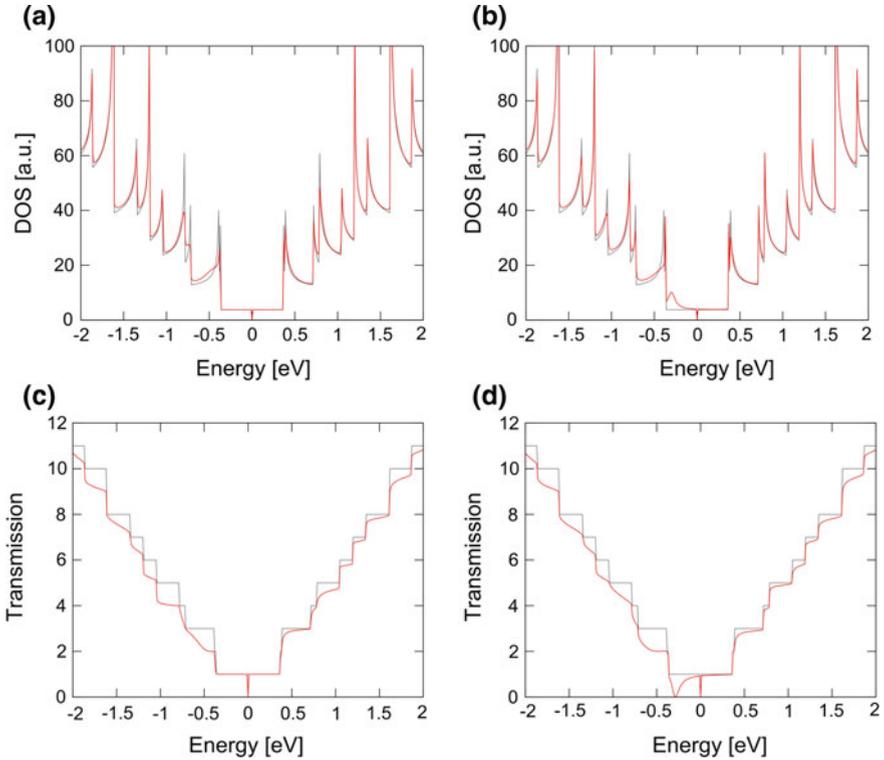
Fig. 13.2 **a** Density of states for the graphene nanoribbon with vacancy V_1. **b** Density of states for the graphene nanoribbon with vacancy V_2. **c** Transmission coefficient for the graphene nanoribbon with vacancy V_1. **d** Transmission coefficient for the graphene nanoribbon with vacancy V_2. *Grey lines* show the ideal values for the non-defected nanoribbon

which is ~ 0.4 eV below the Fermi level (Fig. 13.2d). It is important to point out that in both configurations the ~ -0.4 eV defect state is present, but perturbs the ideal ribbon wave function only in the second case [13]. Above the first plateau, both vacancy configurations show similar characteristics, with a small reduction of the transmission probability mainly for the valence band, due to the presence of a higher number of defect states bellow the Fermi level (in accordance with the acceptor-type character of reconstructed single vacancies in graphene [28]).

We now turn into the band structure calculations for the same graphene nanoribbon as above. Considering the armchair-type lateral confinement and periodicity along its length (*i.e.*, considering that the ribbon has no zig-zag confinement, in accordance with the quantum transport calculations), such ribbon can be described as a supercell with 16 repetitions of the unit cell (see Fig. 13.1). If we unfold the calculated band structure to the primitive Brillouin zone, we obtain the results shown in Fig. 13.3. For the non-defected ribbon, unfolding ideally recovers the unit-cell band structure, assigning a spectral weight value $w = 1$ for the original unit-cell bands, whereas
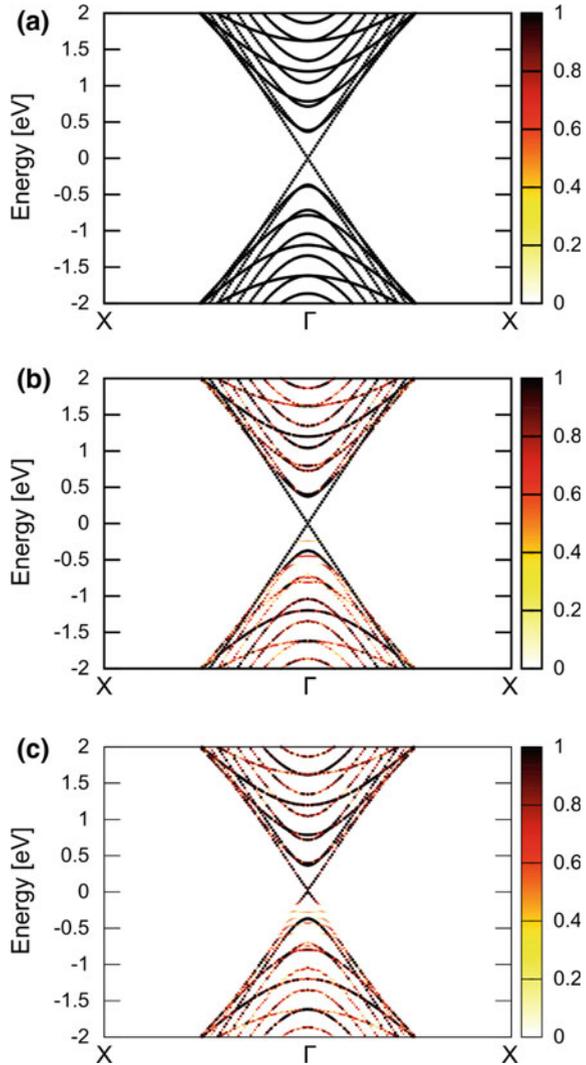
$w = 0$ for the extra bands that are present due to the additional supercell atoms. When performing the same calculation for the defected structures, the spectral weight values range between zero and one also for the original bands, in proportion to the perturbation of the wave functions induced by the defects. We note that for the V_1 vacancy (Fig. 13.2b), the bands around the Fermi level ($-0.45\,\text{eV} < E_F < 0.45\,\text{eV}$) maintain the ideal $w \approx 1$ value. On the contrary, the V_2 vacancy (Fig. 13.3c) shows $w \leq 1$ also for the low-energy spectrum, with a minimum $w \approx 0$ around $-0.4\,\text{eV}$ that corresponds to the resonant energy of the defect state. Similarly as in the transport calculations, for energies above and below the $-0.45\,\text{eV} < E_F < 0.45\,\text{eV}$ range,

both defects give rise to a reduction of the spectral weight value that depends both on the energy as well as on the wavenumber. Also here we note a bigger reduction trend for the valence than for the conduction band. However such reduction is stronger for the spectral weight with respect to the respective lowering of the transmission coefficient.

Based on the previous results, a critical analysis of the two computational methodologies shows strong similarities as well as some subtle differences. Band structure unfolding is a measure of the robustness of the intrinsic bands in the presence of structural/electronic perturbations. The spectral weight measures the resemblance between the unit-cell and the supercell wave functions. In this sense, each perturbative event automatically signals a reduction in the value of the spectral weight. On the other hand, the same perturbative events also give rise to the backscattering of the propagated electrons when calculating the quantum transport. The direct relationship between the cause (structural/electronic perturbation) and the result (diminishment of both the spectral weight and the transmission coefficient) is at the origin of the similarities for the two quantum mechanical methodologies. However, as noted above, a general trend shows that the reduction of the spectral weight in the unfolded band structure is stronger than the respective lowering of the transmission coefficient in quantum transport calculations. The most probable explanation of such quantitative misalignment reflects the fact that disordered states can still contribute in the propagation of electrons through hopping interactions between the perturbed states. This aspect should lower the negative impact of the perturbation even in the presence of moderate disorder. On the other hand, the spectral weight is a direct measure of the same disorder but does not account for tunneling or wave-function overlapping phenomena as the transport formalism does. Notwithstanding such difference, the two methodologies appear complementary, with band structure unfolding giving a robust interpretation of the motivations for electron back-scattering in quantum transport calculations.

Discussion

In this article we have compared the results obtained from two different computational methodologies that account for the effects of defects on the electronic and transport properties of graphene nanoribbons. The first one calculates the quantum transport properties within the non-equilibrium Green's function formalism, whereas the second one computes the spectral weight of the ribbon's unfolded band structure. Our analysis has evidenced clear qualitative similarities between the transmission coefficient of the propagated electrons and the spectral weight of the unfolded bands, as well as some quantitative divergences due to the intrinsic differences between the purely electronic and the quantum transport-related features. On the other hand, a strong complementary character has emerged, with the spectral weight giving a robust theoretical explanation for the origin of conductance degradation in the quantum transport calculations of defected/disordered systems.

References

1. A.K. Geim, K.S. Novoselov, Nat. Mater. **6**(3), 183 (2007). DOI 10.1038/nmat1849
2. A.S. Mayorov, R.V. Gorbachev, S.V. Morozov, L. Britnell, R. Jalil, L.A. Ponomarenko, P. Blake, K.S. Novoselov, K. Watanabe, T. Taniguchi, A.K. Geim, Nano Lett. **11**(6), 2396 (2011). DOI 10.1021/nl200758b
3. X. Du, I. Skachko, A. Barker, E.Y. Andrei, Nat. Nanotech. **3**(8), 491 (2008). DOI 10.1038/nnano.2008.199
4. J. Baringhaus, M. Ruan, F. Edler, A. Tejeda, M. Sicot, A.T. Ibrahimi, Z. Jiang, E. Conrad, C. Berger, C. Tegenkamp, W.A. de Heer, Nature **56**, 349 (2014). DOI 10.1038/nature12952
5. L. Banszerus, M. Schmitz, S. Engels, M. Goldsche, K. Watanabe, T. Taniguchi, B. Beschoten, C. Stampfer, Nano Lett. **16**(2), 1387 (2016). DOI 10.1021/acs.nanolett.5b04840
6. Y. Zhang, Y. Tan, H.L. Stormer, P. Kim, Nature **438**(7065), 201 (2005). DOI 10.1038/nature04235
7. K.S. Novoselov, Z. Jiang, Y. Zhang, S.V. Morozov, H.L. Stormer, U. Zeitler, J.C. Maan, G.S. Boebinger, P. Kim, A.K. Geim, Science **315**(5817), 1379 (2007). DOI 10.1126/science.1137201
8. K.S. Novoselov, E. McCann, S. Morozov, V. Falko, M.I. Katsnelson, U. Zeitler, D. Jiang, F. Schedin, A.K. Geim, Nat. Phys. **2**, 177 (2006). DOI 10.1038/nphys245
9. C.L. Kane, E.J. Mele, Phys. Rev. Lett. **95**, 226801 (2005). DOI 10.1103/PhysRevLett.95.226801
10. V.P. Gusynin, S.G. Sharapov, Phys. Rev. Lett. **95**, 146801 (2005). DOI 10.1103/PhysRevLett.95.146801
11. Y.M. Lin, V. Perebeinos, Z. Chen, P. Avouris, Phys. Rev. B **78**, 161409 (2008). DOI 10.1103/PhysRevB.78.161409
12. H. González-Herrero, J.M. Gómez-Rodríguez, P. Mallet, M. Moaied, J.J. Palacios, C. Salgado, M.M. Ugeda, J.Y. Veuillen, F. Yndurain, I. Brihuega, Science **352**(6284), 437 (2016). DOI 10.1126/science.aad8038
13. I. Deretzis, G. Fiori, G. Iannaccone, A. La Magna, Phys. Rev. B **81**, 085427 (2010). DOI 10.1103/PhysRevB.81.085427
14. S. Datta, *Electronic transport in mesoscopic systems* (Cambridge University Press, Cambridge, 1997)
15. A. La Magna, I. Deretzis, G. Forte, R. Pucci, Phys. Rev. B **80**, 195413 (2009). DOI 10.1103/PhysRevB.80.195413
16. I. Deretzis, G. Fiori, G. Iannaccone, A. La Magna, Phys. Rev. B **82**, 161413 (2010). DOI 10.1103/PhysRevB.82.161413
17. F. Giannazzo, I. Deretzis, A. La Magna, F. Roccaforte, R. Yakimova, Phys. Rev. B **86**, 235422 (2012). DOI 10.1103/PhysRevB.86.235422
18. W. Ku, T. Berlijn, C.C. Lee, Phys. Rev. Lett. **104**, 216401 (2010). DOI 10.1103/PhysRevLett.104.216401
19. V. Popescu, A. Zunger, Phys. Rev. Lett. **104**, 236403 (2010). DOI 10.1103/PhysRevLett.104.236403
20. V. Popescu, A. Zunger, Phys. Rev. B **85**, 085201 (2012). DOI 10.1103/PhysRevB.85.085201
21. P.V.C. Medeiros, S. Stafström, J. Björk, Phys. Rev. B **89**, 041407 (2014). DOI 10.1103/PhysRevB.89.041407
22. T.B. Boykin, G. Klimeck, Phys. Rev. B **71**, 115215 (2005). DOI 10.1103/PhysRevB.71.115215
23. P.B. Allen, T. Berlijn, D.A. Casavant, J.M. Soler, Phys. Rev. B **87**, 085322 (2013). DOI 10.1103/PhysRevB.87.085322

24. C.C. Lee, Y. Yamada-Takamura, T. Ozaki, J. Phys.: Cond. Matter **25**(34), 345501 (2013). DOI 10.1088/0953-8984/25/34/345501
25. M.W. Haverkort, I.S. Elfimov, G.A. Sawatzky, Electronic structure and self energies of randomly substituted solids using density functional theory and model calculations (2011). Preprint arXiv:1109.4036 [cond-mat.mtrl-sci]
26. I. Deretzis, G. Calogero, G.G.N. Angilella, A. La Magna, Europhys. Lett. **107**(2), 27006 (2014). DOI 10.1209/0295-5075/107/27006
27. C. Bena, G. Montambaux, New J. Phys. **11**(9), 095003 (2009). DOI 10.1088/1367-2630/11/9/095003
28. I. Deretzis, G. Forte, A. Grassi, A. La Magna, G. Piccitto, R. Pucci, J. Phys.: Cond. Matter **22**(9), 095504 (2010). DOI 10.1088/0953-8984/22/9/095504

Chapter 14
Defect-Induced Magnetism in Graphene: An Ab Initio Study

A. Pidatella and R. Mazzarello

Abstract Graphene is an amazing two-dimensional system with exceptional physical and chemical properties. Potential applications in quantum information processing have been proposed for C-based materials, in particular for graphene system, where electron spin is a promising candidate for a solid-state qubit. The preservation of long spin coherence time is the fundamental feature to get for efficient working spin-qubit system. Despite graphene environment seems to suit the goal, defects in the structure, interactions with impurities and edge states can be a source of alteration of quantum information, since they could enhance the decoherence effects. The present work is a computational analysis of defective systems. It focuses on the investigations of various prototypical defect states (vacancies) and impurities interacting with graphene surface (hydrogen, boron, nitrogen, and oxygen) by means of density functional theory (DFT). We provide a preliminary study about the effects of these interactions. Vacancy-type defects give rise to a breaking of graphene symmetry, promoting a localized state with a magnetic moment whose magnitude is concentration-dependent. Hydrogen promotes a local hybridization of the structure, providing a localized magnetic moment and giving rise to an enhancement of spin–orbit interaction of about three orders of magnitude, showing the impact of hydrogen on spin-relaxation time. Among boron, nitrogen, and oxygen, the work has shown that the only one which returns a magnetic ground state is nitrogen. Boron provides an n-doping of defective-graphene. Oxygen leads to a hybridization of carbon atoms bonding, but its electronic structure does not allow a magnetic system. In the particular case of a bridge-like adsorption site. Among the different configurations for the adsorption sites, the bridge-site is energetically the most stable one, showing as in

A. Pidatella (✉)
Institut für Festkörperphysik, Technische Universität Dresden,
Zellescherweg, 17, 01069 Dresden, Germany
e-mail: angelo.pidatella@tu-dresden.de

R. Mazzarello
Institut für Theoretische Festkörperphysik, RWTH Aachen University,
Otto-Blumenthalstraße, 52074 Aachen, Germany
e-mail: mazzarello@physik.rwth-aachen.de

the other configurations for nitrogen, a magnetic system. Nitrogen adatoms develop a magnetic order (at zero temperature) which is always ferromagnetic independently from the distance between two adjacent nitrogen atoms.

A Bird's-Eye View at Graphene

Among systems with only carbon atoms, *graphene*—a two dimensional (2D) allotrope of carbon—plays an important role since it is the basis for the understanding of the properties of other carbon allotropes. Graphene is the name given to single 2-dimensional sheet of graphite which was isolated experimentally for the first time in 2004 [1]. Before that, graphene had been investigated theoretically for over sixty years, although its existence as a 2D crystal in free space was thought impossible [2–4]. Since the isolation of a single-layer of graphene and the demonstration of its excellent conductivity and optical properties, the research aiming at determining the electronic properties and potential applications of the same, progressed at a rapid pace. In the last ten years, graphene has been the most studied material in the world, catching the attention of the scientific community for its fascinating physical and chemical properties. Because graphene is made of only carbon atoms, pure graphene is free of *nuclear spins* and should be an attractive material for *electron-spin* based quantum circuits. However, carbon atom has no magnetic moment, hence realizing magnetism in graphene is challenging. The first reason is the absence of d and f shell electrons which are responsible for the magnetic coupling in conventional ferromagnets, so that ideal graphene is in itself *nonmagnetic*. Despite this fact, both theory and experiments suggest that a magnetic order can exist in these carbon structures under particular circumstances, related to the presence of impurities or defects in the pristine structure. Magnetic moments originate from *localized states* in the presence of point defects in the graphene lattice. One example of this kind is missing C atoms, called *vacancies*, even if the mechanism at fundamental of the presence of magnetic moments could be also linked to the interactions between graphene layer and other elements (e.g., metals) moving on the surface as *adatoms* [5]. In the light of these findings, the introduction of defects in graphene becomes a tool to investigate the presence of relevant magnetic phenomena.

Defects in nanographites can be created intentionally by irradiating the material with electrons or ions. By manipulating the conditions of irradiation, it is possible to tune, in a flexible way, the properties of the carbon-based materials. Using these techniques carbon atom can be removed from the π conjugation network of the graphene sheet. The single-atom defects on the graphene lattice give rise to quasilocalized states at the Fermi level. As known, graphene lattice is a *bipartite lattice*, so that when a defect is created in the A sublattice, only the p_z orbitals of C atoms in the B sublattice contribute to the quasilocalized state, and vice versa. These states extend over several nanometers around the defect and the fact that these quasilocalized states lie at the Fermi energy suggests that Stoner magnetism can be induced by the electron exchange instability [5].

Overview on Graphene for Quantum Information

Some potential applications in *quantum information* processing have been proposed for graphene systems, where the electron spin is a very promising candidate for a solid-state *qubit* [6]. Alternative devices, during the years, were introduced, and great expectations were posed about semiconductor heterostructures exploiting GaAs technology for forming quantum-dot- based devices. One of the most important parameter to control to ensure the characteristics required for the quantum information, is the *spin coherence.* The preservation of this feature is important with the aim to obtain efficient and working systems. However, many sources of spin decoherence have been identified, among these, the major ones are for sure the *spin–orbit interaction,* coupling the electron's spin to the magnetic field generated by the electron's orbit and the *hyperfine interaction* of the electron spin with the surrounding nuclear spins [7]. Therefore, it would be good to form qubits in quantum dots based on other materials, where spin-orbit coupling and hyperfine interaction are considerably weaker. Carbon-based materials are good candidates for many reasons. First, spin-orbit coupling is weak in carbon due to its relatively low atomic weight, second, natural C atoms consist predominantly of the zero-spin isotope ^{12}C, for which the hyperfine interaction is absent. Graphene quantum dot, in addition, seems to present the perfect environment for developing this kind of device. Spin qubits in graphene can not only be coupled (via Heisenberg exchange) between nearest-neighbor quantum dots but also over long distances, giving rise to a system ideal for fault-tolerant quantum computation and offering a low error rate due to the weak decoherence, in combination with a high error threshold due to the possibility of long-range coupling. Despite all these fascinating features offered by graphene, the presence of *edges* and *defects* can drastically perturb the spin-coherence and spin-relaxation times. Furthermore, the presence of adatoms on graphene surface has been shown to enhance the spin-orbit coupling [8]. Edges and defects, such as vacancy, and hydrogen adatoms lead to the formation of localized electronic states [5]. Localized states give rise to magnetic moment, due to the presence of a midgap state, occupied by just one electron, thus providing an unpaired spin. One could think these additional spins as pseudo-nuclear spins, as they are expected to interact with the qubit spin in the quantum dot. It is expected that the central spin-defect spin interaction is much stronger than for nuclear spins and that defect spins interact with each other too, so in the end they could represent the dominant source of spin relaxation.

The absence of a corresponding theoretical study of interactions between defect states and defect-induced magnetic impurities within the context of graphene qubits urges to a better understanding of these processes. This work aims at a better understanding of the decoherence effects in graphene spin qubits caused by the presence of localized states at the Fermi energy which inevitably interact with the latter ones. In the following the results on the investigations of various prototypical defect states by means of the density functional theory (DFT) are shortly reported. These can be considered as a preparatory work for a proactive target concerning the development of realistic tight-binding models which describe defects and interactions between them.

A greater understanding will be achieved once one derives low-energy theories from the tight-binding models and by solving them, in order to establish graphene qubits lifetimes. However, in this framework only the significant role that defects play is shown.

Why it is important to control spin coherence *for developing adequate system for quantum information processing?*

Only if sufficiently long computations can be performed, quantum algorithms, which exploit the special properties of a quantum computer, can be implemented successfully. This includes the storage of quantum information as well as the number of performable operations on the quantum bits. By exploiting two fundamental principles of quantum mechanics, namely *superposition* and *entanglement*, a quantum computer is able to solve specific problems with much higher efficiency than a classical device. In a quantum computer, all information is stored in qubits which are quantum mechanical two-level systems. Two complex numbers can be stored in one single qubit. A set of N qubits is initialized in linear time and has 2^N basis states due to the superposition principle. In addition, a transformation can be applied to all qubits at the same time which saves 2^N steps compared to an individual application, feature which is often discussed under the keyword *quantum parallelism.*

The requirements for an implementations of a quantum computer are summarized by the famous DiVincenzo's criteria [9]:

1. Scalability and well-defined qubits.
2. Well-defined initialization of the qubits in a simple state.
3. Long coherence times.
4. A universal set of quantum gates.
5. Measurement of selected qubits.

According to these criteria, quantum dots with electron spins as qubits have a great potential for the realization of a quantum computer. The spin of an electron confined inside the dot is a *well-defined qubit.* Besides the identification of a single spin as qubit, it is also possible to define a qubit in the basis of two or three spins. Solid-systems are *scalable* by construction, but the implementation of systems with a large number of qubits could certainly be challenging. The *initialization* of a lateral quantum dot is achievable by relaxation or by populating the dots with selected spin states. In addition, these system have relatively *long coherence times.*

Anyway, only if the coherence time of the qubit is sufficiently long, information can be stored and an adequate number of computations can be executed. Concerning the success of QIP, a fundamental concept is provided in *quantum threshold theorem* (or *quantum fault-tolerance theorem*): if the average error rate (noise) of the quantum gates is kept below a critical value, arbitrary long computations will be possible due to quantum error correction, i.e., a quantum computer with noise can quickly and accurately simulates an ideal quantum computer, provided the level of noise is below the threshold value. This theorem implies that the error in quantum computers can be controlled as the number of qubits scales up. Thus a lot of effort is put into the development of quantum error correction. Alternatively, one can eliminate or

reduce the sources of errors, aspect which implies a thorough investigation of the decoherence effect induced by external influences.

For a single electron confined in a quantum dot, the hyperfine interaction between the electron spin and the nuclear spins, together with spin-orbit coupling, seem to be the dominating sources of decoherence of the electron spin. By spin–orbit interaction, the spin couples to the electronic degrees of freedom which are exposed to various perturbing effects such as *impurity scattering.* Previous works have proven how if spin-orbit coupling *alone* is the dominating mechanism for the decoherence, a single electron spin confined in a quantum dot would exhibits very long dephasing times, order of ms. However, experimental observations contradict with this theoretical prediction. Measurements revealed strongly reduced dephasing times compared to the relaxation time, thus spin-orbit coupling alone cannot be the major source of decoherence for these systems. In general, it seems that decoherence is dominated by the Fermi contact hyperfine interaction between the electron spin and the nuclear spins in the dot. Although in general in quantum dot system based on semiconducting heterostructures, hyperfine interaction seems to dominate decoherence effect, overcoming the spin–orbit interaction, in graphene quantum dot the hyperfine interaction is *absent*, due to the presence of zero-nuclear spin C atoms.

What Is a Graphene Quantum Dot?

Two fundamental problems need to be overcome before graphene can be used to form spin qubits. The first one, is the difficulty in tuning the quantum dot due to the absence of a gap in the spectrum. The phenomenon of *Klein tunneling* makes it hard to realize a good confinement of particles. Furthermore, owing to the valley degeneracy that exists in graphene, it is not trivial to form two-qubits gates using the Heisenberg exchange coupling for spins in tunnel-coupled dots. These problems can be solved by defining an appropriate system, i.e., in order to define a quantum dot (QD) one can use the confinement provided by the transverse states in a graphene nanoribbon [10] (as shown in Fig. 14.1). In such a device the valley degeneracy is lifted, giving way to the appearance of Heisenberg exchange coupling for spins in tunnel-coupled dots, enabling one to use graphene QDs for spin qubits. Moreover, the coupling makes use of conduction-band to valence-band tunneling processes, giving rise to an interaction at long distance, mechanism mainly based on the Klein paradox in graphene. By solving the Dirac equation for the system it is possible to establish how the boundary conditions impose a set of solutions which gives rise to a confined structure, the QD, and to define the relation between the values of the gate voltage and barrier voltage in order to shift appropriately the energy levels with the final aim to create a resonance between them and thus a tunnel-connection (for further details, see [10]).

It is interesting to see if the spins S_i of these two QDs electrons ($i = 1, 2$) are coupled through an exchange coupling, $H_{exch} = J S_1 \cdot S_2$, in the same way as for regular semiconductor QDs, because this coupling is, in combination with
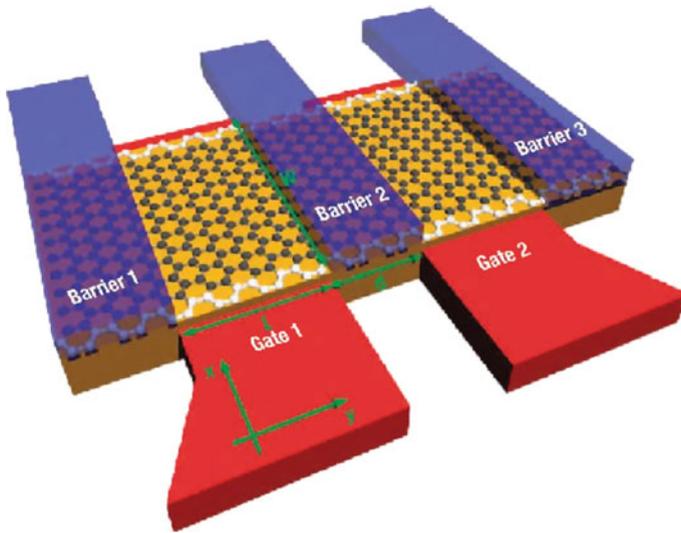
Fig. 14.1 Schematic diagram of a graphene double quantum dot. Dots with length L and width W are based on ribbon of graphene *(grey)* cut, in this case, with armchair edges *(white)*. Particles are confined by tuning the voltages applied to the 'barrier' gates *(blue)* along the y direction, whereas the confinement along x is achieved by means of edges. Additional gates *(red)* allow to shift the energy levels of the dots [10] (color figure online)

single-spin rotations, sufficient to generate all quantum gates required for universal quantum computation [6]. For the case with more than two dots in a line, it turns out that we can couple any two of them with the others being decoupled by *detuning*. Figure 14.2 illustrates the situation of three dots in a line, where the left one and the
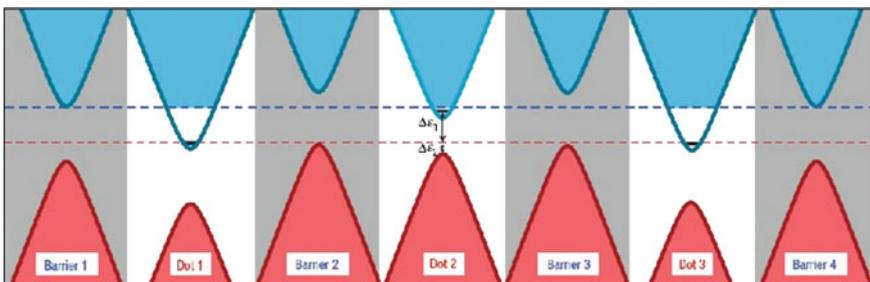
Fig. 14.2 Triple-quantum-dot setup. Illustration of the energy bands of a triple-quantum-dot setup in which dot 1 is coupled to dot 3 via co-tunneling processes through the valence bands of barrier 2, barrier 3, and dot 2. The central dot is decoupled by detuning. The triple-dot example shows that in a line of QDs, it is possible to strongly couple any two of them and decouple the others simply by detuning, which is a unique feature of graphene and cannot be achieved in semiconductors, such as GaAs, due to its larger gap [10]

right one are strongly coupled whereas the central dot is decoupled by detuning. The tunnel coupling between dot 1 and dot 3 is achieved via Klein tunneling through the valence band of the two central barriers and the valence band of the central dot. The most important fact is that the long-distance exchange coupling between 1 and 3 is obtained via the valence band and not via the qubit level of the central dot, leaving the qubit state of dot 2 unchanged, in other words, the qubit level in dot 2 is barely used to couple dot 1 and dot 3. This is a unique feature of graphene QDs due to the small and highly symmetric bandgap, which is not known to exist in other semiconducting materials. The availability of non-local interactions is important in the context of quantum error correction, as it raises the error threshold for fault-tolerant quantum computation.

About the coherence time for graphene spin qubits, in the light of the only graphene properties, one has clearly promising expectations which seem to ensure long-lasting coherence times. However, as said previously, it is important to recognize the role played by the defects in these structures to make stronger the model for realizing it experimentally. It is this the right starting point for our investigations.

Defective-Graphene in a Nutshell

The analysis has been focused on the energetic, structural, electronic, and magnetic properties of bulk defects (vacancies and adatoms). Considering vacancy defect and several types of adatoms relevant for recent experiments carried out on these systems [5, 11–13], including boron, nitrogen, and oxygen adatoms, it has provided a large class of samples making in some way this study eventually extendable to other systems, always with some precautions. To establish the robustness of calculations, some preliminary simulations were needed to determine the most appropriate DFT functional. Although a systematic study of the energetics of adatoms on graphene and on their migration energies has been recently obtained using LDA functionals [14], this work does not provide a thorough picture of the electronic structure of the systems, nor discusses the limitations of the LDA approximation. Moreover, they only consider *non-magnetic* configurations. Then, one could think that the use of a GGA functional could return improved results, but in principle both functionals are equipped with advantages and disadvantages, dependent on the physical context under analysis. Generally, when not specified, we have chosen to perform calculations using GGA functional of Perdew–Burke–Ernzerhof (PBE) [15], according to certain reliable results from other works [5, 11, 16] and to some of our simulations on the same structures with different functionals, which have been a proof of the strength of GGA approximation for these systems. All the ab initio simulations have been performed using Quantum Espresso (QE) [17], a parallelized DFT code developed at SISSA, Trieste (Italy).

The general effect of this induced magnetic ground state is related to a possible coupling between the spin-defect and the central-spin of the QD, worsening and destroying the coherence. My investigation led on different possible sources of
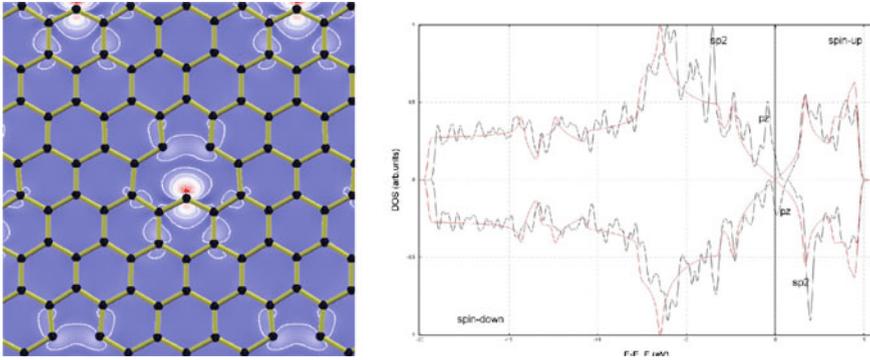
Fig. 14.3 Vacancy-type defect spin-polarized magnetization density and density of states (DOS). *Left* projection on xy-plane around the vacancy defect. Isolines are representative of the magnetic influence of the defect on the odd-nearest neighbor atoms, whereas even-nearest neighbors are anti-aligned. *Right* spin-resolved total density of states plot. In the figure, the *red line* shows the density of states of the ideal graphene, whereas the *black* one represents the DOS of the defective graphene. Labels indicate the character of the defect states. For the single-vacancy, together with the quasilocalized states generated by p electrons and reported as p_z orbital contribution, other localized states are generated by the presence of σ dangling bonds, which contribute as sp^2 hybridized orbital (color figure online)

decoherence has provided that the simplest point defect in the solid state matter, vacancy, breaks down the particle-hole symmetry of the graphene system, giving rise to localized states which provide a magnetic moment different from zero (Fig. 14.3). The total magnetic moment for these systems is determined by the contribution ($1\ \mu_B$) of the localized sp^2 dangling bond state and the fractional contribution ($<1\ \mu_B$) of the *extended defect state* $\psi_d(r)$ localized close to the Fermi energy. The partial spin polarization of $\psi_d(r)$ could be explained in terms of *self-doping* (charge transfer from the bulk to $\psi_d(r)$, corresponding to p-doping of the pristine graphene, which arises from the stabilization of the defect state). It has a magnitude depending on the defect concentration, the greater the concentration, the lower the magnitude, reaching a saturation value of 1.5 μ_B per cell. Another interesting feature is related to the distributions of the electron spin magnetization density in the vicinity of the defect, which shows a breaking of the characteristic $\sqrt{3} \times \sqrt{3}$ pattern with an enhancement of the localized magnetic moment associated with the dangling bond of the unsaturated C atom. For this type of defect, as we said, the threefold symmetry (manifested for example in the hydrogenated samples) is broken due to the Jahn–Teller distortion and this is exhibited also in the spin magnetization density, from which one can also evaluate a typical ferromagnetic configuration for atoms belonging to the same sublattice, whereas atoms which belong to different sublattices will have opposite spin projection, up (down). According to the Stoner scheme, the magnetic ordering is driven by the *exchange energy* $E_x \sim -\sum_i M_j^2$, where M_j is the magnetization of the p_z orbital of the jth carbon atom. It is possible to show that ferromagnetic ordering is the only possibility for the magnetism originating from
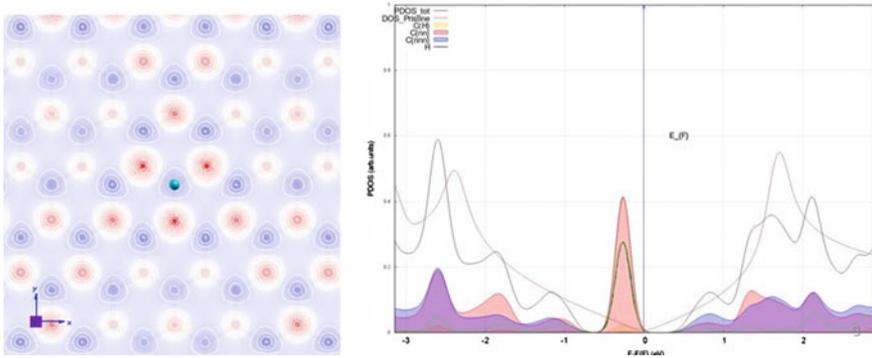
Fig. 14.4 Graphene + H on top adsorption site, spin-resolved magnetization density, and projected-DOS. *Left* projection on the xy plane. The effect induced by the hybridization does not affect the characteristic $\sqrt{3} \times \sqrt{3}$ pattern, respecting the threefold symmetry of the system. Magnetic moment is mainly distributed around the H atom site, isolines show the influence of the adatom on the magnetic alignment of the C-spins (*red* indicates spin-up density, blue denotes spin-down density). Only the C atoms which occupies the sites belonging to the complementary sublattice of the defect state are oriented in parallel with H spin orientation, the other ones are anti-aligned. *Right* PDOS of system. *Black line* denotes the total PDOS; *golden*, *red*, and *blue shaded* regions indicate the contribution to DOS due to C atom below H, first neighbor and second-neighbor C atoms, respectively. *Green line* represents the contribution of H atom. Where the first-neighbors are strongly hybridized by the H atom, clearly visible in the perfect agreement and overlap among these states, H atom does not influence the second-neighbors. The sp^3 hybridization induced by H provides a localized state at the Fermi energy E_F, which is spin-up component is totally occupied, whereas the spin-down component (not reported in the figure) is above the E_F and hence, not occupied (color figure online)

quasilocalized states induced by defects in the *same* sublattice because of the *non-oscillating* behavior of both (i) M_j within the same sublattice and (ii) the indirect (RKKY) coupling due to the semimetallic properties of graphene [18, 19]. On the contrary, it is possible to see that for the case in which defect states belong to different sublattice, E_x is minimized when the coupling is *antiferromagnetic*, due to the indirect spin-polarization effect. The strength of the coupling between the defect-induced magnetic moments located in different sublattices depends on the defect concentration since $E_x \sim -\sum_i M_j^2$ and the contribution of the magnetic moment associated with a single defect is $\sum_i M_j^2 \sim \sum_j |\psi_d(r)|^4 \sim \log^{-2}(N)$, where $1/N$ is the defect concentration [5].

Hydrogen interaction with graphene surface (physisorption on specific sites) produces similar effects, hybridizing locally the system and giving rise to a magnetic moment (Fig. 14.4). The H adsorption causes the appearance of peak in the DOS at the Fermi level which spin-splits due to electron–electron interactions. Remarkably, this result is compatible with Lieb's theorem for the Hubbard model on bipartite lattices. According to this theorem, the removal of a single site in the bipartite lattice gives rise to a ground state with $S = 1/2$. The covalent bond between the H atom and the C atom underneath effectively suppresses the "site" (the p_z orbital, where the π

network conjugation is broken due to the sp^3-hybridized C, establishing a localized state with an unpaired electron, consequently subjected to a spin polarization), creating a vacancy in the underlying low-energy Hamiltonian. It is worth noticing how this result contrasts with that obtained for a vacancy. Vacancies could in principle give rise to similar magnetic states. The difference with respect to the case of H adsorption is that vacancies tend to reconstruct and the magnetic moment generated can actually vanish depending on the concentrations. Furthermore, H enhances the spin-orbit coupling of about three orders of magnitude with respect to the intrinsic spin–orbit interaction. Adatom locally breaks the reflection symmetry across the graphene plane leading to an out-of-plane distortion by an angle ϕ relative to the plane. For $\phi \neq 0$, the distortion mixes σ and π orbitals that are no longer orthogonal (for full sp^3 these states are degenerate) and hence the SO interaction becomes a first-order effect leading to a large enhancement of SOC for covalently bonded hydrogen impurities in graphene. From the theoretical point of view, it is possible to demonstrate that while the general reason to this enhancement of the SOC is related to sp^3 hybridization, specifically, the main enhancement appears due to pseudospin inversion asymmetry (PIA), which generates novel spin-orbit coupling for conduction electrons [8]. The SOC Hamiltonian can be derived by inspecting the reduction of the graphene point group symmetry D_{6h}—which allows for the intrinsic spin-orbit coupling only—to the one corresponding to adatoms-graphene system. First, the C $-$ H covalent bonds break the space inversion symmetry and the point group reduces to C_{6v}. Second, the hydrogenated carbons on sublattice α cannot be interchanged with the nonhydrogenated carbons on sublattice β. This breaks the pseudospin inversion symmetry, and $C_{6v} \rightarrow C_{3v}$ (further details can be found in Ref. [8]). Because of this enhancement, we are interested to evaluate if SOC can perturb our system. I have taken account of the breaking of the space inversion symmetry due to sp^3 hybridization, in order to establish if energetically SOC gives rise to important differences in the magnetic anisotropy energies of the ground state. I got that there are for sure energetic differences due to the magnetic anisotropy of system, but they are very small (order of about μeV), so that SOC is not able to remove the ground-state degeneracy, at least at finite temperature. Studying the interactions between two hydrogen atoms on the same supercell (playing with different configurations: same sublattices $\alpha\alpha$ or different sublattices $\alpha\beta$ sites for positioning H), one can see the formation of a ferromagnetic order when adatoms are located on the same sublattice, regardless of the distance between them, whereas putting these on different sublattices, according to Lieb's theorem [20] derived for a half-filled single-band Hubbard model, the most stable magnetic order is the antiferromagnetic one, appreciable only if the distance between the adatoms is sufficient to avoid hybridization mechanisms which couple the unpaired electrons belonging to the two independent impurities (≥ 16 Å). All these results are in agreement with previous works in the literature.

First-Principles Study of Adatom Adsorption on Graphene

Etched graphene quantum dots on hexagonal boron nitride (h-BN) are currently investigated experimentally in many works in order to assess their suitability as spin qubits. Stampfer et al. [13] have recently shown that these systems display very stable single-dot characteristics (more stable than those of graphene nanostructures deposited on SiO_2) due to the low concentration of substrate-induced defects and the small strain thanks to the lattice compatibility with the substrate. Investigations on adatoms element such boron B, nitrogen N, and oxygen O are for sure preparatory for a better understanding of the interaction and effects that presence of these elements could generate on the system.

Let us try to summarize the results obtained. About the elements studied, and their interactions with graphene monolayer system, in general we can say that the adsorption energy evaluated for all the elements is always not less than 1 eV, which means that boron, nitrogen, and oxygen in atomic form have a relatively strong stability when they sit on the bridge site. The presence of adatoms onto graphene can functionalize graphene in different ways, depending on the character and type of interaction. Whereas for the boron B adatom has been shown that the connection with graphene surface is defined by means of physisorption, which bonds electrostatically the adatom to the surface, but without great deformation of the latter, in the nitrogen N and oxygen O cases, both bonds with C atoms are covalent bonds, defining a chemisorption of the latter ones, inducing a new status of hybridization of the system, breaking in some way the ordinary scheme for graphene lattice.

The character of interaction is also related to the type of bonding, in fact among the three elements the only one which does not show a covalent interaction, is B, that establishes a ionic bonding, with a subsequent charge transfer from adatom to graphene. In B case, this transfer leads to n-doping of graphene, changing the band structure of the system. Moreover, the transfer of electronic charge to graphene implies the loss of a possible magnetization mechanism related to spin-polarizability of the localized p-state near the Fermi energy of system, so that the B has a nonmagnetic ground-state. Boron is characterized by 5 electrons, four of them are in the inner s shells ($1s$ and $2s$) and the remaining one belongs to $2p$ orbital. Figure 14.5 shows the PDOS for B on the bridge-site. One can see that the Dirac point of the boron-graphene system (labeled with E_D) is shifted at lower energy value with respect to Fermi energy (which is set at zero energy). In other words, the Fermi level is shifted higher in energy relative to E_D, reflecting a partial occupation of graphene conduction-band states. The $2p$ state of the B lies 1 eV above E_F, where a small part of this state is below the Fermi energy and showing a minimal interaction between graphene states and boron state.

Although, the B case may look similar to that of H (since just one electron is unpaired in the outer electronic shell), the type of interaction is extremely different. Since H has just one available electron in the $1s$ shell, one cannot think that there could be a charge transfer from H to graphene in order to provide an ionic bonding, independently by the geometrical features, because physically it would be
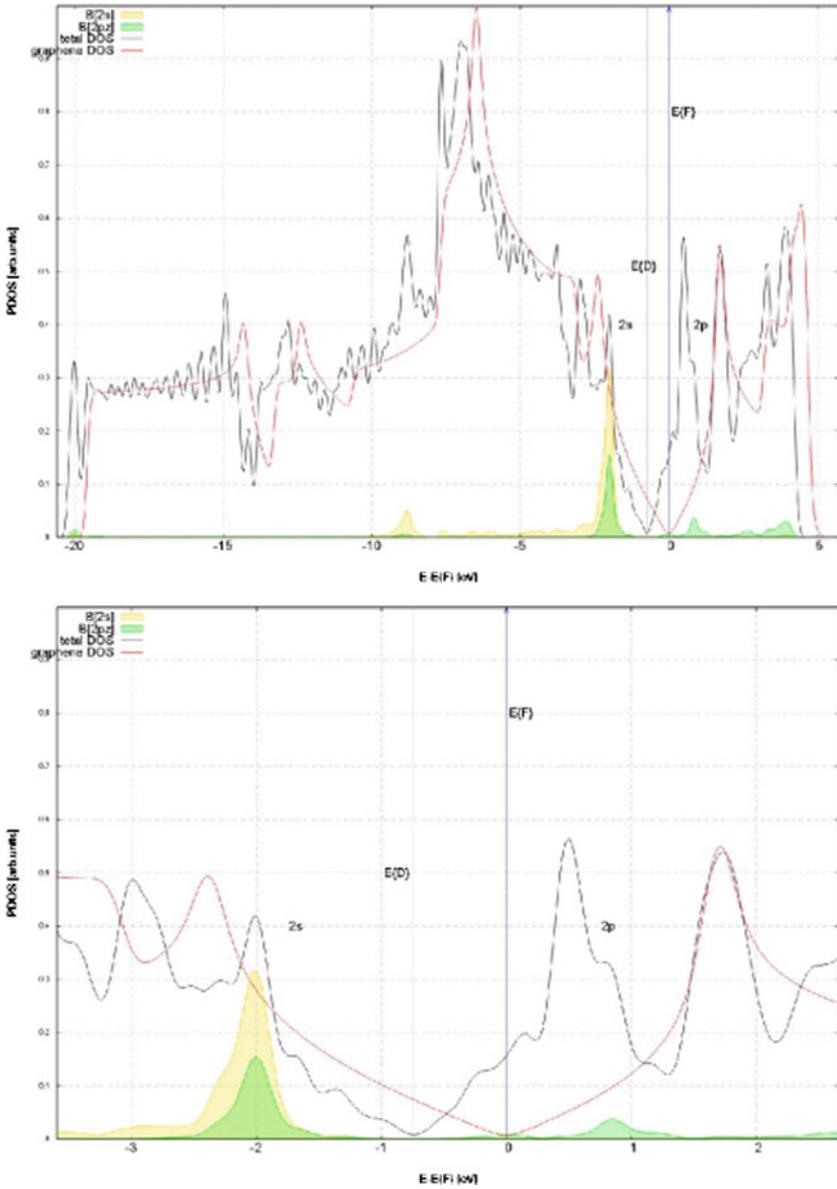
Fig. 14.5 Spin-up-resolved total projected-DOS for graphene monolayer with single-B atom on bridge-site. *Top* PDOS of system. *Black line* denotes the total DOS; *golden* and *green-shaded* region indicate the contribution to PDOS due to orbitals $2s$ and $2p_z$, respectively, of B atom. *Red line* represents the contribution of pristine graphene DOS. It shows the Dirac point of the defective system, labeled in the plot with $E\{D\}$, and the Fermi energy, $E\{F\}$ at zero energy. Although it is clear a hybridization between graphene states and states from $2s$ orbital of B at about -2.5 eV and -9 eV, these states do not play a relevant role in magnetism, which is in principle due to the localized peak of $2p_z$ orbital, in this case above the Fermi energy, but partially occupied, showing a n-doping provided by B to graphene. *Bottom* zoom-in of the same PDOS (color figure online)

impossible for H to release its only one electron to graphene. Hence, the most stable solution is forming a covalent bond. In the B case ionic bonding seems to be the best solution, this also explains the small distance of B from the graphene sheet, to reduce the electrostatic energy. Despite hybridization mechanisms are possible for the s electrons of B with graphene states, these are irrelevant for properties which depend mostly upon the states at E_F, such as transport and magnetism. In agreement with previous works [21], for all the III-group elements, including B, because of the little occupation of this valence p-peak, the up- and down-spin states are degenerate and there is no net magnetic moment for this group of elements.

In this direction, charge transfer for this kind of elements seems to be the benchmark to understand the annihilation of magnetic moment, resulting in a nonmagnetic ground state. Extracting a quantitative value for charge transfer from ab initio calculations is useful for comparing with both experiment and simple models. Charge transfer is meaningful in the context of ionic bonding, since in the covalent-bonding case, charge is shared in the bond between adsorbate and substrate, and thus the charge-transfer concept is less relevant. However, charge transfer is an ambiguous quantity and there is no unique definition.

In order to evaluate this quantitatively, in this work we use a definition of charge transfer based on the shift in E_F relative to the graphene DOS resulting from the adatom adsorption. The DOS of the adatom-graphene system can be used to determine the charge transfer, assuming that graphene states remain unchanged due to the adatom adsorption (aside form a rigid shift). This is a good assumption in the B case. For the III-group-elements systems, electrons are donated to the graphene, filling the rigid graphene states and thereby shifting the Fermi level up from the Dirac point. To calculate the charge transfer from the adatom-graphene DOS, the Dirac point is identified, and the shift in Fermi level

$$\Delta E_F = E_F - E_D \tag{1}$$

is determined. The charge transfer Δq_{DOS} is given by the integral of the isolated graphene DOS from E_D to $E_D + \Delta E_F$. The Fermi-level shift ΔE_F and charge-transfer Δq_{DOS} are provided in Table 14.1 for each configuration, in order to show the significant role played by geometrical configuration in terms of the charge transfer, even if our interest is devoted to the bridge-site configuration.

Table 14.1 Charge-transfer DOS method. Fermi level shift relative to the graphene states ΔE_F, charge transfer determined from the DOS Δq_{DOS}. In red, charge transfer for bridge-site case. A positive charge transfer indicates transfer of electrons from the adatom to graphene; small differences in ΔE_F can give significant differences in Δq_{DOS}

Atom (site)	ΔE_F (eV)	Δq_{DOS}(e)
B (B)	0.7439	0.2818
B (T)	0.8904	0.3942
B (H)	0.9117	0.6309

Although by comparing with previous work of Chan et al. [21], referring to their analysis about III-group elements on hollow-site adsorption site, the value of charge transfer obtained from our calculations in hollow-site configuration is somehow in agreement with their results for other elements belonging to the III-group, the values obtained for top-site and bridge-site are lower than this latter configuration. In particular, one can expect that where the Fermi-level shift is similar, one has to obtain the same charge transfer, since it is provided by an integral on the same energy range of the same isolated-graphene sample. Analyzing the chemical properties of B along group and period, in the periodic table, focusing on the electronegativity values (commonly referred to as the Pauling scale) and ionization energies, one can see that:

- along the III-group: B has the highest electronegativity and ionization-energy values in the group, which are 2.0 and ∼800 kJ/mol, respectively, against average values for Al, Ga, and In which are 1.65 and ∼570 kJ/mol;
- along the first period: the benchmark element is Li, comparing results on it [21] with our calculations on B. According to these results, Li brings about the highest charge transfer among metals analyzed (about 1 electron per adatom), with a ΔE_F very similar to that of III-group elements. Although along the same period the geometrical features of atoms do not significantly change, due to the low electronegativity and ionization energy (1.0, ∼500 kJ/mol, values very close to those ones of III-group) lithium donates very easily its electron to graphene, unlike B.

Summarizing, what we suppose is that unlike elements belonging to the same group of periodic table, the expectations concerning charge transfer in the case of B should be reevaluated. This element appears to be more resistant to charge transfer, since the energy required to rip out valence electron is almost double with respect to other elements. We can conclude stressing that, the work of Chan et al. cannot provide us a complete vision of the charge-transfer problem, since their investigations do not include analysis of boron case and moreover, they study only the hollow-site configuration for the other elements. So that, the use of their results as reference for our ones might not return the best picture of the problem, since, assuming the reliability of the method, they do not provide any information on B and the values for charge transfer provided for other III-group elements refer to an integration of graphene DOS which can differ from our, dependently by parameters used in DFT calculations.

As proved by many works [16], substitutional N impurity in the graphene network, thanks to its very similar geometric features with C atoms, does not affect the graphene structure, resulting in a sp^2-hybridized N atom, where the remaining two electrons of the latter participate in the π-electron network of C atoms. The absence of a breaking of symmetry or alteration of the structure, returns a ground state of substitutional nitrogen which is always nonmagnetic.

If a N dopant does not occupy the substitutional sp^2 site in the graphene network, it will be adsorbed on graphene. Nitrogen is chemisorbed on graphene, binding to irregular carbon structures in sp^3 sites. The N adatom forms a bridgelike structure

on the graphene surface. In fact the adsorption energy has its highest value in correspondence of this configuration. The length of the $C − N$ bond is 1.24 Å and the length of the bond between the two C atoms bonded to the N is 1.54 Å. These two C atoms become sp^3-hybridized, whereas the $C − N − C$ angles of all bridge adsorption indicate sp^2 hybridization of N, without breaking the $C − C$ bond underneath the N atom. The underlying $C − C$ bonds are weakened due to the local tension caused by the N atom adsorption, in fact, contrary to the B case, a stronger deformation of the in-plane distribution of C atoms is induced by the presence of N. This is a signal that chemisorption of nitrogen has a strong influence on the structure, with a consequent modification of electronic features.

In contrast with the substitutional case, significant magnetism has been observed based upon spin-polarized calculations. The origin of the strong magnetism may be the spin-polarization of the chemisorbed atom as well as symmetry-breaking perturbations introduced by the chemisorbed atoms, and hence it may depend also on the position of the adatom in crystal lattice. Let us try to provide a complete scheme concerning magnetism in nitrogen-graphene system, first through the analysis of interaction of single-N adatom onto graphene, second giving some results about the interaction between two N adatoms in the same supercell, stressing differences in the magnetic order.

For the bridgelike chemisorption the N adatom forms two single bonds with the neighboring C atoms via sp^2 hybridization (perpendicular to the graphene surface). Two electrons participate in the single bonds with the C atoms while the other two remain as lone-pair electrons. One electron resides in the remaining un-hybridized p orbital which is oriented perpendicular to the local $C − N − C$ plane (as shown in the spin density projection in Fig. 14.6, providing, in the end the magnetization ∼1 μ_B. In general, one can assert that the magnitude of the magnetic moment is related to the coupling of this p orbital with the surface π orbitals, the stronger coupling, the smaller the moment, since the charge transfer (or in this case, charge sharing) can reduce the spin-polarization intensity. With respect to Krasheninnikov et al. [16], the magnitude of magnetic moment is not fractional (their value is 0.57 μ_B/cell). Assuming that calculations developed using PBE functional ensure robust results, in order to have a better understanding about origins of this discrepancy, we provide some results from calculations developed on a system with the same characteristics of that used by Krasheninnikov et al. (a 5×5 supercell, grid of 5 k-points), but with the exchange-correlation functional PBE, instead of Perdew–Wang [22] used in that work. The results return a magnetic ground state with a fractional magnetic moment of 0.54 μ_B, value very close to that provided by Krasheninnikov et al. Since we have used a different functional with respect to their work, obtaining the same magnetic moment value, we can not ascribe the origin of fractional magnetic moment to this difference, but more probably the main cause is related to a too small k-points grid, where a low-density grid could not correctly resolve the magnetic moment of the system. To verify this suspicion we studied the same system, with a more dense grid of k-points (about 20 points). This calculation returned a value of magnetic moment close to 1 μ_B, providing an important confirmation of our hypothesis and a
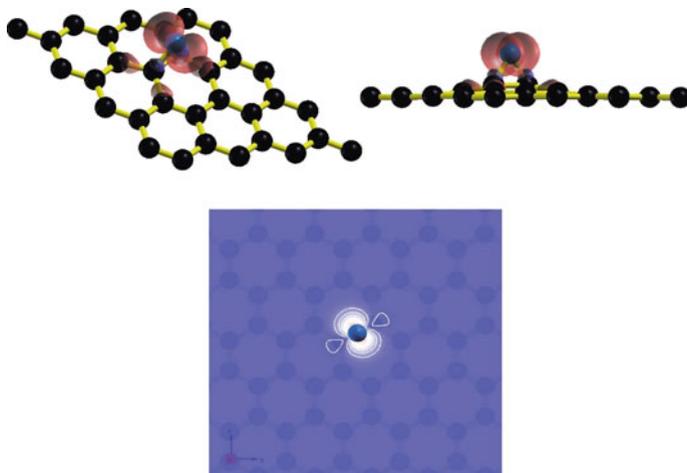
Fig. 14.6 Spin density projection (in μ_B/a.u.^{2}). *Top-Left* 3D picture of the spin magnetization density. The unpaired electrons are distributed orthogonally to the sp^2-hybridization plane $(C - N - C,$ along the bonding line), giving rise to spin-up polarization, denoted by *red* isosurfaces. *Top-Right xz*-plane section of the same. *Bottom* 2D-xy projection of spin density. White region represents the spin-density accumulation which promotes the up-polarization; isolines, strongly localized around the adatom, are a measure of how the interaction affect only the region surrounding N atom (color figure online)

clearer scheme of the possible cause for these "conflicting" results. Although these discrepancies, our findings seem to be well described by this model.

Some interesting results about interactions between a pair of N adatoms adsorbed on graphene can be shown. We consider a bridgelike configuration. The analysis of properties for the single-N-graphene systems has been conducted in a 4×4 supercell, since we regarded that the distance between adatoms periodic images was sufficient to avoid any interaction. Here we consider a larger system. For a first analysis a 6×6 supercell is used (returning a coverage of 2 adatoms per 72 C atoms), where the distance between the two N atoms is about 10 Å (similar to the previous single-adatom case). We expect that in the absence of strong interactions among the adatoms, the behavior of these two is almost independent, as in the case of single-adatom adsorption, giving rise to a most stable ferromagnetic order (in this particular geometrical configuration that in some way overlooks magnetic effects due to the bipartite lattice, proved by Lieb theorem). Our calculations indicate that the most stable magnetic order is the ferromagnetic one (the antiferromagnetic configuration shows a total energy which is ∼15 meV higher than ferromagnetic one. Although at finite temperature this stability could be affected, at low temperature the system prefers the expected ferromagnetic order, providing a total magnetization of 1.88 μ_B (in Fig. 14.7 the spin-density is shown for this system). The resulting fractional magnetic moment raised the suspicion that the magnitude of the magnetic moment of the system could depend on the size of system analyzed, since, if the two adatoms do
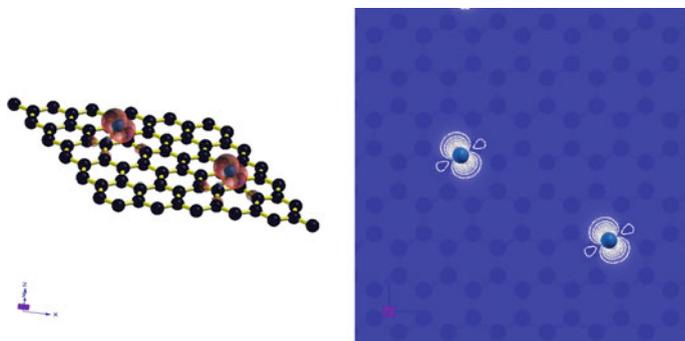
Fig. 14.7 Spin-magnetization density for 2N-graphene 6×6 supercell system. *Left* 3D picture of the spin polarization of states around the two adatoms. The magnetic order is ferromagnetic, returning a fractional magnetic moment of 1.88 μ_B. *Right* 2D xy-plane section of spin density. In agreement with the similarity concerning the spin distribution around the adatom in the single-nitrogen system (Fig. 14.6), one can think that the distance between these N atoms is enough to ensure that each one is independent from the other

not interact with each other and behave as two single adsorbed atoms, there should not be reason to expect a total magnetic moment different with respect to the sum of single moments, i.e., equal to 2 μ_B.

In order to investigate in this direction, we use a larger supercell, made up of 200 C atoms, a 10×10 supercell, so that one has a more diluted concentration of adatoms onto the surface. By setting the two N atoms at different distances in the same supercell, the smallest distance is $\sim$3 Å, whilst the greatest one is $\sim$23 Å, we obtain that the solution is always ferromagnetic regardless of distance among N atoms, providing a fractional magnetic moment of 1.48 μ_B, which is lower than the previous one calculated for the 6×6 supercell. In some way not entirely clear, the dilution of adatom concentration gives rise to a reduction of the magnitude of magnetic moment in this system.

The bridgelike structure implies that O interacts with the two C bonded atoms beneath, without breaking the bond among these, but just deforming the natural sp^2 hybridization status of the latter ones into C sp^3-hybridized. This gives rise to covalent bonding between C atoms and the sp^2-hybridized O atom, in the plane $C - O - C$ perpendicular to the graphene surface. Oxygen has 6 electrons in the outer shells, divided into 2 paired electrons in the $2s$ shell and the other 4 which occupy the $2p$ orbitals, following the Hund's rule. So that, 2 electrons of these 4, remain unpaired. Unlike the nitrogen case, where one has three unpaired electron, two of whose form bonds with C atoms beneath, and the other one remain unpaired, providing a spin-polarizability of the system, here, O pairs the two unpaired electrons belonging to $2p$ orbitals sharing the other two with the p_z electron of C atoms, i.e., there is a saturation of the bonds, returning a nonmagnetic state. Figure 14.8 shows the charge localization around the bonded structure, representative of what we said so far.
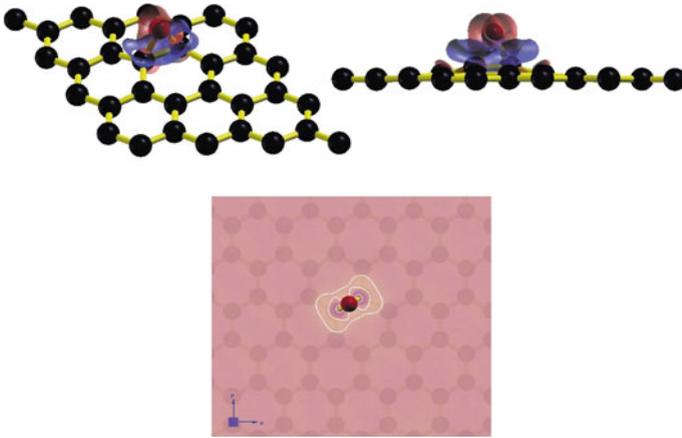
Fig. 14.8 Electronic charge density representation for Oxygen-graphene system. *Top-Left* 3D vision of the charge distribution around the O adatom (*red*). Isosurfaces [*red* (*blue*), charge accumulation (depletion)] are indicative of the spatial localization of the charge shared by the total system. *Top-Right* xz-plane section of the same. *Bottom* xy-plane section of the charge density. It is clear how charge is more localized with respect to the B case, in agreement with the absence of a doping mechanism and charge transfer. Isolines denote this aspect. The *yellow* region is indicative of the charge sharing region, whilst the two violet spots denote the accumulation of charge which surrounds the two covalent bonds with C atoms (color figure online)

Analyzing the PDOS of the system one can see that the presence of O does not affect the DOS of clean graphene, except for peaks localized below and above E_F in the energy range of $|1.5|$ eV (see Fig. 14.9). These peaks refer to the contribution of
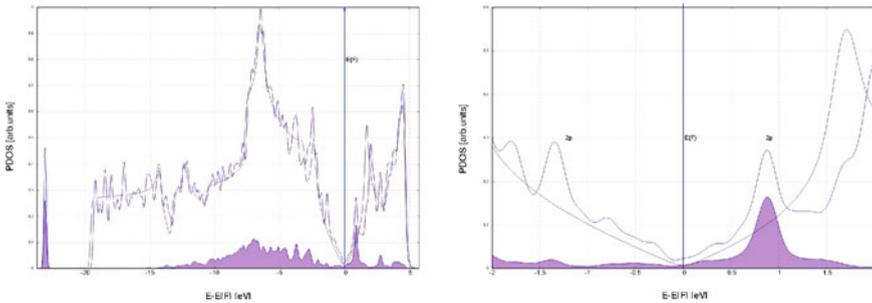
Fig. 14.9 Spin-resolved total projected-DOS for graphene monolayer with single-O atom on bridge-site. *Top Black line* denotes the clean-graphene DOS, whereas the *violet* one represents the spin-resolved total projection of oxygen-graphene density of states (spin-up and spin-down projection are perfectly overlapped). *Violet shaded region* denotes the contribution to DOS (*up and down*) from O p-orbitals. *Bottom* we stress the surrounding of E_F with the aim to mark two splitted peaks (exchange interaction) representative of the localized state induced by the hybridization of C atom below O. The interaction with oxygen does not alter the Fermi level position which matches that one of the pristine sample (color figure online)

two sp^3-hybridized carbon atoms which bind with O. Since electrons of these states are completely paired thanks to the sharing with electrons of O adatom, so that up- and down-spin states are mirror images, with the same number of states/energy, the resulting system is in a nonmagnetic ground state.

Conclusions

Graphene quantum dots (GQDs) are potential candidates for quantum information processing thanks to their properties which ensure the highest coherence time for electron spin confined inside the dot and the tunability of quantum dots (QDs) in order to couple specific electron spins, even allowing a coupling in a long-range distance between two different QDs.

Nevertheless, defects in the structure or interactions with impurities lead to the formation of localized electronic states which give rise to magnetic moment, due to the occupation of the state by just one unpaired electron (i.e., unpaired spin). This localized spin can interact with the qubit spin in the QD, providing decoherence, resulting in possible loss of information. In fact, the general effect of this induced magnetic ground state is related to a possible coupling between the spin-defect and the central-spin of QD, worsening and destroying the coherence. This investigation led on different possible sources of decoherence has provided that vacancy, the simplest point defect, breaks down the symmetry of the graphene system, giving rise to localized states which provide a magnetic moment different from zero. It has a magnitude depending on the defect concentration, the greater the concentration, the lower the magnitude. Hydrogen interaction produces similar effects, hybridizing locally the system and giving rise to a magnetic moment. Furthermore, it enhances the spin-orbit coupling of about three orders of magnitude with respect to the intrinsic spin–orbit interaction. Studying the interaction between two hydrogen atoms on the same supercell, one can see the formation of a ferromagnetic order when adatoms are located on the same sublattice, regardless of the distance between them, whereas putting these on different sublattices, according to Lieb's theorem, the most stable magnetic order is antiferromagnetic, appreciable only if the distance between the adatoms is sufficient to avoid hybridization mechanisms which couple the unpaired electrons belonging to the two independent impurities.

About the adatom interactions, boron, nitrogen, and oxygen have been considered and their influence on the generation of magnetism in graphene has been studied. In the end the only one element, in the bridgelike configuration, which provides a magnetic character to graphene seems to be nitrogen N. The size of cell within the adatom is adsorbed, i.e., the concentration of adatoms/cell seems to influence the magnitude of the magnetic moment, in particular, fixing the number of adatom per cell, the bigger is cell, the lower is the magnitude. Therefore, one should pay attention to these results, since that in order to obtain a graphene quantum dot system with the minimal influence on spin-relaxation and spin-coherence times for forming reliable spin qubits, one should have no decoherence sources. Here N could be

the main responsible of perturbation of spin qubit, since this can interact with the electron-spin of adatom-graphene system, promoting the deterioration of quantum information.

References

1. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, A.A. Firsov, Science **306**, 666 (2004). DOI 10.1126/science.1102896
2. P.C. Hohenberg, Phys. Rev. **158**(2), 383 (1967). DOI 10.1103/PhysRev.158.383
3. L.D. Landau, Phys. Z. Sowjetunion **11**, 26 (1937)
4. N.D. Mermin, H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966). DOI 10.1103/PhysRevLett.17.1133. [Erratum *ibid.***17**, 1307 (1966)]
5. O.V. Yazyev, L. Helm, Phys. Rev. B **75**, 125408 (2007). DOI 10.1103/PhysRevB.75.125408
6. D. Loss, D.P. DiVincenzo, Phys. Rev. A **57**, 120 (1998). DOI 10.1103/PhysRevA.57.120
7. M.A. Ruderman, C. Kittel, Phys. Rev. **96**, 99 (1954). DOI 10.1103/PhysRev.96.99
8. M. Gmitra, D. Kochan, J. Fabian, Phys. Rev. Lett. **110**, 246602 (2013). DOI 10.1103/PhysRevLett.110.246602
9. D.P. DiVincenzo, Fortschritte der Physik **48**(9–11), 771 (2000). DOI 10.1002/1521-3978(200009)48:9/11<771::AID-PROP771>3.0.CO;2-E
10. B. Trauzettel, D.V. Bulaev, D. Loss, G. Burkard, Nat Phys **3**(3), 192 (2007). DOI 10.1038/nphys544
11. P.O. Lehtinen, A.S. Foster, Y. Ma, A.V. Krasheninnikov, R.M. Nieminen, Phys. Rev. Lett. **93**, 187202 (2004). DOI 10.1103/PhysRevLett.93.187202
12. W. Han, R.K. Kawakami, M. Gmitra, J. Fabian, Nat Nano **9**(10), 794 (2014). DOI 10.1038/nnano.2014.214. Review
13. S. Engels, A. Epping, C. Volk, S. Korte, B. Voigtländer, K. Watanabe, T. Taniguchi, S. Trellenkamp, C. Stampfer, Appl. Phys. Lett. **103**(7), 073113 (2013). DOI http://dx.doi.org/10.1063/1.4818627
14. K. Nakada, A. Ishii, Solid State Commun. **151**(1), 13 (2011). DOI 10.1016/j.ssc.2010.10.036
15. J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996). DOI 10.1103/PhysRevLett.77.3865. [Erratum Phys. Rev. Lett. **78**, 1396 (1997)]
16. Y. Ma, A.S. Foster, A.V. Krasheninnikov, R.M. Nieminen, Phys. Rev. B **72**, 205416 (2005). DOI 10.1103/PhysRevB.72.205416
17. P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A.D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, J. Phys.: Cond. Matter **21**(39), 395502 (2009). DOI 10.1088/0953-8984/21/39/395502
18. M.A.H. Vozmediano, M.P. López-Sancho, T. Stauber, F. Guinea, Phys. Rev. B **72**, 155121 (2005). DOI 10.1103/PhysRevB.72.155121
19. S. Saremi, Phys. Rev. B **76**, 184430 (2007). DOI 10.1103/PhysRevB.76.184430
20. E.H. Lieb, Phys. Rev. Lett. **62**, 1201 (1989). DOI 10.1103/PhysRevLett.62.1201
21. K.T. Chan, J.B. Neaton, M.L. Cohen, Phys. Rev. B **77**(23), 235430 (2008). DOI 10.1103/PhysRevB.77.235430
22. J.P. Perdew, W. Yue, Phys. Rev. B **33**, 8800 (1986). DOI 10.1103/PhysRevB.33.8800

Chapter 15
Electron Correlations in Molecular Chains

L.S. Brizhik

Abstract It is shown that in a one-dimensional molecular chain two extra electrons with opposite spins interacting with acoustical phonons in the harmonic approximation form a bisoliton, which is a bound singlet state of the electrons trapped by a self-created deformation of the chain. The harmonic approximation in description of lattice vibrations can be violated for molecular chains which are relatively rigid, for systems with relatively strong electron–lattice interaction (resulting in large deformation of the chain) or under the external pressure applied to the system. In this case, the deformation can be described taking into account anharmonic terms. It is shown that the account of these terms leads to stabilization of bisolitons up to the sound velocity in the chain, and to formation of bisolectrons, which can propagate with the supersound velocity. In the case of finite number of electrons in a molecular chain their ground state is a self-consistent cnoidal wave in the deformation potential. This cnoidal wave at sufficiently small doping level corresponds to the condensate of autolocalized bisolitons (bisolectrons). The dependence of the condensate gap width on concentration of charge carriers is shown to have a nonmonotonic character. The pressure dependence of the condensate critical temperature is analyzed and shown to be of the type characteristic for HTSCs under pressure.

Introduction

It is widely known fact that low-dimensional molecular systems possess very unusual physical properties. In many cases this is due to electron–lattice interaction which in such systems can be relatively strong, plays a significant role and can result in formation of polarons in such systems [1]. Examples are conducting polymers, polydiacetylene, macromolecules, organic, and high-temperature superconducting systems. Many of such systems satisfy the adiabatic approximation and possess relatively

L.S. Brizhik (✉)
Bogolyubov Institute for Theoretical Physics, Nonlinear Condensed
Matter Physics Department, National Academy of Sciences of Ukraine,
14-b Metrolohichna str., Kiev 03680, Ukraine
e-mail: brizhik@bitp.kiev.ua

strong electron–lattice coupling. It has been shown that under these conditions a quasiparticle (extra electron or hole, or molecular excitation) is self-trapped in the deformation potential and forms a soliton [2].

Two extra electrons with opposite spins interacting with acoustical phonons in a one-dimensional molecular chain form a bisoliton which is a bound singlet state of the electrons trapped by a self-created deformation of the chain [3]. In the case of finite number of electrons in a molecular chain their ground state is a self-consistent cnoidal wave in the deformation potential [4].

These results are obtained in the harmonic approximation for the lattice vibrations. In many cases this approximation can be violated. Examples are relatively rigid molecular chains, systems with relatively strong electron–lattice interaction which results in large deformation of the chain, exposure of the system to external pressure, etc. In these cases the deformation can be described taking into account anharmonic terms describing lattice vibrations [5–7].

Some low-dimensional systems show significant dependence of their properties on the charge carrier concentration and pressure, like, for instance, high-T_c superconductors (HTSC) [8–11]. So far to my knowledge the first theory which has expressed doping and pressure influences on the critical temperature of HTSCs was suggested in [12], and it was based on the bisoliton model [13].

Below we study properties of a bisoliton in a harmonic chain (section "Bisolitons in Harmonic Lattices"), bisoliton condensate (section "Bisoliton Cnoidal Wave"), the dependence of the bisoliton energy gap on pressure and doping (section "Pressure Dependence of the Critical Temperature") and role of the lattice anharmonicity in stabilizing bisolitons (section "Bisolitons in Anharmonic Lattices").

Bisolitons in Harmonic Lattices

Let us consider an infinitely long relatively soft molecular chain. For simplicity we assume there is a one atom in a unit cell of the chain, or, in other words, we assume there is one active phonon mode, which is acoustical mode. At intermediate values of the electron–lattice coupling constant the ground state of an extra electron in it corresponds to the Davydov soliton, which is a localized solution of the system of the nonlinear equations. In the linear approximation such a soliton corresponds to the large polaron [2]. Such a system can be described by the Fröhlich Hamiltonian, which in the case of extra two extra electrons with opposite spins can be written in the following form:

$$H = H_{el} + H_{ph} + H_{el-ph} + H_{Coul}, \tag{15.1}$$

Here, the electron Hamiltonian, H_{el}, has the form

$$H_{el} = \sum_{n,s=1,2} \left[E_0 A_{n,s}^\dagger A_{n,s} - J A_{n,s}^\dagger \left(A_{n+1,s} + A_{n-1,s} \right) \right], \tag{15.2}$$

where $A_{n,s}^{\dagger}$ ($A_{n,s}$) are Fermi operators of creation (annihilation) an electron with spin s on the lattice site n, E_0 is the on-site electron energy and J is the electron exchange interaction energy.

In the adiabatic approximation the lattice can be described in semiclassical approximation. In this case the Hamiltonian of acoustical phonons can be written as follows:

$$H_{\text{ph}} = \sum_n \left[\frac{\hat{p}_n^2}{2M} + \hat{U}\left(\hat{\beta}_n\right) \right]. \tag{15.3}$$

Here, $\hat{\beta}_n$ is the operator of the displacement of the nth unit cell from its equilibrium position, $\hat{p}_n$ is the operator of the canonically conjugated momentum, M is the mass of a unit cell. In Eq. (15.3) the operator $\hat{U}$ is the operator of the potential energy of the lattice. Let us first consider the lattice in the harmonic approximation. Then this operator reads as

$$\hat{U}\left(\hat{\beta}_n\right) = \frac{1}{2}w\left(\hat{\beta}_n - \hat{\beta}_{n-1}\right)^2, \tag{15.4}$$

where w is the coefficient of the elasticity of the chain.

We will consider the case when the dependence of the on-site electron energy on the longitudinal displacements of unit cells exceeds the dependence of the inter-cite electron energy. In this case, the Hamiltonian of the electron–lattice interaction is given by the expression

$$H_{\text{el-ph}} = \chi \sum_{n,s=1,2} \left(\hat{\beta}_{n+1} - \hat{\beta}_{n-1}\right) A_{n,s}^{\dagger} A_{n,s}, \tag{15.5}$$

where χ is the electron–lattice coupling constant.

The Coulomb repulsion between the electrons is given by the Hubbard-type Hamiltonian

$$H_{\text{Coul}} = \sum_{n,m,s=1,2} V_{nm} A_{n,s}^{\dagger} A_{n,s} A_{m,s}^{\dagger} A_{m,s}, \tag{15.6}$$

where V_{nm} is the corresponding matrix element of the Coulomb interaction. Below, we will omit this term of the Hamiltonian assuming that the Coulomb repulsion is screened in the lattice, and will take it into account later on. The screening can be significant in lattices with several atoms in a unit cell and in relatively soft molecular chains so that the localization of electron wave functions is extended over several lattice sites.

In the adiabatic approximation, we can choose the vector state of the system as the product of the vector state of electrons and vector state of the lattice:

$$|\Psi(t)\rangle = |\Psi_{\text{el}}(t)\rangle |\Psi_{\text{ph}}(t)\rangle. \tag{15.7}$$

In its turn, the vector state of the lattice has the form of the product of the operator of coherent displacements of unit cells and vacuum state of the lattice, $|0\rangle_{ph}$,

$$|\Psi_{\mathrm{ph}}(t)\rangle = \exp\left\{-\frac{i}{\hbar}\sum_n\left[\beta_n(t)\hat{p}_n - p_n(t)\hat{\beta}_n\right]\right\}|0\rangle_{ph}, \qquad (15.8)$$

where $\beta_n(t)$ and $p_n(t)$ are, respectively, the mean values of the displacements of unit cells from their equilibrium positions and their canonically conjugated momenta in the state Eq. (15.7).

The electron state vector for two excess electrons has the form

$$|\Psi_{\mathrm{el}}(t)\rangle = \sum_{n_1,n_2,s_1,s_2}\Psi(n_1,n_2,s_1,s_2;t)A^{\dagger}_{n_1,s_1}A^{\dagger}_{n_2,s_2}|0\rangle_{\mathrm{el}}. \qquad (15.9)$$

The two-electron wave function $\Psi(n_1, n_2, s_1, s_2; t)$ for electrons with antiparallel spins can be represented as the product of the symmetric coordinate function $\Psi(n_1, n_2, t)$ and antisymmetric spin function $\chi(s_1, s_2)$. Here, we set

$$\Psi(n_1,n_2,t) = \frac{1}{\sqrt{2}}\left[\psi_1(n_1,t)\psi_2(n_2,t) + \psi_2(n_1,t)\psi_1(n_2,t)\right], \qquad (15.10)$$

where quasi-one-electron wave functions $\psi_j(n,t)$ satisfy the normalization condition

$$\sum_n|\psi_j(n,t)|^2 = 1. \qquad (15.11)$$

Recall, we use the term 'quasi-one-electron wave functions' for functions $\psi_j(x,t)$ to distinguish them from a proper one-electron wave functions to indicate that they determine the two-electron wave function in the given ansatz, Eq. (15.10), and only formally look like a one-electron wave functions. It is worth to recall here that this is not the only possible ansatz for the wave function, obviously. One can introduce the function which depends on the center of mass coordinate and relative coordinate, but this ansatz describes the state which has higher energy than the one we use here, as it has been shown in [14].

For analytical study we will use the continuum approximation for which we introduce the continuum variable $z \equiv na$ instead of a discreet index n, assuming that electron and lattice wave functions vary slowly in space. In fact such assumption is valid at the intermediate values of the electron–lattice interaction, when self-trapped soliton state is the ground state of an extra electron in the lattice, as comparing with the small polaron state formed at strong electron–lattice interaction, and almost free electron state in a limit of a very weak electron–lattice interaction. In our case the soliton state is extended over few lattice cites, as we will see later. Therefore, the continuum description is a good approximation to describe solitons.

The state vector, Eq. (15.7), defines the Hamiltonian functional

$$\mathscr{H} = \langle\Psi(t)|H|\Psi(t)\rangle,$$

corresponding to the Hamiltonian operator, Eq. (15.1). Minimization of the functional $\mathcal{H}$ with respect to electron and phonon variables gives us the system of coupled equations

$$
i\hbar\frac{\partial\Psi(z_1,z_2,t)}{\partial t} + J\left(\frac{\partial^2\Psi(z_1,z_2,t)}{\partial z_1^2} + \frac{\partial^2\Psi(z_1,z_2,t)}{\partial z_2^2}\right)
$$
$$
+ \chi a\left(\frac{\partial\beta(z,t)}{\partial z}\Big|_{z=z_1} + \frac{\partial\beta(z,t)}{\partial z}\Big|_{z=z_2}\right)\Psi(z_1,z_2,t) = 0,
$$
$$(15.12\text{a})$$

$$
\frac{\partial^2\beta}{\partial t^2} - V_{ac}^2\frac{\partial^2\beta}{\partial z^2} =
$$
$$
\frac{a}{M}\chi\left(\int dz_2\frac{\partial|\Psi(z_1,z_2,t)|^2}{\partial z_1}\Big|_{z_1=z} + \int dz_1\frac{\partial|\Psi(z_1,z_2,t)|^2}{\partial z_2}\Big|_{z_2=z}\right).
$$
$$(15.12\text{b})$$

Here V_{ac} is the sound velocity in the chain, $V_{ac} = a\sqrt{w/M}$.

Let us introduce the function of the lattice deformation $\rho(z,t) = -\partial\beta(z,t)/\partial z$. It is easy to see that for the ansatz Eq. (15.10) the system of equations (15.12) can be reduced to the following system of coupled equations for quasi-one-electron wave functions $\Psi_j(z,t)$ ($j = 1,2$) and lattice deformation:

$$
i\hbar\frac{\partial\psi_j(z,t)}{\partial t} + J\frac{\partial^2\psi_j(z,t)}{\partial z^2} + \chi a\rho(z,t)\psi_j(z,t) = 0, \qquad (15.13\text{a})
$$

$$
\frac{\partial^2\rho(z,t)}{\partial t^2} - V_{ac}^2\frac{\partial^2\rho(z,t)}{\partial z^2} = -a\frac{\chi}{M}\frac{\partial^2}{\partial z^2}\sum_{j=1,2}|\psi_j(z,t)|^2. \qquad (15.13\text{b})
$$

The latter equation has the solution

$$
\rho(\xi) = \frac{\chi}{w(1-s^2)}\sum_{j=1,2}|\psi_j(\xi)|^2, \qquad (15.14)
$$

where ξ is the running-wave coordinate, $\xi \equiv (z - z_0 - Vt)/a$, and the notation is used $s^2 \equiv V^2/V_{ac}^2$.

Substituting the solution, Eq. (15.14), into Eq. (15.13a), we obtain the two-component nonlinear Schrödinger equation

$$
i\hbar\frac{\partial\psi_j(z,t)}{\partial t} + J\frac{\partial^2\Psi_j(z,t)}{\partial z^2} + 2Jg\left[|\psi_1(z,t)|^2 + |\Psi_2(z,t)|^2\right]\psi_j(z,t) = 0,
$$
$$(15.15)$$

where the dimensionless electron–phonon interaction coefficient is introduced

$$g = \frac{\chi^2}{2Jaw(1-s^2)} = \frac{g_0}{1-s^2}, \qquad g_0 \equiv \frac{\chi^2}{2Jaw}. \qquad (15.16)$$

Here, we are interested in the localized lowest energy solutions and, taking into account the identity of two electrons, except their spins, we can choose the following ansatz $|\psi_2(z,t)|^2 = |\psi_1(z+la,t)|^2$ and, hence, can omit index j. In the class of localized functions, whose modulus depends on ξ, we set $\psi(z,t) = \phi(\xi)\exp[im^*Vz/\hbar - i\varphi(t)]$. Here m^* is the effective electron mass, $m^* = \hbar^2/(2Ja^2)$. Substituting this into Eq. (15.15), we come to the equation

$$\frac{d^2\phi(\xi)}{d\xi^2} + 2g\left[\phi^2(\xi) + \phi^2(\xi+l)\right]\phi(\xi) = -\varepsilon_l\phi(\xi), \qquad (15.17)$$

where the value

$$\varepsilon_l \equiv \frac{E-E_0}{J} = \int d\xi \left\{ \left(\frac{d\phi}{d\xi}\right)^2 - 2g\phi^2(\xi)\left[\phi^2(\xi) + \phi^2(\xi+l)\right] \right\} \qquad (15.18)$$

corresponds to the eigenenergy of the solutions, measured in units of J.

It has been shown in [3] that the highest energy solution of Eq. (15.17) corresponds to the case of two almost independent solitons of nonlinear Schrödinger equation (NLSE), localized infinitely far from one another, i.e. to $l \to \infty$:

$$\phi(\xi) = \phi_s(\xi) \equiv \frac{\sqrt{g}}{2}\,\mathrm{sech}\left(\frac{g\xi}{2}\right). \qquad (15.19)$$

The lowest energy solution of Eq. (15.17) corresponds to the case when the distance between center of mass coordinates of quasi-one-electron wave functions vanishes, $l = 0$,

$$\phi(\xi) = \phi_{bs}(\xi) \equiv \sqrt{\frac{g}{2}}\,\mathrm{sech}(g\xi) \qquad (15.20)$$

This localized bound state of two electrons is called 'bisoliton.' Its eigenenergy, as can be easily calculated from Eq. (15.18), is

$$\varepsilon_{bs} \equiv \varepsilon_{l\to 0} = -g^2. \qquad (15.21)$$

It exceeds the double eigenenergy of an isolated soliton,

$$\varepsilon_s \equiv \varepsilon_{l\to\infty} = -\frac{g^2}{4}. \qquad (15.22)$$

The binding energy of the bisoliton which is determined as the difference of the total energy of the bisoliton solution, Eq. (15.20), and total energy of the almost independent infinitely far localized solitons Eq. (15.19), including the energy of the lattice deformation in the corresponding states, $E_{\text{bind}}(V) \equiv E_{l\to\infty}(V) - E_{l\to0}(V)$, is a function of the bisoliton velocity

$$E_{\text{bind}} = \frac{1}{2}Jg_0^2\,\frac{1-5s^2}{(1-s^2)^3}. \tag{15.23}$$

One can easily see, that the binding energy is positive only at $|V| < V_{\text{ac}}/\sqrt{5}$, when the loss of the energy because of the increase of the energy of the lattice deformation in the bisoliton state, is less than the gain of the eigenenergy due to the binding of two electrons. We remind here that this result is obtained in the harmonic approximation for the lattice vibrations.

It is remarkable that the width of the bisoliton, $\kappa_{\text{bs}} = g/\pi$, is smaller than the width of an isolated soliton, $\kappa_{\text{s}} = 2g/\pi$, according to Eqs. (15.20) and (15.19). This width determines also the width of the lattice deformation in the corresponding state, according to Eq. (15.14). Therefore, we conclude that in the bisoliton state electron-induced lattice deformation is stronger, and, respectively, the deformational potential is deeper than in the case of electrons in quasi-independent soliton state.

In the above considerations we have omitted the Coulomb repulsion between the electrons. This repulsion in molecular chains is screened and can be taken into account using the perturbation method. Thus, it has been shown in [3], that in the result of this repulsion center of mass coordinates of quasi-one-electron wave functions, l, in the bisoliton state takes some finite value, which depends on the interplay of the gain in energy due to the binding of two electrons and loss of the energy due to the Coulomb repulsion:

$$l = l_0 \equiv \left(\frac{e^2}{4\pi\varepsilon a\zeta}\right)^{1/3}, \tag{15.24}$$

where

$$\zeta = \left[\frac{4}{3}\frac{\chi^2 a^2 \rho_0^2 g}{MV_{\text{ac}}^2} - \frac{1}{3}wa^2\rho_0^2 g^2.\right] \tag{15.25}$$

Here, e is the effective electron charge with account of its screening in the molecular chain due to the surrounding and complex structure of a unit cite, $\varepsilon = \varepsilon_{\text{m}}\varepsilon_0$ is the dielectric constant of the chain, which contains the dielectric constant ε_{m} of the medium, ρ_0 is the maximum value of the lattice deformation. Substituting in Eqs. (15.24) and (15.25) the value

$$\rho_0 = \frac{\chi g}{w}, \tag{15.26}$$

which follows from (15.13b), l_0 can be approximated by the following expression:

$$l_0 = \left(\frac{3e^2 w}{4\pi \varepsilon \chi^2 a^3 g^3} \right)^{1/3}. \tag{15.27}$$

Due to the Coulomb repulsion, the bisoliton wave function takes the form of a single-bell envelope at small values of l_0 and the double-bell envelope at large values of l_0 with the distance between the two maxima of the envelope, equal to l_0.

Bisoliton Cnoidal Wave

Let us now consider in the same harmonic approximation the case when in the molecular chain there are $2N_e$ charge carriers (electrons or holes) in the state when their total spin is zero. For simplicity we will refer below to charge carriers as to electrons. We have shown above that in the zero adiabatic approximation two electrons with opposite spins interacting with the local deformation of a chain, bind in the autolocalized singlet spin bisoliton state. Although bisoliton is a singlet state, strictly speaking, it is not a Bose-particle, there is spin interaction between the electrons. It has been shown in [15] that, due to the Pauli principle, two bisolitons repel one another because of this spin interaction. As the result, at some finite concentration of electrons, δ_e, the periodic cnoidal wave is formed in a self-consistent deformation potential of the chain [4]. Generalizing the vector state ansatz Eqs. (15.9) and (15.10) for the case of $2N_e$ electrons, we can show that the wave function Ψ_{2N_e} of this cnoidal wave in the continuum approximations satisfies the nonlinear Schrödinger equation

$$\left\{ i\hbar \frac{\partial}{\partial t} + J \frac{\partial^2}{\partial z^2} + 4Jg|\Psi_{2N_e}(z, t)|^2 \right\} \Psi_{2N_e}(z, t) = 0, \tag{15.28}$$

where the dimensionless electron–phonon coupling coefficient has been determined above in Eq. (15.16).

Two extra conditions are imposed on the wave function, the periodicity, and normalization conditions, namely,

$$\Psi(z, t) = \Psi(z + La, t), \quad \frac{1}{a} \int_0^{La} |\Psi(z, t)|^2 dz = 1, \tag{15.29}$$

in which $L = 1/\delta_e$ is the period of the cnoidal wave measured in units of the lattice constant a.

Equation (15.28) under the conditions Eq. (15.29) admits two qualitatively different types of solutions, delocalized and localized ones. These solutions are separated by the energy gap.

In particular, the delocalized solution is described by the function with constant envelope, and has the energy

$$|\Psi_{\text{del}}(z,t)|^2 = \frac{1}{L}, \quad E_{\text{del}}(V) = -\frac{4Jg}{L} + m^*V^2, \tag{15.30}$$

respectively. Here m^* is the effective mass of an electron, $m^* = \hbar^2/2Ja^2$.

The second type solution is described by the cnoidal wave function

$$\Psi_{\text{cn}}(z,t) = \phi_{\text{cn}}(\xi) \exp\left\{\frac{i}{\hbar}\left[m^*Vz - (E + \frac{m^*V^2}{2})t\right]\right\}, \quad \xi = (z - z_0 - Vt)/a, \tag{15.31}$$

where E is the energy of the solution per period and the envelope function is determined by the Jacobi elliptic function:

$$\phi_{\text{cn}}(\xi) = \sqrt{\frac{g}{2}E^{-1}(m)}\,\text{dn}\left[\frac{g\xi}{E(m)}, m\right], \tag{15.32}$$

where m is the modulus of the elliptic function, and $E(m)$ is the complete elliptic integrals of the second kind. In some textbooks on the elliptic functions the notation is used $m = k^2$.

The modulus m of the elliptic function is determined by the density of electrons from the equation [4]

$$E(m)K(m) = \frac{g}{2\delta_e}, \tag{15.33}$$

in which $K(m)$ is the complete elliptic integrals of the first kind. The same equation can be interpreted in another way as the equation which determines the electron density as the function of the modulus of the elliptic function. It is shown in Fig. 15.1.

The numerical analysis of expression Eq. (15.37) shows that there exists the critical value of electron density

$$\delta_e^{(\text{cr})} = \frac{2g_0}{\pi^2}, \tag{15.34}$$

above which the energy gap width vanishes (see also Fig. 15.1).

The localized solution corresponds to the condensate of a periodically distributed lattice of 'bisolitons' moving coherently in the deformation potential along the chain. The total energy of the system, calculated per period, includes the eigenenergy of the cnoidal wave per period and energy of the lattice deformation:

$$E_{\text{cn}}(V) = m^*V^2 + 2J\int_0^L\left[\left(\frac{d\phi_{\text{cn}}}{d\xi}\right)^2 - 2g\frac{1 - 3s^2}{1 - s^2}\phi_{\text{cn}}^4\right]d\xi. \tag{15.35}$$
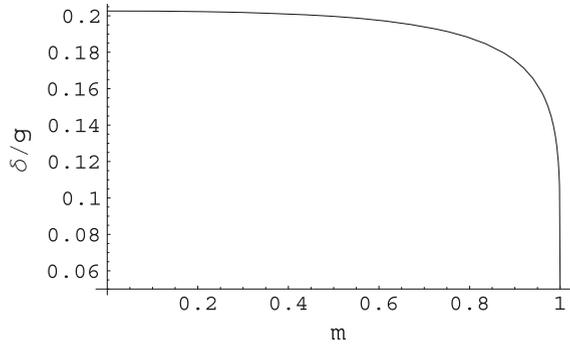
Substituting in this expression the explicit form of the cnoidal wave function and carrying the integration, we get

$$E_{\mathrm{cn}}(V) \ = \ - \frac{2}{3} J g^2 \frac{\mathrm{E}(m)(2-m) \ + \ \mathrm{K}(m)(1-m)}{\mathrm{E}^3(m)} \, . \qquad (15.36)$$

The width of the energy gap of the bisoliton condensate is determined by the difference of the densities of energies corresponding to the solutions (15.30) and (15.31) at zero velocities, and, hence, is the function of the electron density,

$$\Delta \ = \ \delta_e \left[E_{\mathrm{del}}(0) \ - \ E_{\mathrm{cn}}(0) \right] \ = \ \frac{1}{3} J g_0^3 F(\delta_e) \, , \qquad (15.37)$$

where the notation is used

$$F(\delta_e) \ \equiv \ \frac{\mathrm{E}(m)(2-m) + \mathrm{K}(m)(1-m)}{\mathrm{E}^4(m)\mathrm{K}(m)} \ - \ \frac{3}{\mathrm{E}^2(m)\mathrm{K}^2(m)} \, . \qquad (15.38)$$

It follows from Eq. (15.37), that the gap attains the maximum value at small values of density, when, according to Eq. (15.33), the modulus of the elliptic function takes values close to one. Using the expansions of complete elliptic integrals in the corresponding limit, equivalent to the limit of vanishing additional modulus, $m_1 \equiv 1 - m \ll 1$, we get the following expression

$$\Delta \approx \frac{2}{3} J g_0 \delta_e (g_0 - 6\delta_e) \, , \qquad \delta_e \approx \frac{g_0}{2 \ln(4/\sqrt{m_1})} \, , \qquad m_1 \ll 1. \qquad (15.39)$$

Figure 15.2 shows the dependence of the energy gap width on the density of charge carriers.

From Eq. (15.39) we calculate the optimal electron density value and the width of the corresponding energy gap
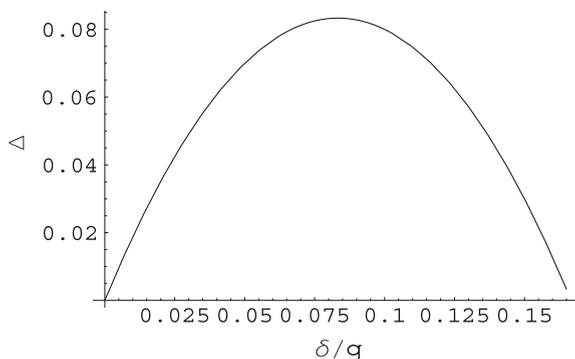
$$\delta_e^{(opt)} = \frac{g_0}{12}, \qquad \Delta_{max} = \frac{Jg_0^3}{36}. \qquad (15.40)$$

Notice that in frame of the bisoliton model of superconductivity the critical temperature T_c is proportional to the energy gap width at zero temperature [13], $\Delta = \alpha T_c$, where α is a numerical constant. By fitting the maximum value of the critical temperature obtained in our model to the experimental value ($T_c = 38$ K for lanthanum compound and $T_c = 90$ K for yttrium compound), we get [12] $Q = 460$ and $Q = 1090$, respectively, where Q is the characteristic dimensionless parameter:

$$Q \equiv \frac{Jg^3}{3\alpha}. \qquad (15.41)$$

These results are in good agreement with the experimental data on the dependence of T_c on δ_e, as it has been demonstrated in [12] (see Fig. 2 in [12]). So far to our knowledge, it was the first model of high-temperature superconductivity which explained the experimental data on the dependence of the critical temperature on the density of charge carriers.

Pressure Dependence of the Critical Temperature

Another distinct feature of the high-temperature superconductivity is the dependence of the critical temperature on applied pressure. Not only high-temperature superconductors show unusual impact of pressure. There are some organic low-dimensional materials (Bechgaard salts) which undergo superconducting transition, when they are exposed to external pressure. It is worth to recall that the common feature of these two classes of materials is their strong anisotropy, or, in other words, their low-dimensionality. So it seemed natural to extend the bisoliton model to explain the pressure dependence of the critical temperature [12].

For this aim the following relation is useful

$$\frac{\partial T_c}{\partial P} = -\frac{T_c}{B}\frac{\partial \ln T_c}{\partial \ln V} \approx -\frac{T_c}{3B}\frac{\partial \ln T_c}{\partial \ln a}, \tag{15.42}$$

where P, V, B, a are pressure, volume, bulk modulus, and lattice constant, respectively.

As mentioned above, the critical temperature is connected with the energy gap width at zero temperature by the relation $\Delta = \alpha T_c$, in which α is a numerical constant [4]. The logarithmic derivative of the energy gap follows from Eq. (15.39),

$$\frac{\partial \ln \Delta}{\partial \ln a} = \frac{\partial \ln J}{\partial \ln a} + \left(1 + \frac{g_0}{g_0 - 6\delta_e}\right)\frac{\partial \ln g_0}{\partial \ln a} + \left(1 - \frac{6\delta_e}{g_0 - 6\delta_e}\right)\frac{\partial \ln \delta_e}{\partial \ln a}. \tag{15.43}$$

Here $\delta_e < \delta_e^{(cr)}$, where $\delta_e^{(cr)}$ is determined in Eq. (15.34).

In its turn, according to Eq. (15.16), the following equality is valid

$$\frac{\partial \ln g_0}{\partial \ln a} = 2\frac{\partial \ln \chi}{\partial \ln a} - \frac{\partial \ln J}{\partial \ln a} - \frac{\partial \ln w}{\partial \ln a}. \tag{15.44}$$

Substituting relation (15.44) into Eq. (15.42), we get

$$\frac{\partial T_c}{\partial P} = -\frac{Q}{3B}\left\{F\left(\frac{\partial \ln J}{\partial \ln a} + \frac{\partial \ln g_0}{\partial \ln a} + \frac{\partial \ln \delta_e}{\partial \ln a}\right) + \frac{2\delta_e}{g_0}\frac{\partial \ln g_0}{\partial \ln a} - \frac{12\delta_e^2}{g_0^2}\frac{\partial \ln g_0}{\partial \ln a}\right\}, \tag{15.45}$$

where F is the function defined in (15.38), and parameter Q is defined in Eq. (15.41).

The parameters $p = g$, δ_e, χ, w, J, appearing in Eq. (15.45), can be estimated as follows. The elasticity coefficient can be obtained from the potential energy of the copper–oxygen ion interaction, which can be described by the Buckingham potential $U = -A_1 r^{-6} + A_2 \exp(-cr)$,

$$w = \left[\frac{d^2 U}{dr^2}\right]_{r=a} = \frac{6A_1}{a^8}(ca - 7), \tag{15.46}$$

from which the logarithmic derivative follows

$$\frac{\partial \ln w}{\partial \ln a} = -\frac{7(ca - 8)}{ca - 7} \approx -7. \tag{15.47}$$

The electron–phonon coupling constant can be estimated as the coefficient of the term $(u_{2n+2} - u_{2n})$ in the expansion of the energy of extra electron at an oxygen ion in the chain

$$E = -\frac{Ze^2}{|R_{2n+1} - R_{2n}|} - \frac{Ze^2}{|R_{2n+1} - R_{2n+2}|}. \tag{15.48}$$

Here, Ze is the charge of the copper ion, $R_j = R_j^0 + u_j$ is the coordinate of the jth ion, u_j is the displacement of copper ($j = 2n$) or oxygen ($j = 2n + 1$) ion from its equilibrium position. Performing the expansion of expression (15.48), we get

$$\chi = \frac{4Ze^2}{a^2}, \quad \frac{\partial \ln \chi}{\partial \ln a} = -2. \tag{15.49}$$

Finally, the copper–oxygen exchange integral in the tight-binding approximation can be written as

$$J = -\int \Psi_O^*(\mathbf{r})[W(\mathbf{r}) - w(\mathbf{r})]\Psi_{Cu}(\mathbf{r})d\mathbf{r}, \tag{15.50}$$

where $W(\mathbf{r}) - w(\mathbf{r})$ is the potential energy of electron interaction with all ions except the one it occupies, $\Psi_O(\mathbf{r})$ and $\Psi_{Cu}(\mathbf{r})$ are the wave functions of oxygen and copper ions in states $2p^5$, $3d^{10}$, respectively. Substituting the explicit form of radial parts of functions into Eq. (15.50) and carrying out the integration, we get

$$J \approx A\left(\frac{12}{\rho}\sinh\frac{\rho}{12} - \cosh\frac{\rho}{12}\right) e^{-5\rho/12} \approx ca^2\exp(-\beta a), \tag{15.51}$$

where $\rho = a/2a_B \approx 3.7$ with a_B being the Bohr radius, $\beta = 5/(24a_B)$, A, and c are constants independent of the lattice constant. Differentiating expression (15.51), we find

$$\frac{\partial \ln J}{\partial \ln a} = 2 - \beta a \approx 0.45, \tag{15.52}$$

which, together with Eqs. (15.44), (15.47) and (15.48), gives $(\partial \ln g)/(\partial \ln a) \approx 2.4$.

Let us now analyze the pressure dependence of the charge carrier density δ_e. using the values obtained experimentally from Hall effect measurements [16]. For instance, for yttrium compound this value was estimated in [11] as $(\partial \ln \delta_e)/(\partial \ln a) = -90$, somewhat less value of $(\partial \ln \delta_e)/(\partial \ln a)$ is found for lanthanum compound (see [17]).

Provided the charge density is less than the critical value determined in Eq. (15.34), the approximate equality follows from Eq. (15.45)

$$\frac{\partial T_c}{\partial P} \approx -\frac{2C\delta_e}{3Bg_0}\left[\frac{\partial \ln J}{\partial \ln a} + 2\frac{\partial \ln g_0}{\partial \ln a} + \frac{\partial \ln \delta_e}{\partial \ln a} - 6\frac{\delta_e}{g_0}(\frac{\partial \ln J}{\partial \ln a} + \frac{\partial \ln g_0}{\partial \ln a} + 2\frac{\partial \ln \delta_e}{\partial \ln a})\right]. \tag{15.53}$$

Substituting into the last equation the values of the logarithmic derivatives of the parameters J and g obtained above, and the experimental value of the logarithmic derivative of δ_e, we can conclude that the critical temperature of the bisoliton condensate can increase or decrease with the increase of pressure depending on the value of the charge carrier density and its logarithmic derivative, which vary from experiment
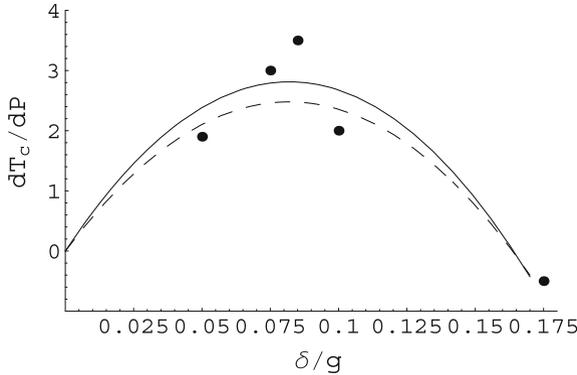
Fig. 15.3 Pressure derivative of critical temperature $\partial T_c/\partial P$ (K/GPa) as a function of composition x in La$_{2-x}$M$_x$CuO$_4$ compound. *Solid line* corresponds to logarithmic derivative of the electron density equal to -80, *dashed line* to -90, bulk modulus is $B = 180$ GPa. The experimental data are taken from [11]

to an experiment. In or opinion, this explains the wide scattering of the experimental data on the pressure dependence of the critical temperature of HTSCs.

In conclusion, we represent in Fig. 15.3 the pressure derivative of the critical temperature, $\partial T_c/\partial P$, as a function of doping x for the compound La$_{2-x}$M$_x$CuO$_4$ given by Eq. (15.53). The experimental data are taken from Fig. 12 in [17]. Note that the bisoliton model reproduces well the general trend of the experimental data. Strictly speaking, our model should be compared with the experimental data on uniaxial compression of superconducting samples instead of hydrostatic pressure. But, first, at present such experiments are very few, and second, the comparison we have done indicates that the bisoliton model even in a one-dimensional formulation without attraction of any extra mechanism or assumption describes very well the behavior of superconducting oxides under pressure.

Bisolitons in Anharmonic Lattices

In the previous sections we have considered bisolitons and bisoliton condensate in molecular chains in the harmonic approximation for the lattice vibrations. This approximation is enough to describe soft chains, such as formed by hydrogen bonds. In more rigid chains, such as Cu–O chains in copper oxides, or in systems under high pressure this approximation can be violated. In such a case one has to take into account the anharmonic terms. The simplest cases are cubic and/or quartic terms, added to the harmonic potential. In particular, the operator of the lattice potential $\hat{U}\left(\hat{\beta}_n\right)$ in the phonon Hamiltonian Eq. (15.3) in these cases gives the following expressions for the phonon potential energy.

To get the explicit solutions we have to specify the lattice potential. Below we will consider two cases of c, and, respectively, we will assign to the functions:

$$U_c(\rho) = \frac{1}{2}\rho^2 + \frac{\alpha}{3}\rho^3, \quad U_q(\rho) = \frac{1}{2}\rho^2 + \frac{\beta}{4}\rho^4, \tag{15.54}$$

where ρ is the lattice deformation, α and β are anharmonicity coefficients in cubic and quartic anharmonic potentials, and subscripts c or q correspond to cubic and quartic anharmonic potentials, respectively.

It has been shown in [6, 7] that the anharmonic terms of the potential modify the maximum value of the lattice deformation. While in the harmonic lattice this value was given by the expression Eq. (15.26), in the anharmonic lattices it becomes a function of the dynamically modulated inverse anharmonic stiffness coefficient δ, $\nu = c, q$, where

$$\delta_c = \frac{1 - s^2}{\alpha}, \quad \delta_q = \frac{1 - s^2}{\beta}. \tag{15.55}$$

It follows from the numerical analysis of the expressions for $\rho_{0(\nu)}$, that (i) the maximum lattice deformation depends on the soliton velocity; (ii) the soliton amplitude increases and its width decreases with the velocity increasing, attaining some finite values at the sound velocity, $V = V_{ac}$ (i.e., when $\delta = 0$); (iii) the soliton amplitude increases with the electron–lattice coupling increasing; (iv) the quartic anharmonicity is dominant at small values of δ (large velocities), while cubic anharmonicity is dominant at larger values of δ (small velocities).

We will omit here the details of the solution of the system of coupled equations for electrons and lattice variables with the account of the anharmonic terms. Instead we cite the final expressions final expressions for the bisoliton wave functions (see [18])

$$\phi_c(\xi) = \sqrt{\frac{\rho_{0(c)}}{2D}} \text{Sech}(\kappa_c\xi)\sqrt{1 - s^2 + \alpha\rho_{0(c)}\text{Sech}^2(\kappa_c\xi)}, \tag{15.56a}$$

$$\phi_q(\xi) = \sqrt{\frac{\rho_{0(q)}}{2D}} \text{Sech}(\kappa_c\xi)\sqrt{1 - s^2 + \beta\rho_{0(q)}^2\text{Sech}^4(\kappa_q\xi)}, \tag{15.56b}$$

where

$$\kappa_c = \sqrt{\frac{\sigma\rho_{0(c)}}{2}} \frac{\sqrt{4\rho_{0(c)}\left(\rho_{0(c)} + 2\delta_c\right)/3 + \delta_c^2}}{2\rho_{0(c)} + \delta_c}, \tag{15.57a}$$

$$\kappa_q = \frac{1}{2}\sqrt{\frac{\sigma\rho_{0(q)}\left(3\rho_{0(q)}^2 + 2\delta_q\right)}{\rho_{0(q)}^2 + 2\delta_q}}. \tag{15.57b}$$

Here the notation is used $D = \chi a/(M V_{ac}^2)$,

The eigenenergy and the energy of the lattice deformation in lattice with cubic anharmonicity are

$$E_c^{(bs)}(V) = -DMV_{ac}^2 \rho_{0(c)} \frac{4\rho_{0(c)} + 3\delta_c}{3(\rho_{0(c)} + \delta_c)},$$ (15.58)

$$W_c(V) \approx \frac{MV_{ac}^2}{3\sqrt{2\sigma}} \rho_{0(c)}^{3/2} \left(\frac{8}{15}\alpha\rho_{0(c)} + 1 + s^2 \right),$$ (15.59)

respectively, and in lattice with quartic anharmonicity these two energies are

$$E_q^{(bs)}(V) = -\frac{1}{2}DMV_{ac}^2 \rho_{0(q)} \frac{3\rho_{0(q)}^3 + 2\delta_q}{\rho_{0(q)}^2 + \delta_q},$$ (15.60)

$$W_q(V) \approx 8\frac{MV_{ac}^2}{\sqrt{2\sigma}} \rho_{0(q)}^{3/2} \left[\frac{1}{3} \left(s^2 + \frac{1}{2}\delta\beta \right) + \frac{2}{35}\beta\rho_{0(q)}^2 \right].$$ (15.61)

Two important conclusions follow from the above expressions. First of all, comparing the bisolectron energies with the energies of solectrons (see [5]), we conclude, that there is positive binding energy of the bisolectron in the whole interval of velocities $V^2 \leq V_{ac}^2$

$$E_{bind(\nu)}^{(bs)}(V) = 2E_{tot(\nu)}^{(s)}(V) - E_{tot(\nu)}^{(bs)}(V), \quad \nu = c, q$$ (15.62)

Here $E_{tot(\nu)}^{(bs)}(V)$ is the total energy of the system in the bisolectron state with account of the energy of the lattice deformation, and $E_{tot(\nu)}^{(s)}(V)$ is the energy of the system with one electron in a soliton state with account of the energy of the lattice deformation. This means that anharmonic lattice soliton can capture two electrons with opposite spins and that such a bisolectron state is energetically favorable as comparing with two independent solectrons (lattice soliton bound with one electron).

Another important conclusion from the above results is that bisolectrons are stable in the whole interval of the velocity up to the sound velocity in the chain. The bisolectron energy and the energy of the lattice deformation take finite values at the velocity of the bisolectron equal to the velocity of the sound in the chain, namely:

$$E_{tot(c)}^{(bs)}(V_{ac}) = mV_{ac}^2 - \frac{2}{3}\chi a\rho_{0(c)} + \frac{16}{45}\chi a\alpha\rho_{0(c)}^2,$$ (15.63)

$$E_{tot(q)}^{(bs)}(V_{ac}) = mV_{ac}^2 - \frac{3}{2}\chi a\rho_{0(q)}^2 + \frac{8}{35}\chi a\beta\rho_{0(q)}^3.$$ (15.64)

Recall, in harmonic lattice bisolitons are stable at small velocities, $s^2 < 1/5$, i.e., at $|V| < V_{ac}/\sqrt{5}$, as it follows from Eq. (15.23).

These results can be directly generalized for the periodic solutions of the corresponding system of equations for $2N_e$ electrons with total spin equal zero.

Conclusions

Thus, we have shown that two extra electrons with opposite spins interacting with acoustical phonons in a one-dimensional molecular chain in the harmonic approximation for the lattice vibrations form a bisoliton. Such a bisoliton is a bound singlet state of the electrons trapped by a self-created deformation of the chain. Bisolitons in harmonic lattices are stable at small enough velocities due to the fast increase of the energy of lattice deformation with the increase of the velocity.

In the case of finite number of electrons in a molecular chain their ground state is a self-consistent cnoidal wave in the deformation potential. We have calculated the bisoliton condensate energy gap and analyzed the dependence of the critical temperature of the bisoliton condensate on pressure and doping. These results are close to the corresponding experimental data for HTSCs.

We have also shown that in anharmonic lattices (with cubic or quartic anharmonicity), the bisolitons are stable in the whole interval of the velocities, up to the velocity of the sound in the chain.

References

1. A.S. Alexandrov (ed.), *Polarons in Advanced Materials* (Springer, Dordrecht, 2007)
2. A.S. Davydov, Phys. Rep. **190**(4), 191 (1990). DOI 10.1016/0370-1573(90)90061-6
3. L.S. Brizhik, A.S. Davydov, Fiz. Nizk. Temp. **10**(7), 748 (1984)
4. L.S. Brizhik, A.S. Davydov, Fiz. Nizk. Temp. **10**(4), 358 (1984)
5. L.S. Brizhik, A.S. Davydov, Physica Status Sol. B **115**(2), 615 (1983). DOI 10.1002/pssb.2221150233
6. M.G. Velarde, L. Brizhik, A.P. Chetverikov, L. Cruzeiro, W. Ebeling, G. Röpke, Int. J. Quantum Chem. **112**(13), 2591 (2012). DOI 10.1002/qua.23282
7. M.G. Velarde, L. Brizhik, A.P. Chetverikov, L. Cruzeiro, W. Ebeling, G. Röpke, Int. J. Quantum Chem. **112**(2), 551 (2012). DOI 10.1002/qua.23008
8. Y. Tokura, J.B. Torrance, T.C. Huang, A.I. Nazzal, Phys. Rev. B **38**, 7156 (1988). DOI 10.1103/PhysRevB.38.7156
9. J.B. Torrance, Y. Tokura, A.I. Nazzal, A. Bezinge, T.C. Huang, S.S.P. Parkin, Phys. Rev. Lett. **61**, 1127 (1988). DOI 10.1103/PhysRevLett.61.1127

10. R. Griessen, Phys. Rev. B **36**, 5284 (1987). DOI 10.1103/PhysRevB.36.5284
11. R.J. Wijngaarden, E.N. van Eenige, J.J. Scholtz, H.K. Hemmes, R. Griessen, in *Molecular systems under high pressure: Proceedings of the II Archimedes workshop on Molecular solids under high pressure, Catania, Italy, May 28–31, 1990*, ed. by R. Pucci, G. Piccitto (North-Holland, Amsterdam, 1991), pp. 157–180
12. L. Brizhik, A. La Magna, R. Pucci, F. Siringo, High Press. Res. **11**(6), 375 (1994). DOI 10.1080/08957959408203165
13. L.S. Brizhik, A.S. Davydov, Physica Status Sol. B **143**(2), 689 (1987). DOI 10.1002/pssb.2221430228
14. L.S. Brizhik, A.A. Eremko, Physica D **81**(3), 295 (1995). DOI 10.1016/0167-2789(94)00206-6
15. L.S. Brizhik, A.A. Eremko, Ukr. J. Phys. **44**(8), 1022 (1999)
16. I.D. Parker, R.H. Friend, J. Phys. C: Solid State Phys. **21**(11), L345 (1988). DOI 10.1088/0022-3719/21/11/002
17. R.J. Wijngaarden, R. Griessen, in *Studies of high temperature superconductors*, ed. by A.V. Narlikar (Nova Science, New York, 1993), p. 160
18. L.S. Brizhik, A.P. Chetverikov, V. Ebeling, G. Röpke, M.G. Velarde, Ukr. J. Phys. **58**, 562 (2013)

Chapter 16
Hydrogen-Bonded Systems Under Intense Electric Fields

G. Cassone, F. Saija, A.M. Saitta and P.V. Giaquinta

Introduction

Thermal effects induced on matter by electric fields are well known and largely understood [1]. The same is not true for nonthermal effects and for the related microscopic mechanisms that are triggered by intense electric fields [2, 3]. Our ignorance on this matter is manifest over different scales. In fact, nonthermal effects are poorly understood both at a subatomic, quantum, level as well as at molecular and macroscopic levels. An account of the current state of the art in this specific field has recently appeared in the literature [4]. Aragones et al. have given the first experimental evidence that an electric field can control chemical reactions, showing that

G. Cassone
Institute of Biophysics, Czech Academy of Sciences, Královopolská 135,
61265 Brno, Czech Republic
e-mail: giuseppe.cassone@impmc.upmc.fr

F. Saija
CNR-IPCF, Viale F. Stagno d'Alcontres 37, 98158 Messina, Italy
e-mail: saija@ipcf.cnr.it

A.M. Saitta
Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC),
Sorbonne Universités, Université Pierre et Marie Curie Paris 06,
Unité Mixte de Recherche 7590, 75005 Paris, France
e-mail: marco.saitta@impmc.upmc.fr

A.M. Saitta
CNRS, UMR 7590 IMPMC, 75005 Paris, France

P.V. Giaquinta (✉)
Dipartimento di Scienze Matematiche e Informatiche, Scienze Fisiche e Scienze della Terra,
Università degli Studi di Messina, Contrada Papardo, 98166 Messina, Italy
e-mail: paolo.giaquinta@unime.it

the field strength, as well as its polarity, can actually drive, enhance, or even inhibit a given reaction (e.g., a Diels–Alder reaction).

In order to explain such a "delay" in investigating and clarifying phenomenological aspects of materials which, in some respects, would sound logically plausible – if not absolutely trivial – even to a nonexpert, we should consider some experimental and theoretical boundary conditions on the nature of the problem. First of all, very intense local electric fields are necessary in order to induce a rearrangement of covalent bonds, definitely not an easy task on the experimental side. Second, only very recently reliable theoretical approaches have been implemented and developed, which are capable of treating, in the framework of a quantum description of matter, field-induced effects. Such tools, thanks to the fast growth of available computing power, have eventually allowed for the discovery of some fundamental "pieces" of physics and chemistry when matter is being irradiated with electric fields.

The basic interaction between a stationary, spatially homogeneous electric field, and one single molecule in the vacuum is well understood over a wide range of intensities. By resorting to the theoretical tools provided by quantum mechanics and Density Functional Theory (DFT), in conjunction with advanced computing techniques, it is possible to simulate in a realistic way and, thus, finely describe all the relevant degrees of freedom of such a simple model system. However, the situation becomes much more intricate and complicated when dealing with condensed-matter systems, in which also complex interactions between molecules have to be taken into account. The subtle interplay between field-induced and neighbor-induced polarization effects, in addition to the role of thermodynamic conditions, makes the task of the computational physicist much harder. But an even more fundamental reason has represented, on this side, a tough hurdle when attempting at modelling and simulating realistic material samples under the effect of static electric fields. Indeed, typical numerical experiments carried out over homogeneous samples are ordinarily performed on a cell with periodic boundary conditions (PBC); in this way, one avoids, to a large extent, the effects induced by the artificial spatial confinement of the elementary interacting constituents, which would obviously inhibit any significant comparison of the numerical results with the properties of a real-life homogeneous material. In classical molecular dynamics (MD) simulations the electric field is implemented as an additional force acting on each particle. However, the serious problem in dealing with such force fields resides in the effective intermolecular potential which is not able to account, in a reliable way, for polarization effects as well as, more importantly, for the very quantum nature of matter.

Since, as previously pointed out, electric fields are able to induce the cleavage and the formation of (intrinsically quantum) covalent bonds, it is clear that classical simulations fail to be predictive in this specific context in that they do not provide a complete physical representation of the involved phenomena. In order to model and reproduce the whole spectrum of effects induced by the application of an electric field, ab initio molecular dynamics (AIMD) simulations are mandatory for a twofold reason. First, they are able to reproduce correctly intra and intermolecular structural properties in many disparate systems because of the appropriate, more fundamental description of the material (see Ref. [5] and references contained therein). Second,

AIMD simulations are, by their own nature, intrinsically able to treat, in principle, every kind of covalent-bond rearrangement and, hence, to deal with chemical reactions.

Unfortunately, because of the nonperiodic nature of the quantum position operator, the implementation of electric fields in ab initio simulation codes is other than easy. Indeed, only in 2002 Umari and Pasquarello [6] discussed and implemented the first operational theoretical framework which, upon exploiting Berry's phases and the modern theory of polarization [7–9], was able to manage static electric fields in first-principles simulations under PBC.

In this chapter, we review the microscopic phenomenology of hydrogen-bonded liquid and solid systems under the action of a static and homogeneous electric field, showing different shades of the effects produced by the application of the field. Schematically, H-bonds have partial covalent, electrostatic, dispersed, and polarized character [10]. Hence, with such multifaceted properties of electronic matrix, their treatment by means of ab initio techniques is particularly suited. Moreover, as it will be pointed out in the following sections, molecular dissociations can be induced through proton transfers for high enough field intensities; correspondingly, related simple and complex chemical reactions can be modeled and investigated. Hence, from a theoretical perspective, AIMD simulations currently represent a unique technique for adequately treating, reproducing, and predicting many microscopic and macroscopic phenomena which occur when matter is exposed to the action of intense electric fields.

The chapter is structured as follows. In the next sections, various effects manifested for increasing field strengths will be treated with a gradual approach: in section "Low-to-Moderate Field Regime: The Response of the H-Bonded Network" the response of liquid and solid H-bonded systems to low-to-moderate fields will be presented; in section "Moderate-to-Strong Field Regime: Proton Transfer and Molecular Dissociation" several properties of proton transfer and molecular dissociation induced by moderate-to-strong fields will be illustrated; finally, in section "Strong Field Regime: Field-Induced Chemical Reactions" complex chemical reactions produced by very intense electric fields will be analyzed.

Low-to-Moderate Field Regime: The Response of the H-Bonded Network

Liquids

By definition, H-bonds are formed in specific polar solvents, which are composed by molecules that carry an intrinsic dipole moment. The fundamental coupling between a static and homogeneous electric field and a single dipolar entity is a basic physics textbook problem. However, when we consider a set of molecules interacting through H-bonds at a finite temperature, the situation is clearly more complex.

Let us first consider the seemingly most studied liquid in condensed-matter physics, *viz.*, water. What happens when water is subject to relatively low-strength fields? Typically, the radial distribution functions (RDFs), describing the mutual positional order of all the atomic species dissolved in a sample, represent a good tool for analyzing the time-averaged structural changes induced by an external field. An extensive analysis has been performed in neat liquid water at a nominal temperature of 350 K [11]. However, pure water is an abstraction; hence, we shall also discuss the properties of salty water, in which sodium and chlorine ions have been solvated at a molarity of 1.7 M and at a temperature of 315 K [12]. The rearrangement of water molecules induced by the field is quite similar in both liquids since the relatively low molarity of the electrolytic solution does not significantly alter the local structure around a randomly chosen water molecule. In Fig. 16.1 we display the oxygen–oxygen (O–O) and the oxygen–hydrogen (O–H) RDFs in salty water calculated by Cassone et al. [12] for various field strengths. A zero-field Car-Parrinello [13] dynamics about 28 ps long was performed and then a field of increasing intensity was applied. Density correlations between pairs of oxygen atoms and between oxygen and hydrogen atoms get manifestly weaker and weaker as the field strength increases. In fact, the maxima of both O–O and O–H RDFs are depressed, whereas the minima rise; in other words, the aqueous solvent becomes more and more disordered at short and medium distances, as also shown by Saitta, Saija, and Giaquinta [11] for neat water.

The way in which the local structure of water changes in response to an external field does not represent the general behavior necessarily exhibited by all H-bonded systems. In fact, let us consider the simplest alcohol, i.e., methanol (CH_3OH), a one-donor H-bonded system which certainly constitutes an illustrative case in this context. If the same kind of numerical experiment (i.e., exploiting the same DFT setup) is carried out on liquid methanol at ambient temperature (300 K), a dual response
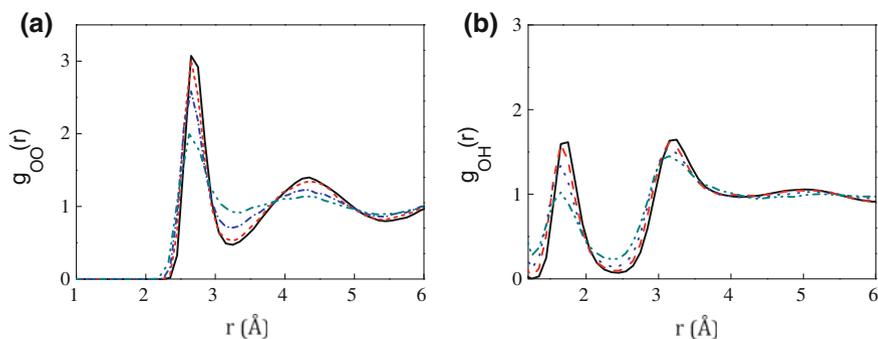
Fig. 16.1 Oxygen–oxygen (**a**) and oxygen–hydrogen (**b**) radial distribution functions in a 1.7 M solution of Na^+ and Cl^- ions in water, calculated for increasing strengths of the electric field [12]; *black solid lines* $E = 0$ V/Å; *red dashed lines* $E = 0.15$ V/Å; *blue dotted-dashed lines* $E = 0.25$ V/Å; *green dotted-dotted-dashed lines* $E = 0.35$ V/Å. The average response of the local structure to the field is similar to that documented in Ref. [11] for neat water (color figure online)
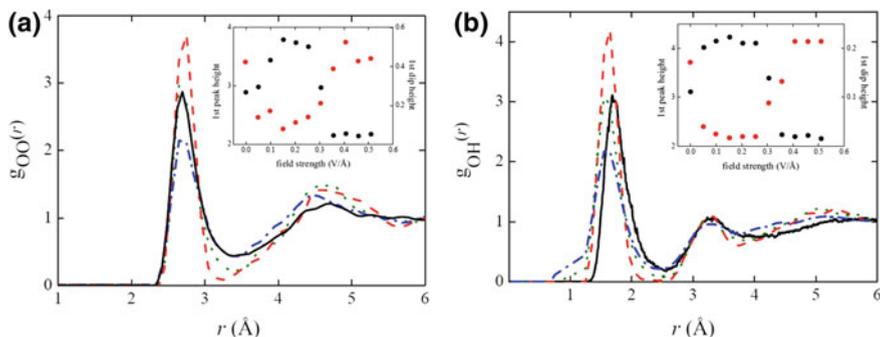
Fig. 16.2 Oxygen–oxygen (**a**) and oxygen–hydrogen (**b**) radial distribution functions of liquid methanol calculated for increasing values of the electric field. *Black solid line E = 0 V/Å; red dashed line E = 0.15 V/Å; green dotted line E = 0.30 V/Å; blue dotted-dashed line E = 0.45 V/Å.* *Inset* heights of the first peak (*black solid circles*) and of the first dip (*red solid circles*) plotted as a function of the field intensity (color figure online)

regime of the H-bonded network can be easily recognized. As shown in Fig. 16.2, at variance with water pair correlations between oxygen atoms are enforced for relatively low-field strengths. Indeed, liquid methanol gets more ordered at short and medium distances as both maxima and minima of the O–O RDF are enhanced and the whole shape is sharpened. However, as the field intensity grows, this trend eventually changes in a significant way: in fact, the first peak of the O–O RDF, after reaching its maximum value for $E \approx 0.15$ V/Å, abruptly drops to about 50% of its top value for field intensities in the range between 0.25 and 0.35 V/Å, and remains approximately constant thereon (see the inset of Fig. 16.2a). Correspondingly, the first dip rises, as also does the second peak. The first peak and the immediately following dip of the O–H RDF behave in a similar way (see Fig. 16.2b). In addition, the first coordination shell moves to moderately shorter distances and a weak shoulder gradually emerges at about 1 Å, which indicates an increasing freedom of hydrogen atoms in sampling shorter relative distances from an oxygen atom. Hence, the H-bonded network of liquid methanol shows an ambivalent response to different regimes of the field intensity. In the low-field regime H-bonds are strengthened by the application of the field whereas, for higher intensities, intermolecular correlations are depressed.

A further correlated consequence produced by the application of static electric fields on H-bonded liquids, such as water and methanol, is the gradual aligning of increasing fractions of molecular dipole moments along the field direction. Let θ be the angle formed by dipoles with the field direction (z axis). As shown in Fig. 16.3a for water and in Fig. 16.3b for methanol, in the absence of the field the distribution $P(\theta)$ of molecular dipole orientations exhibits a maximum for $\theta = 90°$, since it is not normalized to the solid angle. However, when an electric field has been switched on, the maximum of the distribution shifts, for both water and methanol, to lower values of θ since molecules tend to align with the field, the larger the number the stronger the field. Whereas intermolecular correlations are discouraged
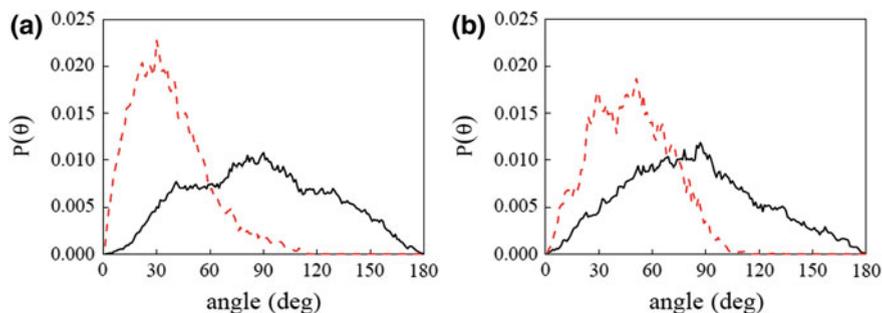
Fig. 16.3 Distribution of the angle θ formed by dipole moment vectors with the electric field axis in water (**a**) and methanol (**b**). *Black curve $E = 0$ V/Å; red broken curve $E = 0.25$ V/Å* (color figure online)

in water at all the investigated field regimes, in methanol the initial molecular aligning process increases the robustness of the H-bonded network because of the peculiar spatial statistical distribution of the H-bond donor/acceptor sites which enhances the probability of formation of H-bonds [14].

Solids

The field-induced aligning effect of molecular dipole moments discussed in the previous section leads to peculiar phenomena in a solid sample. H-bonds are clearly more stable in solids than in liquids. Hence, one would expect less striking effects in solids when a static electric field has been applied. Indeed, in the ambient I_h ice phase of water no significant field-induced molecular rotations or structural modifications have been observed [15, 16]. On the other hand, some effects have been observed in ice XI, the ferroelectric counterpart of ice I_h [15].

The external field produces a further notable effect in that it "flattens" the planes parallel to its direction. This phenomenon is more clearly resolved in ice XI, as shown in Fig. 16.4, and can be assimilated to a sort of "pressure" exerted by the applied field which forces the mutual alignment of molecular dipole moments [11, 17]. The resulting polarization effect ultimately enhances the electrostatic coupling between the molecular dipole moments and the external field in a powered manner. In ice I_h the flattening of the planes was not observed as clearly as in the ferroelectric case, because a larger energy is needed, on average, to align (pseudo)disordered dipole moments with the field direction. The pressure exerted by the field increases the capability of the water molecules of exploring shorter intermolecular distances. This electrostatic mechanism triggers the crucial phenomena of proton transfer and molecular dissociation which will be treated in the next section.
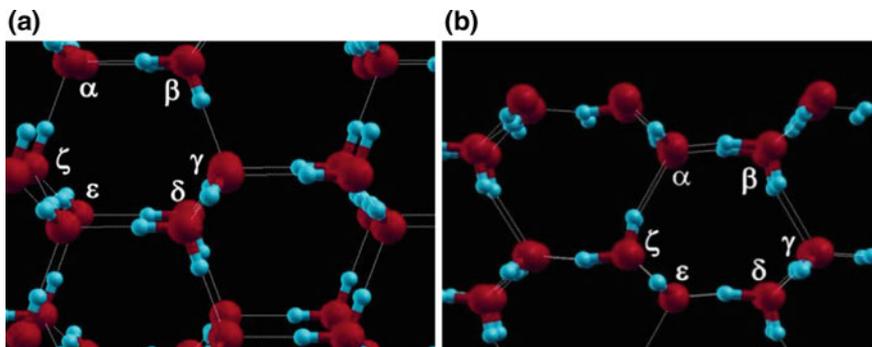
Fig. 16.4 a The standard ice XI structure and **b** the field-induced planes flattening, with pertinent angles labeled as Greek letters; panel **a** $\alpha = \beta = \gamma = \delta = \varepsilon = \zeta = 109.45°$; panel **b**: $\alpha = 113.2°$, $\beta = 112.9°$, $\gamma = 102.3°$, $\delta = 114.8°$, $\varepsilon = 114.8°$, and $\zeta = 102.4°$, a configuration which exhibits a shortening of the in-plane intermolecular lengths (i.e., flattening)

Moderate-to-Strong Field Regime: Proton Transfer and Molecular Dissociation

The strength of quantum-based MD calculations mainly resides with the capability to treat covalent bonds and hence to reproduce (at least, in principle) any chemical reaction. The simplest mechanism by which a rearrangement of covalent bonds occurs is that leading to the ionization of a molecule through the loss or gain of a proton. This phenomenon is of paramount importance since it represents the foundation of the concept of pH of an aqueous solution, of the functioning of many ecological devices which exploit proton transfer (PT), such as Nafion membranes and direct methanol fuel cells (DMFC), as well as of large parts of (neuro)biology and biochemical sciences [18, 19].

The typical textbook example is obviously water. In liquid water the process through which a H_2O molecule dissociates, thus setting the acid or basic character of an aqueous solution, is known as *protolysis* and occurs according to the reaction

$$2H_2O \rightleftharpoons OH^- + H_3O^+ , \tag{16.1}$$

in which, formally, a PT occurs between two water molecules, giving rise to the formation of the hydroxide (OH^-) and hydronium (H_3O^+) ions. This circumstance also leads to an important conjugate redistribution of H-bonds in the condensed phase of water.

We have investigated PT and molecular dissociation mechanisms in water, salty water and water ices, as well as in liquid methanol. Single-donor H-bonds make bulk methanol a crucial system for understanding the PT mechanism which can be basically schematized through the following reaction:

$$2CH_3OH \rightleftharpoons CH_3O^- + CH_3OH_2^+ . \tag{16.2}$$

Like water, liquid methanol displays a fast PT. This process takes place along the H-bond structure via the general Grotthuss mechanism [20], whose exact nature has been (and still is) a source of vivid controversies [21] and open debate.

As is well known, one of the main difficulties which arises when studying PT and molecular dissociation is due to the fact that the protolysis reaction is an extremely rare phenomenon. However, Saitta, Saija and Giaquinta [11] have shown that it is possible to stimulate the PT process – and, hence, to investigate protolysis in a more systematic way – by applying an external electric field. In this respect, the tools provided by theoretical first-principles approaches are very well suited to tackling and treating chemical reactions which occur at such length and time scales. By directly following the atomistic dynamical evolution, it is possible to trace the relevant physical quantities and the local structural changes which assist PT and molecular dissociation.

An information that can be readily earned upon applying sufficiently strong electric fields is the effective threshold associated with PT and molecular dissociation, respectively. In liquid water fields as intense as 0.25 V/Å are needed to cleave covalent bonds, whereas intensities equal or higher than 0.30 V/Å are needed for the same purpose in liquid methanol. Despite of this slight difference, which can be qualitatively correlated with the corresponding difference in the dissociation constants of the two substances, more important insights can be gained through the analysis of the (protonic) conductive regimes. In order to establish a measurable protonic current both in water and in methanol, a field intensity of 0.35 V/Å must be applied. As shown in Refs. [11, 22], a feeble protonic current has been recorded above this field strength threshold. A related ionic conductivity of $7.8 \, S \cdot cm^{-1}$ was calculated in liquid water, a value that is more than one order of magnitude larger than that calculated for methanol (i.e., $0.4 \, S \cdot cm^{-1}$). Such a significant difference in the ionic conduction displayed by water and methanol can be explained as follows: on one hand, each water molecule is statistically bonded to almost 4 other water molecules, whereas each methanol molecule is bonded, on average, to 2 methanol molecules only. A lower chance of proton migration is the outcome of the under-coordinated state of the methanol molecule. On the other hand, some more specific considerations on the charge carriers active in pure water and methanol are also in order. Indeed, although the ionic complexes associated with a cooperative PT in water and in methanol are the Zundel ($H_5O_2^+$) and Zundel-like ($[C_2H_6O_2H_3]^+$) ions (see Fig. 16.5a, b), the microscopic ionic defects that are responsible for charge transport in such two systems are the hydronium and the methyloxonium ions (i.e., H_3O^+ and $CH_3OH_2^+$), respectively. These short-lived ionic species are respectively threefold and twofold coordinated, via H-bond links, to their own local environment. Since for moderate-to-strong field intensities, higher than 0.35 V/Å, the H-bonded network is clearly perturbed, being characterized by lower average life times [22], the fluctuations of the average H-bond number around each molecule increase. In liquid water this circumstance leads to an increase of the PT probability as a result of the increased H-bond fluctuations in the first solvation shell of the H_3O^+ species. In fact, the first coordination shell of
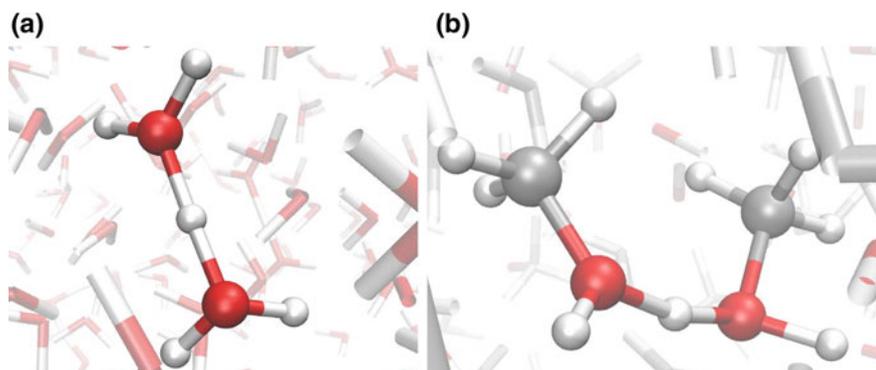
Fig. 16.5 The Zundel ion $H_5O_2^+$ formed **a** by two water molecules with an evenly shared proton and its variant (**b**), observed during proton transfer events, in methanol (two methanol molecules with an evenly shared proton form the ionic complex $[C_2H_6O_2H_3]^+$)

H_3O^+ includes, on average, three water molecules. The formation of a positive ion implies that the coordination number of a water molecule must change from a value of almost 4 (neutral molecule) to 3 (cation), an event that is statistically amplified by the external electric field. In methanol the increase of the rate of H-bond breaking and reforming processes (i.e., of fluctuations in the H-bond network) due to the field cannot lead to a PT enhancement comparable with that of water. In fact, the average coordination number of a neutral methanol molecule is 2 and the same value is obviously found for the protonated species $CH_3OH_2{}^+$. Hence, larger H-bond fluctuations do not give rise to a significant increase of the PT activity as observed in liquid water. This is the additional reason why in methanol just a feeble protonic current
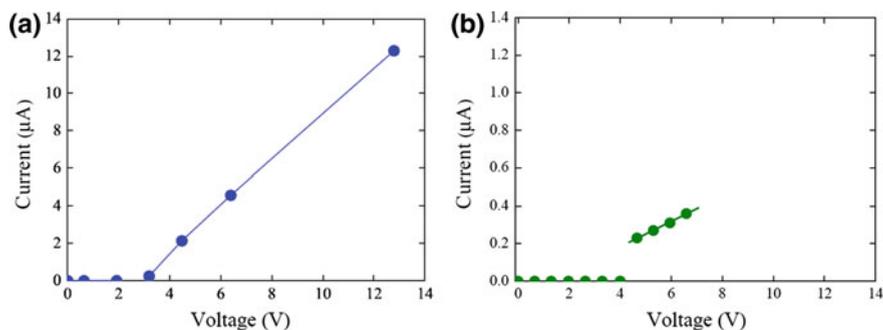
Fig. 16.6 Ionic current–voltage characteristics for cubic cells of comparable sides **a** in water and **b** methanol. The proton conduction process is characterized by an ohmic behavior provided that the current threshold has been overcome. The *dots* represent the calculated values, while the *solid lines* are just guides for the eye. Note that the current intensity scale of panel **b** has been reduced by one order of magnitude with respect to that for water in panel **a**. The estimated protonic conductivities of liquid water and methanol are 7.8 and $0.4\,S{\cdot}cm^{-1}$, respectively

eventually flows even for intense field strengths. In fact, although the onset of a net proton flow has been observed for the same field intensity recorded for liquid water (0.35 V/Å) [11, 22], the efficiency of the process is extremely lower than in water: very weak currents of the order of tenths of μA are induced by fields with a strength of $0.35-0.50$ V/Å, as shown in the current–voltage diagrams of Fig. 16.6.

We finally note that the current–voltage characteristics of both samples show an (ionic) ohmic response once a net proton flow has been activated.

Strong Field Regime: Field-Induced Chemical Reactions

In the previous section, we have already discussed the capability of an external field of driving some "simple" chemical reactions through a local rearrangement of covalent bonds. In fact, electric fields are able to trigger, drive, or promote even very complex reactions, as recently shown, for instance, for a specific Diels–Alder reaction [4]. Actually, in 1953 Stanley Miller reported a series of experiments in which he had observed the spontaneous formation of aminoacids from a mixture of simple molecules reacting under an electric discharge [23]. By all means, this finding gave birth to a new multidisciplinary research field: prebiotic chemistry. Hence, the role played by electric fields in increasing the reactivity of atoms in matter had been already ascertained more than 60 years ago. However, the chemical reactions leading to the formation of aminoacids, as those which occur in Miller's experiments, have never been studied at the atomic level before the advent of the extremely powerful computing techniques that have become available in more recent times. Again, ab initio methods have proven to be particularly suited for this purpose: in 2014 the first in silico Miller experiments have been reported [24], showing for the first time the active role of molecules such as formic acid and formamide as direct precursors of the simplest aminoacid glycine. In this pioneering study, several cubic supercells were set up in order to study and amplify the different reaction steps, as long as they were successively observed. All systems were studied in the canonical NVT ensemble at a constant temperature of 400 K, with a cubic cell side chosen so as to obtain a density of 1.0 g/mL, under the action of progressively stronger external electric fields, up to intensities of the order of about 0.5 V/Å. As pointed out in the previous sections, such high values of the field intensity are needed to dissociate water molecules within the picosecond time-scale of AIMD [11, 15, 16], and are compatible with the experimental values of water dissociation in the presence of electrical discharges [25–27] and, thus, with the presumed conditions of Miller's experiments.

Moreover, the same order of magnitude (and even higher) has been recorded in the naturally occurring electric fields at the surface of minerals [28, 29], also involved in the prebiotic geochemical reactions and in solutions with solvated ionic species [30–32]. Upon choosing appropriate proportions of the atoms composing the different numerical samples so as to mimic different stages of the Miller–Strecker reactions, the expected synthesis of glycine has been observed. In particular, it has
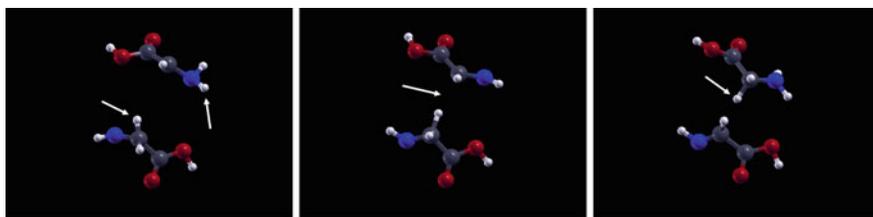
Fig. 16.7 Trajectory snapshots of the electric-field-driven evolution of dehydroglycine into glycine. At first a H_2 molecule splits and the two H atoms bond to the α-carbon atom and to the nitrogen atom of two nearby dehydroglycine molecules, respectively (*left panel*, shown by *arrows*); after about 90 fs the two α-carbon atoms of both molecules approach (*middle panel*, indicated by the *arrow*), and the extra proton jumps from one of the two molecules onto the other one, thus producing glycine (*right panel*, indicated by the *arrow*)

been demonstrated that, starting from water, ammonia, methane, carbon monoxide, and molecular nitrogen, the synthesis of glycine can be achieved with a field intensity of 0.5 V/Å. Complex chemical reactions have been observed even for lower strengths (0.35 V/Å) in such numerical experiments, in which formic acid and formamide molecules have been produced. In such a way, the atomic mechanistic aspects of the synthesis of the simplest building blocks leading to the emergence of life have been revealed for the first time. As shown in Fig. 16.7, the final steps of the formation of glycine have been recorded just after the formation of dehydroglycine from the intermediate molecules that lead to the fundamental carbon-to-carbon bond. The just described field-driven chemical reactions are the top of the iceberg of the full chemical landscape that the application of a static electric field can produce when it is applied to a material. As an example, even in the methanol case molecular dissociations are not the only chemical reactions which take place. Indeed, depending on the field strength, a progressive complexification of the system has been recently emphasized [33].

Another "textbook" chemical reaction has been observed and analyzed with various DFT tools and advanced metadynamics methods [34, 35]. In particular, the following nominal reaction has been revealed at ambient temperature (i.e., at 300 K), for field intensities above 0.5 V/Å,

$$2CH_3OH \xrightarrow{\text{E}} H_2CO + CH_4 + H_2O \,, \tag{16.3}$$
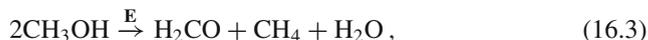
which, leading to the formation of formaldehyde, methane, and water, represents the disproportionation of the methanol molecule. This rather simple chemical reaction has never been observed before and represents a clear example of the way *computational experiments* can pave the way to *laboratory experiments*.

A final remark on the action of an external electric field is now in order. In the reactions described above the mere presence of reactive charged species in solution is not sufficient. This aspect has been carefully checked by running several zero-field simulations of the same initial set of Miller molecules, but replacing all water and ammonia molecules with their ionic OH^- and NH_4^+ counterparts. Although the

proportion of ionic species was thus much larger than in any instantaneous configuration ever observed under a nonzero finite field, no reactions other than proton jumps occurred, until the full neutralization of all molecules had been achieved, which proves the indispensable role of the external field in Miller's reactions. In a way, this finding also confirms that the electric field is not just a mere source of chemical energy and that it plays an "order-maker" role that inherently favors the assembling of larger chemical units from smaller ones and, thus, complexity. This crucial aspect has been confirmed also for the (maybe simpler) methanol case. Moreover, as checked in detail for the chemical reaction observed in liquid methanol, it has been argued that, by switching off the external field once the first reaction products have been formed in the system, they were stable. In particular, by performing a dynamics longer than 20 ps, during which the (de)protonated methanol species neutralize by recombining (i.e., via several proton transfers), persistent amounts of formaldehyde, methane, and water have been found in the sample. It has also been proven that switching on an external electric field drastically modifies the reaction network of methanol, lowering some activation barriers, stabilizing the reaction products, and opening otherwise difficult to achieve chemical routes [33].

Conclusions

In this chapter, we have reviewed the phenomenological evidence that has recently emerged from state-of-the-art *ab initio* simulations on the structural and transport properties of complex molecular systems under the action of a static and homogeneous electric field. The old and challenging problem of treating in detail the interaction between matter and intense electric fields is currently developing in an impressive way thanks to the growth of computational resources and of the predictive character acquired, at a semi-quantitative level, by first-principle numerical calculations. Just a couple of years ago applying these numerical methods to such tough problems, closely related to biophysics and biochemistry, as those discussed in this chapter was almost unthinkable.

References

1. N.J. English, C.J. Waldron, Phys. Chem. Chem. Phys. **17**, 12407 (2015). DOI 10.1039/C5CP00629E
2. D.I. de Pomerai, B. Smith, A. Dawe, K. North, T. Smith, D.B. Archer, I.R. Duce, D. Jones, E.P.M. Candido, FEBS Letters **543**(1–3), 93 (2003). DOI 10.1016/S0014-5793(03)00413-7
3. M. Porcelli, G. Cacciapuoti, S. Fusco, R. Massa, G. d'Ambrosio, C. Bertoldo, M.D. Rosa, V. Zappia, FEBS Letters **402**(2–3), 102 (1997). DOI 10.1016/S0014-5793(96)01505-0
4. A.C. Aragonés, N.L. Haworth, N. Darwish, S. Ciampi, N.J. Bloomfield, G.G. Wallace, I. Diez-Perez, M.L. Coote, Nature **531**(7592), 88 (2016). DOI 10.1038/nature16989
5. D. Marx, J. Hutter, *Ab initio Molecular Dynamics. Basic Theory and Advanced Methods* (Cambridge University Press, Cambridge, 2009)
6. P. Umari, A. Pasquarello, Phys. Rev. Lett. **89**, 157602 (2002). DOI 10.1103/PhysRevLett.89.157602
7. M.V. Berry, Proc. R. Soc. Lond. A **392**(1802), 45 (1984). DOI 10.1098/rspa.1984.0023
8. R.D. King-Smith, D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993). DOI 10.1103/PhysRevB.47.1651
9. R. Resta, Rev. Mod. Phys. **66**, 899 (1994). DOI 10.1103/RevModPhys.66.899
10. G. Desiraju, J. Vittal, A. Ramanan, *Crystal EngignNeering. A textbook* (World Scientific, Singapore, 2011)
11. A.M. Saitta, F. Saija, P.V. Giaquinta, Phys. Rev. Lett. **108**(20), 207801 (2012). DOI 10.1103/PhysRevLett.108.207801
12. G. Cassone, F. Creazzo, P.V. Giaquinta, F. Saija, A. Marco Saitta, Phys. Chem. Chem. Phys. **18**, 23164 (2016). DOI 10.1039/C6CP03926J
13. R. Car, M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985). DOI 10.1103/PhysRevLett.55.2471
14. S.J. Suresh, A.L. Prabhu, A. Arora, J. Chem. Phys. **126**(13), 134502 (2007). DOI 10.1063/1.2647105
15. G. Cassone, P.V. Giaquinta, F. Saija, A.M. Saitta, J. Phys. Chem. B **118**(16), 4419 (2014). DOI 10.1021/jp5021356
16. G. Cassone, P.V. Giaquinta, F. Saija, A.M. Saitta, J. Phys. Chem. B **118**(44), 12717 (2014). DOI 10.1021/jp507376v
17. E. Schwegler, G. Galli, F.m.c. Gygi, R.Q. Hood, Phys. Rev. Lett. **87**, 265501 (2001). DOI 10.1103/PhysRevLett.87.265501
18. K. Kaila, B.R. Ransom, *pH and brain function* (Wiley, New York, 1998)
19. E.I. Zoulias, N. Lymberoupolos, *Hydrogen-Based Autonomous Power Systems* (Springer, Berlin, 2008)
20. C.J.T. Grotthuss, Ann. Chim. **LVIII**, 54 (1806)
21. N. Agmon, J. Phys. Chem. A **109**(1), 13 (2005). DOI 10.1021/jp047465m
22. G. Cassone, P.V. Giaquinta, F. Saija, A.M. Saitta, J. Chem. Phys. **142**(5), 054502 (2015). DOI 10.1063/1.4907010
23. S.L. Miller, Science **117**(3046), 528 (1953). DOI 10.1126/science.117.3046.528
24. A.M. Saitta, F. Saija, Proc. Natl. Acad. Sci USA **111**(38), 13768 (2014). DOI 10.1073/pnas.1402894111
25. Z. Hammadi, J.P. Astier, R. Morin, S. Veesler, Cryst. Growth Des. **7**(8), 1472 (2007). DOI 10.1021/cg070108r
26. Z. Hammadi, J.P. Astier, R. Morin, S. Veesler, Cryst. Growth Des. **9**(8), 3346 (2009). DOI 10.1021/cg900150n
27. Z. Hammadi, M. Descoins, E. Salançon, R. Morin, Appl. Phys. Lett. **101**(24), 243110 (2012). DOI 10.1063/1.4770516
28. S. Laporte, F. Finocchi, L. Paulatto, M. Blanchard, E. Balan, F. Guyot, A.M. Saitta, Phys. Chem. Chem. Phys. **17**, 20382 (2015). DOI 10.1039/C5CP02097B
29. L. Vlcek, Z. Zhang, M.L. Machesky, P. Fenter, J. Rosenqvist, D.J. Wesolowski, L.M. Anovitz, M. Predota, P.T. Cummings, Langmuir **23**(9), 4925 (2007). DOI 10.1021/la063306d

30. Y. Bronstein, P. Depondt, L.E. Bove, R. Gaal, A.M. Saitta, F. Finocchi, Phys. Rev. B **93**, 024104 (2016). DOI 10.1103/PhysRevB.93.024104
31. B. Reischl, J. Köfinger, C. Dellago, Mol. Phys. **107**(4-6), 495 (2009). DOI 10.1080/00268970902865493
32. B. Sellner, M. Valiev, S.M. Kathmann, J. Phys. Chem. B **117**(37), 10869 (2013). DOI 10.1021/jp405578w
33. G. Cassone, F. Pietrucci, F. Saija, F. Guyot, A.M. Saitta, (2016). Submitted
34. A. Laio, M. Parrinello, Proc. Natl. Acad. Sci. USA **99**(20), 12562 (2002). DOI 10.1073/pnas.202427399
35. F. Pietrucci, A.M. Saitta, Proc. Natl. Acad. Sci. USA **112**(49), 15030 (2015). DOI 10.1073/pnas.1512486112

Chapter 17
Miniature Spherical Sapphire Anvil Cell for Small Angle Neutron Scattering

X. Wang, N.A. Parzyk, D.M. Paul, C.D. Dewhurst, G. Giriat and K.V. Kamenev

Abstract We report the design of a compact sapphire anvil cell for small angle neutron scattering. Based on turnbuckle-opposed anvil technique, the dimensions of this cell are minimized for use with a small-bore magnets using finite element analysis. This cell is the smallest sapphire anvil cell to date. The cell body is approximately 14 mm diameter cylinder and 13 mm length, the overall length of the assembled cell is less than 17 mm. Inside the cell, two spherical-shaped sapphire anvils are used to ensure compactness and anvil's strength, and at the same time to provide sufficient sample volume for a neutron scattering experiment. The sample volume of the cell used in the SANS experiment was as large as 0.4 mm^3, which is approximately 40 times larger than the conventional diamond anvil cell. The cell has been used in a SANS study of a niobium single crystal at D22 Institut Laue-Langevin with the maximum pressure of 4 GPa being achieved with the liquid pressure medium. During the tests the cell demonstrated that it was capable of achieving pressure to 6 GPa with a powder sample.

Introduction

Small angle neutron scattering (SANS) is a powerful neutron technique that uses elastic neutron scattering at small scattering angles to study material structure at length scale from 1–1000 nm, which can be coupled with auxiliary equipment to study

X. Wang · K.V. Kamenev (✉)
Centre for Science at Extreme Conditions and School of Engineering,
University of Edinburgh, Edinburgh EH9 3JZ, UK
e-mail: K.Kamenev@ed.ac.uk

N.A. Parzyk · D.M. Paul
Department of Physics, University of Warwick, Coventry CV4 7AL, UK

C.D. Dewhurst
Institut Laue-Langevin, 6 Rue Jules Horowitz, 38042 Grenoble, France

G. Giriat
Laboratory of Nanostructures and Novel Electronic Materials, École Polytechnique
Fédérale de Lausanne, 1015 Lausanne, Switzerland

crystallographic and magnetic properties of materials at the extremes of temperature, field, and pressure. Recently, SANS experiments combined high pressure became a fast growing field of study in material science [1–3], bimolecular systems and soft condensed matter [4–8]. Conventional piston cylinder high pressure cells are not ideal for SANS as the beam passing through the metallic cell body would result in the micro structure of high pressure alloy interfering with the scattering from the sample [7]. Though diamond anvil cells (DAC) had been used in neutron diffraction experiments since 1980s [9], small sample size results in unreasonably long data collection times.

Single crystal sapphires are widely used in construction of high pressure cells for SANS due to the material's high hardness, good neutron transmission, low incoherent background scattering, and optical transparency. Apart from that, large sample volumes can be achieved in cells using sapphire. All but one of the existing high pressure cells for SANS can be categorized as sapphire windows cells [3, 4, 6, 8]. This type of cell is designed to generate pressure less than 500 MPa with large sample volume. The external compressor can be used to provide pressure jump techniques which enable researchers to explore the kinetics of phase behavior and structure formation in complex matter. Opposed anvil cells can generate much higher pressures. Bonetti and Calmettes [5] developed the first sapphire anvil cell for SANS in 2005. The pressure capability is still limited to 530 MPa as large culet sapphire anvils were used. Apart from limited pressure capability, the existing high pressure cells developed for SANS are also massive and complicated as most of loading methods involve use of a hydraulic system. The aim of this work is to design a cell which has higher pressure capability and large enough sample volume for SANS study. The operation of the cell needs to be simple and user friendly. Most importantly for studies of magnetic properties of materials, the dimensions of such a cell needs to be minimised for use with a magnet/cryostat on a SANS beamline.

Design

Cell Structure

One of the smallest existing designs of a pressure cell is that of a turnbuckle diamond anvil cells were pioneered by S.W. Tozer [10–14] initially for use with high pulsed magnetic fields. The design of our cell is shown in Fig. 17.1. It is based on a scale-up version of a previous turnbuckle magnetic diamond anvil cell (TM-DAC) reported by Giriat et al. [15]. The cell consists of two end-nuts, two spherical sapphire anvils and a cell body. The end-nuts are identical but with external M10×0.5 threads cut in opposite direction. These match the internal threads at the top and the bottom of the cell body. Fine resolution of anvil displacement is assured by a fine custom-cut M10×0.5 thread. All the parts are machined out from hardened nonmagnetic beryllium copper alloy (BERYLCO-25), which has a tensile strength of 1.4 GPa. Central
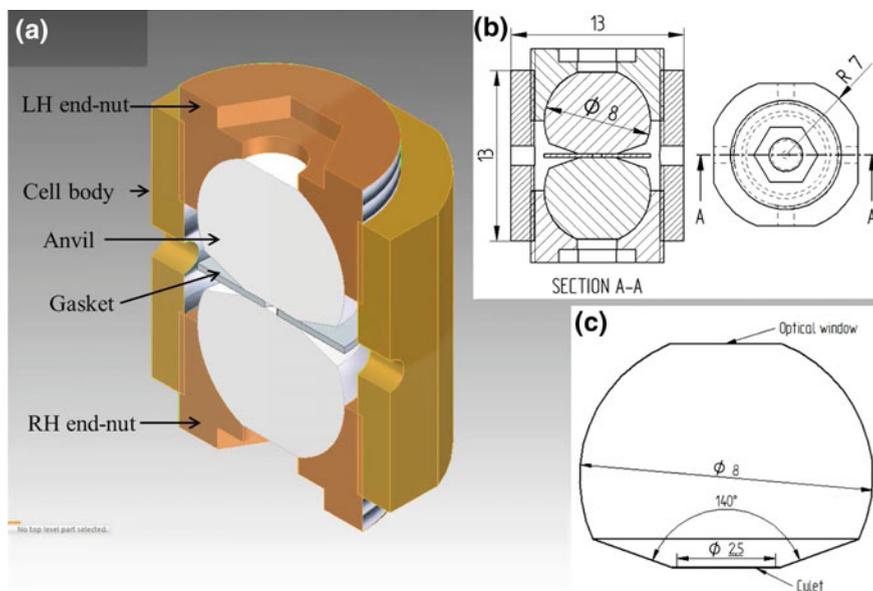
Fig. 17.1 **a** Cross-section view of the assembled pressure cell (LH and RH stand for *left*-hand and *right*-hand thread, respectively); **b** CAD drawing with key dimensions; **c** the dimensions of the spherical anvil

holes are drilled in both end-nuts to allow the neutron beam to pass through axially. In addition, the holes can be used for sample observation and pressure measurement by ruby fluorescence method [16]. The strength of the thread has been assessed by finite element analysis (FEA). The simulation presented in Fig. 17.2 shows the threads are capable of holding up to 20 kN load before the thread starts to yield.

Spherical Sapphire Anvil

Spherical sapphire anvils used in our cell have been used in several designs which demonstrated their performance and cost effectiveness [17–21]. Sapphire balls are widely available commercially and can be easily formed into anvils by grinding two parallel surfaces on them. One face is used as a culet to compress the sample and the gasket. Truncating the opposite end of the sphere eliminates the optical distortion from the curved surface, thus enabling the clear optical access for alignment and pressure measurement by ruby fluorescence. The highest sample pressure of 12.6 GPa was reported by Takano and Wakatsuki [18] with 12.5 kN load. This was achieved using a pair of 3/8 in. sapphire balls with the culet diameter of 1.5 mm and the sample volume being around 10^{-2} mm^3. As shown in Fig. 17.1, we employ two 8 mm diameter spherical sapphire anvils in our cell. To improve the compactness of
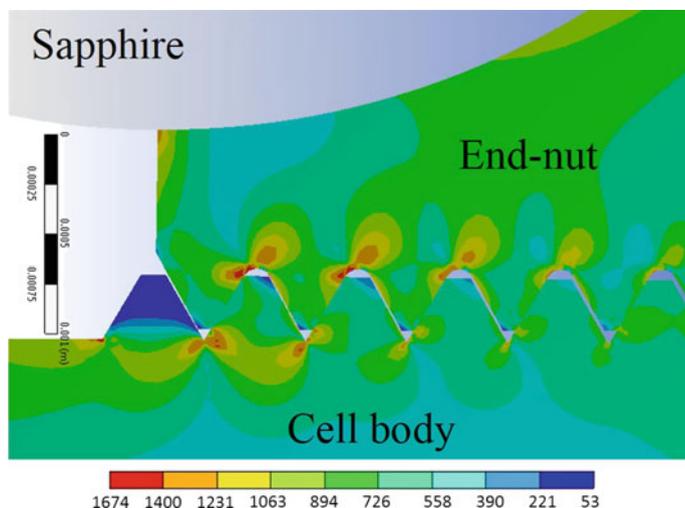
Fig. 17.2 FEA stress analysis shows that the first four turns of the thread is the main stress concentration area. When the anvil is subjected to 20 kN load on the culet, Von-Mises stress becomes equal to the yield strength of the material
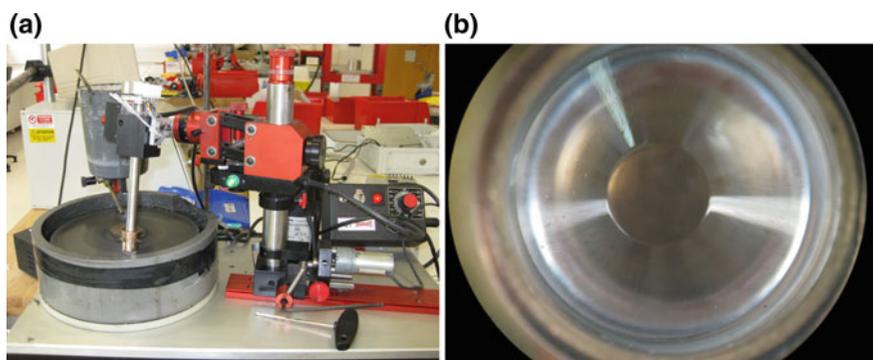
Fig. 17.3 **a** In-house polisher modified to grind conical bevels on the sapphire anvil; **b** surface finish of the beveled spherical sapphire anvil

the cell, 140° bevel angle is lapped on the anvil with in-house polisher as shown in Fig. 17.3a. The sapphire ball is mounted on the swing arm and ground against a spinning disk. Slurry which contains diamond powder was used on the disk to improve the surface finish. To form the bevel, modification was made on the original polisher. A motor was installed on the arm (Fig. 17.3a) to rotate the sapphire around its own axis for lapping the bevel. The culet is polished to 2.5 mm diameter. Another flat surface is lapped on the opposite site of the anvil for optical observation and pressure measurement. As ruby fluorescence method was used to measure sample pressure, the chromium impurity in the ball is reduced to less than 1 ppm by the supplier.

The surface finish on the culet plays a critical role for successful loading. Scratches on the culet would behave as a stress riser which would significantly compromise the strength of the anvil and lead to premature surface failure [22]. To avoid this, the surface finish of the culet needs to be enhanced as much as possible. Visible scratches observed on the culet were manually polished off by a soft cloth impregnated with $1\,\mu m$ diamond paste. Figure 17.3b shows no visible scratches with $4\times$ magnification under microscope.

Loading Test

To assess the performance of the design, several copies of the pressure cell were manufactured for loading and pressure testing (Fig. 17.4a). Similarly to the TM-DAC [15], a bracket and two supports (Fig. 17.4b) are needed to apply load onto the cell as shown in Fig. 17.4c. The pressurization of the cell requires a hydraulic press. In addition, a load cell is placed under the bracket to monitor the load. The load is applied to move the anvils toward each other and to reduce friction within the threads in order to lock the cell. The cell is clamped by rotating the body using a spanner. Sample pressure can be checked after removing the cell from the hydraulic press and the bracket.
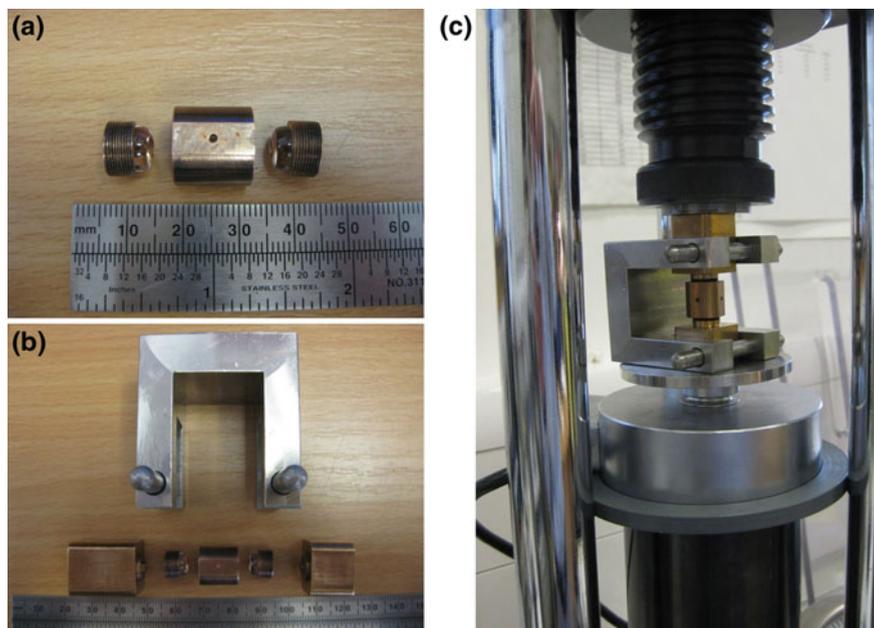
Fig. 17.4 **a** Key parts of the cell; **b** disassembled cell with the bracket and supports; **c** figure of the cell in a loading test, a load cell is placed beneath the bracket to monitor the compressive force from the press
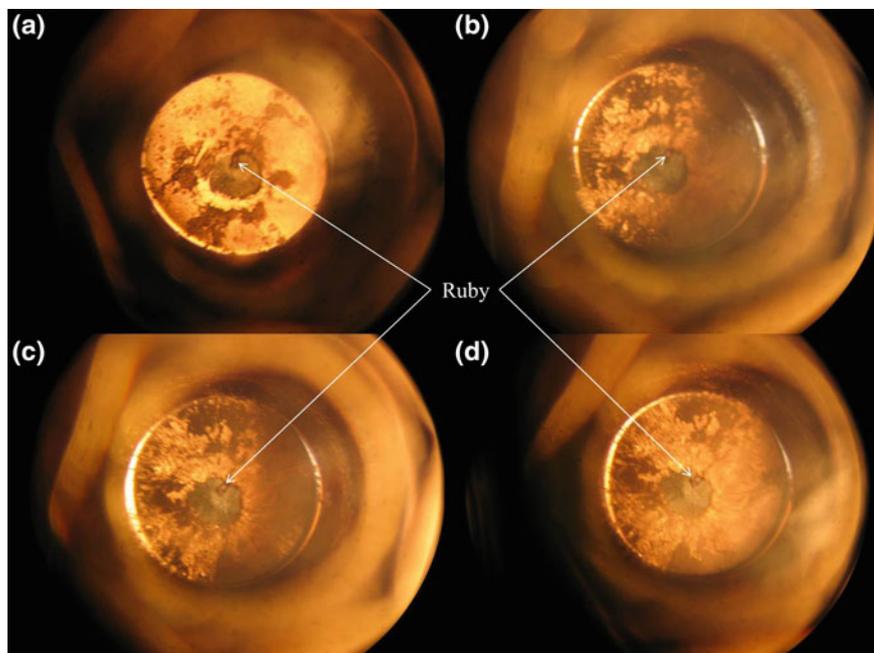
Fig. 17.5 Loading test with a copper gasket which was pre-indented to 220 μm thick; the final thickness varies from 60 to 40 μm; **a** 1.4 GPa with 9.84 kN load; **b** 4.2 GPa with 13.58 kN load; **c** 5.8 GPa with 17.33 kN load; **d** 6.4 GPa with 19.6 kN load

Soft nonmagnetic metal gaskets (copper or unhardened BeCu) were used in tests as harder gaskets were found to cause premature failure. This observation is consistent with the previously reported use of the sapphire anvil cell [17–21]. Hard gasket can induce high shear stress on the culet, which usually lead to surface damages of sapphire anvil (e.g., cracks, scratches, and chippings). In the test shown in Fig. 17.5, the thickness of the gasket was pre-indented from 500 to 220 μm. A 0.6 mm diameter sample hole was drilled in the centre of the gasket to accommodate NaCl powder pressure medium and a ruby chip. With these parameters, the initial sample volume is larger than 0.06 mm^3. In the loading test, the cell was tested close to the maximum allowable load of the cell (as predicted by FEA modeling) and no anvil failure was observed. The highest generated pressure was 6.4 GPa with 19.6 kN load. Figure 17.5 shows the deformation of the sample hole with increasing pressure. The final gasket thickness was between 40–60 μm, which can be used as the means of assessing the quality of the anvil alignment and the taper angle between the culets. The test result shows the attainable pressure of our cell is much higher than the existing pressure cells developed for SANS. Moreover, the anvils demonstrate better performance than the cell reported by Takano [18] by holding 6.4 GPa with similar sample volume.

However, the sample size is limited for the copper gasket because of the thinning of the gasket. Thus, only powder sample or small single crystal can be placed into copper gasket to avoid crushing.

High Pressure SANS Experiment on Single Crystal Niobium

This section demonstrates an application of the cell in high pressure SANS experiment (Fig. 17.6). A single crystal of niobium was loaded into the cell to study the pressure dependence of the flux line lattice (FLL) of the crystal by using SANS technique. The sample was around 1 mm long and 0.3 mm thick as shown in Fig. 17.6a. Accordingly, a thick gasket is needed for this experiment to suit the height of the sample. Copper gasket cannot be used in this case due to its low strength and thinning under high pressure. To contain the sample, a BeCu gasket with 1 mm diameter sample hole and 0.5 mm thick was used. The initial sample volume was equivalent to 0.4 mm^3. Mixture of Fluorinert (FC 72: FC 84) was used as the pressure medium to create the hydrostatic pressure environment. As the sample hole is larger than those in previous tests, the culets of the anvils were polished to 3 mm diameter accordingly to prevent the sample hole from blowing out in the early pressurization stage.

In the experiment, the highest attainable pressure was 4 GPa with 17 kN load. The pressurized cell was mounted on the aluminium sample holder as shown in Fig. 17.6b. To reduce the scattering background, the cell was covered with neutron absorbing cadmium foil. The pressurized cell was loaded into a cryomagnet [23] on
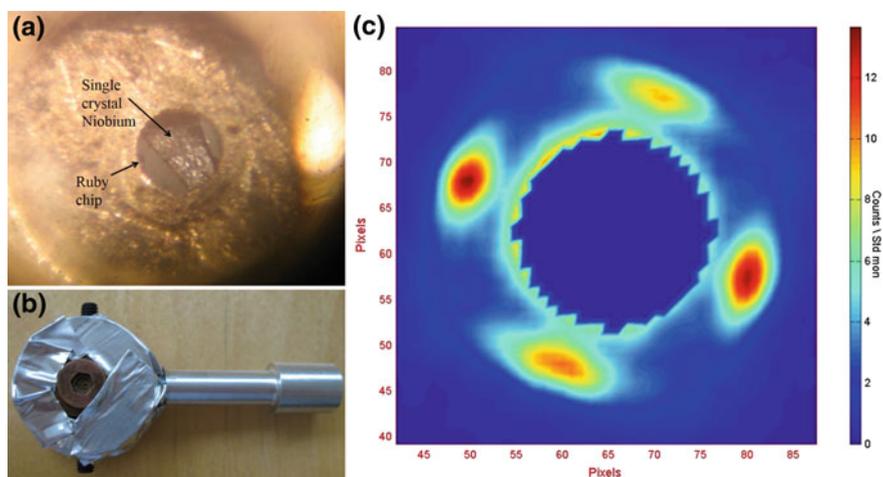
Fig. 17.6 SANS experiment at D22, ILL: **a** the single crystal Niobium in the gasket; **b** the loaded cell covered by cadmium foil; **c** SANS neutron data with flux line lattice appearing as four *red* spots

D22 beamline, Institut Laue–Langevin (ILL). The cell was cooled down to 1.8 K and Fig. 17.6c shows the collected data after the cell was exposed to the neutron beam for 10 min.

Summary

We have developed a compact turnbuckle sapphire anvil cell for SANS experiment. The dimensions of our cell is even smaller than the "egg" cell which was considered as the smallest sapphire cell reported by Goncharenko [9]. This cell can be used with any existing cryomagnet even with a small bore [23] and the cooling rate is much higher due to its small size. The compactness, non-magnetic body and large sample volume (up to $0.4\,mm^3$) facilitate the potential implementation with other experimental techniques. In this design, the weakest part of this cell was found to be the sapphire anvils. It is worthwhile to conduct investigation into how to improve the durability of the spherical sapphire anvil, such as high temperature annealing [22, 24, 25], shape optimization, etc.

References

1. N. Vennemann, M.D. Lechner, R.C. Oberthür, Polymer **28**(10), 1738 (1987). DOI 10.1016/0032-3861(87)90018-8
2. H. Takeno, M. Nagao, Y. Nakayama, H. Hasegawa, T. Hashimoto, H. Seto, M. Imai, Polym. J. **29**(11), 931 (1997). DOI 10.1295/polymj.29.931
3. G. Pépy, P. Baroni, J. Appl. Cryst. **36**(3 Part 1), 814 (2003). DOI 10.1107/S0021889803003935
4. M. Bonetti, P. Calmettes, Rev. Sci. Instrum. **75**(2), 440 (2004). DOI 10.1063/1.1637434
5. M. Bonetti, P. Calmettes, Rev. Sci. Instrum. **76**(4), 043903 (2005). DOI 10.1063/1.1884325
6. A. Gabke, J. Kraineva, R. Köhling, R. Winter, J. Phys.: Cond. Matter **17**(40), S3077 (2005). DOI 10.1088/0953-8984/17/40/011
7. C. Pfleiderer, A.D. Huxley, S.M. Hayden, J. Phys.: Cond. Matter **17**(40), S3111 (2005). DOI 10.1088/0953-8984/17/40/014
8. J. Kohlbrecher, A. Bollhalder, R. Vavrin, G. Meier, Rev. Sci. Instrum. **78**(12), 125101 (2007). DOI 10.1063/1.2817632
9. I.N. Goncharenko, High Press. Res. **24**(1), 193 (2004). DOI 10.1080/08957950410001661882
10. C. Martin, C.C. Agosta, S.W. Tozer, H.A. Radovan, T. Kinoshota, M. Tokumoto, J. Low Temp. Phys. **138**(5), 1025 (2005). DOI 10.1007/s10909-004-2898-8
11. M. Kano, N. Kurita, M. Hedo, Y. Uwatoko, S.W. Tozer, H.S. Suzuki, T. Onimaru, T. Sakakibara, J. Phys. Soc. Jpn. **76**(Suppl. A), 56 (2007). DOI 10.1143/JPSJS.76SA.56

12. D.E. Graf, R.L. Stillwell, K.M. Purcell, S.W. Tozer, High Press. Res. **31**(4), 533 (2011). DOI 10.1080/08957959.2011.633909
13. M. Kano, H. Mori, K. Matsubayashi, M. Itoi, M. Hedo, T.P. Murphy, S.W. Tozer, Y. Uwatoko, T. Nakamura, J. Phys. Soc. Jpn. **81**(2), 024716 (2012). DOI 10.1143/JPSJ.81.024716
14. L. Chen, N. Kurita, M. Hedo, K. Nakazawa, I. Oguro, Y. Uwatoko, T. Matsumura, M. Tokumoto, S.W. Tozer, J. Phys. Soc. Jpn. **76**(Suppl. A), 58 (2007). DOI 10.1143/JPSJS.76SA.58
15. G. Giriat, W. Wang, J.P. Attfield, A.D. Huxley, K.V. Kamenev, Rev. Sci. Instrum. **81**(7), 073905 (2010). DOI 10.1063/1.3465311
16. R.A. Forman, G.J. Piermarini, J.D. Barnett, S. Block, Science **176**(4032), 284 (1972). DOI 10.1126/science.176.4032.284
17. W.F. Kuhs, H. Ahsbahs, D. Londono, J.L. Finney, Physica B: Condens. Matter **156**, 684 (1989). DOI 10.1016/0921-4526(89)90763-1
18. K.J. Takano, M. Wakatsuki, Rev. Sci. Instrum. **62**(6), 1576 (1991). DOI 10.1063/1.1142435
19. W.B. Daniels, M. Lipp, D. Strachan, D. Winters, Z. Yu, in *XIII AIRAPT International Conference on High Pressure Science and Technology* (1992), p. 809
20. W.F. Kuhs, F.C. Bauer, R. Hausmann, H. Ahsbahs, R. Dorwarth, K. Hölzer, High Press. Res. **14**(4–6), 341 (1996). DOI 10.1080/08957959608201420
21. W.F. Kuhs, F.C. Bauer, H. Ahsbahs, G.J. McIntyre, Rev. High Press. Sci. Tech. **7**, 307 (1998). DOI 10.4131/jshpreview.7.307
22. E.R. Dobrovinskaya, L.A. Lytvynov, V. Pishchik, *Sapphire Material, Manufacturing, Applications* (Springer, New York, 2009)
23. A.T. Holmes, G.R. Walsh, E. Blackburn, E.M. Forgan, M. Savey-Bennett, Rev. Sci. Instrum. **83**(2), 023904 (2012). DOI 10.1063/1.3688657
24. S. Klotz, *Techniques in high pressure neutron scattering* (CRC Press, New York, 2012)
25. M. Eremets, *High Pressure Experimental Methods* (Oxford University Press, Oxford, 1996)

Chapter 18
Are Light Alkali Metals Still Metals Under High Pressure?

F. Siringo, R. Pucci and G.G.N. Angilella

Abstract The extended Hubbard Hamiltonian on a bcc lattice is studied at half-filling and for a finite hopping between next-nearest neighbours, in mean-field approximation. An ionic insulating broken-symmetry phase is predicted for any hydrogenoid bcc solid in the density range $1.0 < r_s < 2.6$. The occurrence of an ionic phase would explain the failure to achieve hydrogen metallization at high pressures. Moreover, a metal–insulator transition is expected for sodium in the 100 GPa region.

Alkali metals crystallize in a bcc phase under ordinary thermodynamic conditions. In 1935, Wigner and Huntington [1] proposed that even hydrogen should undergo a metal–insulator (MI) transition from a molecular to a monatomic bcc crystal, under high pressure in analogy to alkali metals. During the last years new excitement arose after several claims for the reach of hydrogen's metallization [2–6]. However, it is

F. Siringo (✉) · R. Pucci · G.G.N. Angilella
Dipartimento di Fisica e Astronomia, Università di Catania, Via S. Sofia, 64, 95123 Catania, Italy
e-mail: fabio.siringo@ct.infn.it

R. Pucci
e-mail: renato.pucci@ct.infn.it

G.G.N. Angilella
e-mail: giuseppe.angilella@ct.infn.it

F. Siringo · G.G.N. Angilella
INFN, Sezione di Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci · G.G.N. Angilella
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci · G.G.N. Angilella
IMM-CNR, UdR Catania, Via S. Sofia 64, 95123 Catania, Italy

G.G.N. Angilella
Scuola Superiore di Catania, Università di Catania, Via Valdisavoia, 9, 95123 Catania, Italy

now out of doubt that all the observed phases of hydrogen are still molecular, and there is no evidence of metallization induced by a band overlap mechanism up to 191 GPa [7]. From this point of view, hydrogen seems to be quite different from the halogens, for which MI transitions have been observed [8, 9].

The nature of the high pressure *A* phase of hydrogen still remains unexplained, and several hypotheses have been advanced. The recent proposal by Baranowski [10] that the hydrogen molecules may develop electric dipole moments is challenging for different reasons: (i) first of all, such a phase would be per se interesting, being a broken-symmetry ground state of the symmetric H_2 molecular system; (ii) in addition, the existence of such an ionic phase could move a possible monatomic phase further towards higher pressures; (iii) finally, given the similarity of hydrogen with the alkali metals, the possible existence of an ionic instability should show up even in lithium or sodium under the proper thermodynamic conditions.

Provided that such a broken-symmetry ionic crystal does exist, it would be desirable to fix the boundaries between the ionic and the monatomic bcc phases. Approaching the boundary from the bcc phase, we look for an ionic instability of the monatomic crystal, or, in other words, for a charge density wave (CDW) instability, commensurate with the cubic lattice. The existence of such a CDW has been recently observed in the ground state of sodium and potassium, under ordinary thermodynamic conditions [11, 12]. Such instabilities of the Fermi gas had been predicted [13] as a consequence of the Coulomb electron–electron (*e–e*) repulsion, but do not give rise to any MI transition, since the CDW is not commensurate with the lattice. Such very small effect does not prevent us from considering the alkali metals as simple 'free-electron' metals for most aspects. On the other hand, a nearest neighbour tight-binding model on a bipartite bcc lattice gives rise to a perfectly nested cubic Fermi surface at half-filling, and any small *e–e* repulsion would drive the system towards a spin density wave (SDW) or towards a CDW commensurate with the lattice.

In this Letter we show that, even without nesting of the Fermi surface, a hydrogenoid bcc crystal undergoes a MI transition towards a broken-symmetry commensurate CDW phase, for an appropriate bounded range of density values. Such a conclusion emerges from a careful analysis of the mean-field phase diagram for an extended Hubbard Hamiltonian, modified in order to take in due account the long range Coulomb interactions and the hopping between next-nearest neighbours. Even for a spherical Fermi surface, the model predicts the occurrence of a broken-symmetry insulating ground state, provided that the nearest neighbour repulsive interaction *V* exceeds some critical value. The latter is a function of the other energy scales and mainly of the on-site Hubbard repulsion *U* between two electrons sharing the same lattice site. While *U* is only slightly affected by any increase in density, *V* scales as a^{-1}, being *a* the cubic lattice spacing. Under high pressure, *V* may reach its critical value, giving rise to a MI transition, albeit in an intermediate density range; then, at very high densities, the large increase of the Fermi energy, scaling as a^{-2}, would eventually stabilize the monatomic phase. In other words, we expect that under high pressure both lithium and sodium should undergo a MI transition from a simple metal to an ionic insulator. On the other hand, such a high density instability of the bcc

crystal would suggest that, in order to stabilize a hydrogen monatomic phase, higher densities are required than previously estimated.

The extended Hubbard Hamiltonian reads as:

$$H = -t_1 \sum_{\langle ij \rangle \sigma} c_{i\sigma}^\dagger c_{j\sigma} - t_2 \sum_{\langle\langle ij \rangle\rangle \sigma} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} + \frac{1}{2} \sum_{i \neq j \, \sigma\sigma'} V_{ij} n_{i\sigma} n_{j\sigma'},$$

(18.1)

where $c_{i\sigma}$ ($c_{i\sigma}^\dagger$) denote the annihilation (creation) operators for an electron in the Wannier state centred on the ith site of a bcc lattice, with spin projection $\sigma \in \{\uparrow, \downarrow\}$ $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$, $t_1, t_2 > 0$, and $\langle ij \rangle$, $\langle\langle ij \rangle\rangle$ restricting the sums over nearest and next-nearest neighbour couples, respectively.

The Hamiltonian, Eq. (18.1), incorporates two major approximations: (i) it neglects all bond-bond and bond-ion interactions; (ii) it neglects any hopping term other than those between nearest or next-nearest neighbours. Approximating the Wannier states by atomic hydrogen ground state wave functions, for a density corresponding to $r_s \approx 2.2$, the larger bond-ion interaction term does not exceed the 30% of the corresponding ion–ion interaction. Of course, any extrapolation to higher densities would require some caution. Regarding the neglected hopping terms, we must notice that the insertion of such exponentially decreasing terms does not change the shape of the Fermi surface in a significant way. A ratio $t_2/t_1 \approx 0.6 \div 0.7$ allows for an almost spherical Fermi surface, up to a 3% deviation.

The model may be solved in mean-field (MF) approximation by inserting $\langle n_{i\sigma} \rangle = \frac{1}{2} + \Delta_{i\sigma}$ and neglecting second-order terms in the fluctuations $\delta n_{i\sigma} = n_{i\sigma} - \langle n_{i\sigma} \rangle$. Since we are looking for a commensurate CDW instability, we assume $\Delta_{i\sigma} = \Delta \cos(\mathbf{Q} \cdot \mathbf{R}_i)$, being $\mathbf{Q} = \left(\frac{2\pi}{a}, 0, 0\right)$ the exact nesting vector of the Fermi surface when $t_2 = 0$. Working in the reciprocal lattice, i.e., introducing:

$$c_{i\sigma}^\dagger = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_i} c_{\mathbf{k}\sigma}^\dagger,$$

(18.2)

with $\mathbf{k}$ summed over the N points inside the first Brillouin zone, and neglecting second-order fluctuation terms, the Hamiltonian Eq. (18.1) reads (up to a constant) as:

$$H_{\mathrm{MF}} = \sum_{\mathbf{k}\sigma} \varepsilon(\mathbf{k}) c_{\mathbf{k}\sigma}^\dagger c_{\mathbf{k}\sigma} + \Gamma \sum_{\mathbf{k}\sigma} c_{\mathbf{k}+\mathbf{Q}\sigma}^\dagger c_{\mathbf{k}\sigma} - N\Gamma\Delta,$$

(18.3)

where $\Gamma = \Delta(U - 16W)$, $\varepsilon(\mathbf{k}) = \varepsilon_1(\mathbf{k}) + \varepsilon_2(\mathbf{k})$, and:

$$\varepsilon_1(\mathbf{k}) = -4t_1 \cos\left(\frac{k_x a}{2}\right) \cos\left(\frac{k_y a}{2}\right) \cos\left(\frac{k_z a}{2}\right),$$

(18.4a)

$$\varepsilon_2(\mathbf{k}) = -t_2[\cos(k_x a) + \cos(k_y a) + \cos(k_z a)].$$

(18.4b)

Here W is a renormalized interaction parameter summing up all the long range Coulomb interactions:

$$W = \frac{1}{8} \sum_{m=1}^{\infty} (-1)^{m+1} z_m V^{(m)}, \tag{18.5}$$

where, for $m = 1, 2, \ldots \infty$, $V^{(m)}$ denotes the V_{ij} interaction term for nearest neighbours, next-nearest neighbours, etc., and z_m is the corresponding coordination number.

The MF Hamiltonian, Eq. (18.3), is easily diagonalized by a canonical transformation. Let us introduce the spinorial notation $\Psi_{\mathbf{k}\sigma}^{\dagger} = \left(c_{\mathbf{k}\sigma}^{\dagger}, \ c_{\mathbf{k}+\mathbf{Q}\sigma}^{\dagger} \right)$, with $\mathbf{k}$ restricted inside the cube $|k_\alpha| < \pi/a$ ($\alpha = x, y, z$), which is exactly half the first Brillouin zone. Actually the transformation $\mathbf{k} \to \mathbf{k} + \mathbf{Q}$ maps such reduced zone onto the complementary half-zone. The Hamiltonian Eq. (18.3) now reads:

$$H_{\mathrm{MF}} = \sum_{\mathbf{k}\sigma} \Psi_{\mathbf{k}\sigma}^{\dagger} h(\mathbf{k}) \Psi_{\mathbf{k}\sigma} - N\Gamma\Delta, \tag{18.6}$$

where the 2×2 matrix $h(\mathbf{k})$ is defined as:

$$h(\mathbf{k}) = \begin{pmatrix} \varepsilon_1(\mathbf{k}) + \varepsilon_2(\mathbf{k}) & \Delta(U - 16W) \\ \Delta(U - 16W) & \varepsilon_2(\mathbf{k}) - \varepsilon_1(\mathbf{k}) \end{pmatrix}, \tag{18.7}$$

and is promptly diagonalized yielding the spectrum:

$$E^{\pm}(\mathbf{k}) = \varepsilon_2(\mathbf{k}) \pm \sqrt{\varepsilon_1(\mathbf{k})^2 + \Gamma^2}. \tag{18.8}$$

A gap opens between the two bands, $E^{\pm}(\mathbf{k})$, whenever $\Gamma > 2t_2$: in such a regime, the system is an insulator, and the total energy E_{tot} is readily evaluated by summing E^- over the doubly occupied half-zone:

$$E_{\mathrm{tot}} = \sum_{\mathbf{k}\sigma} [\varepsilon_2(\mathbf{k}) - \sqrt{\varepsilon_1(\mathbf{k})^2 + \Gamma^2}] - N\Gamma\Delta. \tag{18.9}$$

A gap equation is obtained by differentiating E_{tot} with respect to the order parameter, Δ:

$$\frac{1}{16W - U} = \frac{a^3}{2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{1}{\sqrt{\varepsilon_1(\mathbf{k})^2 + \Gamma^2}}. \tag{18.10}$$

A finite Γ always solves the latter condition at any coupling strengths but, for sake of consistency with the above assumption of dealing with an insulating phase, Γ must exceed the critical value, $\Gamma_c = 2t_2$. A consistent minimum for the total energy is, e.g., found for $(16W - U)/t_1 > 3.84$ if $t_2/t_1 = 0.8$.

Fig. 18.1 $U-V$ phase diagram for a CDW instability, for $t_2/t_1 = 0.8$. The boundary between normal metal (M) and ionic insulator (I) is reported as a solid line. The *dashed line* represents the equation of state for a light alkali as hydrogen: the squares correspond to $r_s = 0.6 \div 2.8$
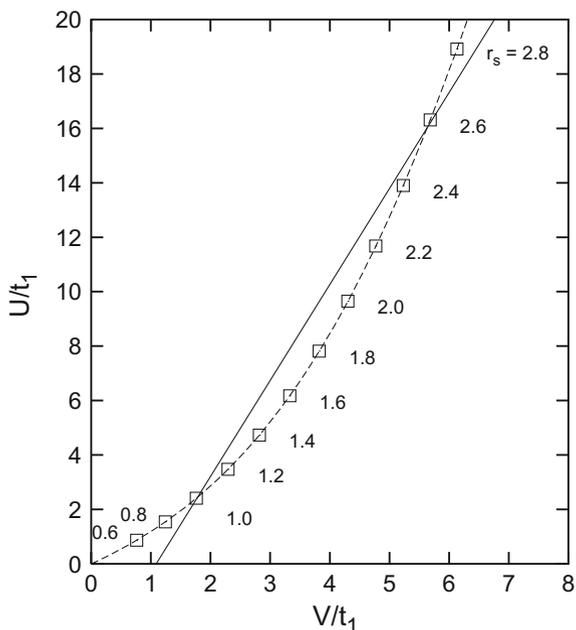
In the insulating phase, the Coulomb interaction is not screened by the conduction electrons, and long range contributions cannot be neglected. We should evaluate the parameters U, $V^{(m)}$ as diagonal matrix elements of the bare Coulomb interaction in the Wannier representation. If we assume $V_{ij} \sim 1/|\mathbf{R}_i - \mathbf{R}_j|$, the renormalized parameter W follows from its definition, Eq. (18.5), as $W \approx \frac{1}{8}\alpha_M V$, being α_M the Cs–Cl Madelung constant, $\alpha_M = 1.763$ (see, e.g., [14]), and $V \equiv V^{(1)}$ the first term in the expansion, Eq. (18.5), i.e., the nearest neighbour repulsive interaction. The problem is thus mapped back onto the standard extended Hubbard model with an effective number of nearest neighbour sites $z^* = \alpha_M$, to be compared with the bcc value $z_1 = 8$. For the reasonable choice $t_2/t_1 = 0.8$,[1] the $U-V$ phase diagram for the CDW instability is shown in Fig. 18.1.

The boundary between the metallic and the ionic insulating phases is given by the simple linear relation $U/t_1 = -\gamma + 2\alpha_M V/t_1$, being γ the minimum value of the ratio $(16W - U)/t_1$, as emerging from Eq. (18.10) for $\Gamma = \Gamma_c = 2t_2$ ($\gamma = 3.84$ for $t_2/t_1 = 0.8$). In principle, an ionic metallic phase may exist just over the boundary, since the gap closure is due to band overlap, while a finite order parameter Δ always arises from the gap Eq. (18.10). However, in presence of a band overlap, the total energy, Eq. (18.9), is incorrect, since the energy levels must be summed up to the Fermi value inside both the bands $E^{\pm}$. Thus, Eq. (18.10) is not correct in the metallic

---

[1] A slight change of the ratio t_2/t_1 does not affect too much the phase diagram. In general, a smaller value than 0.8 would be expected, thus giving rise to an enlargement of the ionic region in the phase diagram.

phase and the existence of a stable broken-symmetry ground state is questionable in the metallic regime. Moreover, all the Coulomb interaction terms would be strongly screened by the conduction electrons, so that the symmetric $\Delta = 0$ ground state is expected to be more favoured for the metallic phase. The phase diagram is incomplete, since we have not taken in consideration the possible occurrence of SDW instabilities, which are likely to be present for $U \gg V$, though irrelevant for the following considerations.

Let us first discuss the U–V phase diagram in relation with the behaviour of solid hydrogen under pressure. The possible bcc phase of dense hydrogen would be a simple 'free-electron' metal, with an almost spherical Fermi surface and a Fermi energy comparable with the free electron value $E_F = 1.84/r_s^2$ a.u., which also reproduces the observed Fermi energies for the alkali metals. In the metallic phase, neglecting any interaction term, the model Fermi energy arises from the unperturbed spectrum given by Eqs. (18.4a) and (18.4b) in terms of the parameters t_1, t_2 as $E_F = \varepsilon(\mathbf{k}_F) - \varepsilon(0)$, where $\mathbf{k}_F$ is the Fermi vector, $ak_F \approx (6\pi^2)^{1/3}$. Comparing the latter with the definition for E_F in the free electron case yields an estimate of t_1 for any chosen ratio t_2/t_1, and for any fixed density. If we fix $t_2/t_1 = 0.8$, then $t_1 = 0.259/r_s^2$ a.u. Approximating the Wannier states by $1s$ hydrogen wave functions, both U and V follow, in a.u. [15], as $U = \frac{5}{8}$, $V = \frac{1}{R} - e^{-2R}\left(\frac{1}{R} + \frac{11}{8} + \frac{3}{4}R + \frac{1}{6}R^2\right)$, where the nearest neighbour distance $R = (\sqrt{3}\pi)^{1/3}r_s$ has been used. For large r_s values, $U/t_1 \sim r_s^2$, while $V/t_1 \sim r_s$, so that the equation of state for hydrogen in the U/t_1–V/t_1 phase diagram of Fig. 18.1 is just a parabola. Some possible states of dense hydrogen are reported on Fig. 18.1 At very high density, the equation of state deviates from the parabolic behaviour, since V saturates for $r_s \to 0$. However, the very high density limit is questionable, and must be considered as an extrapolation out of the range where the adopted approximations are reasonable.

The phase diagram is not significatively altered by a change of the ratio t_2/t_1, since both the boundary line and the equation of state are shifted in the same direction and their relative changes compensate.

If we rely on the emerging scenario, even in the very high density limit, then we would predict that hydrogen metallization requires $r_s < 1$, since the bcc phase would be unstable towards an insulating ionic phase for $1.0 < r_s < 2.6$. The occurrence of such an ionic phase could explain the failure of all the attempts to reach the monatomic state at the currently achievable densities. In fact, as discussed by Chen et al. [7] and Ashcroft [16, 17], the occurrence of an IR active vibron mode could be justified by the presence of permanent dipole moments. Besides, the occurrence of any other molecular ionic phase cannot be ruled out by our approach, which only prevents the stability of the monatomic bcc structure for a given high density range. However, the high density limit $r_s \simeq 1$ is only qualitatively correct, as previously discussed, being the approximations out of control in this very high density regime. Therefore, we do not find any contrast with the Monte Carlo prediction [18] of a transition towards the monatomic phase for $r_s \approx 1.3$. Moreover, that numerical calculation evidenced the equivalence of the monatomic ground state energies in both the fcc and in the

bcc phases. Thus the ionic instability, lowering the ground state energy, should be relevant even if an fcc structure were the most stable monatomic phase.

While the lower bound of the ionic phase is only qualitatively determined by the present approach, the upper limit $r_s < 2.6$ is much more reliable. Any hydrogenoid bcc solid should undergo a MI transition around that critical density value,[2] thus realizing an unexpected broken-symmetry ionic phase. At room pressure, $r_s = 3.94$ for sodium and $r_s = 3.25$ for lithium, so that according to Fig. 18.1 both the elements are correctly predicted to be stable in the monatomic bcc phase. In order to reach the critical density $r_s \approx 2.6$, a very high pressure is required [19–21]. Such a pressure could be really prohibitive for K, Rb and Cs. Besides, these heavier alkali elements undergo several pressure induced structural transitions which are believed to be driven by the electronic transfer to upper empty bands [19]. Therefore, our attention should be focused on the lightest alkali, since such transfer mechanism is negligible and the required pressure could be reached by modern diamond anvil cell apparatus. Lithium has been compressed up to $r_s \approx 2.8$ [20], and a phase transition from bcc to fcc has been observed for $P = 6.9\,\text{GPa}$ ($r_s \approx 2.9$). The occurrence of the fcc phase could in principle invalidate our discussion, even if the ground state energies of such cubic monatomic structures are so close that the ionic instability cannot be ruled out at higher densities. Mostly, the best alkali with which we may compare our prediction is sodium: (i) it has an almost spherical Fermi surface; (ii) its $3s$ orbital is expected to be comparable for extension with the $1s$ orbital of hydrogen[3]; (iii) no structural phase transition has been observed up to $r_s \approx 2.9$ [21]; (iv) the first empty d band is far from the Fermi energy. A structural phase transition has been predicted [23] from bcc to hcp at the very high density $r_s \lesssim 2.5$, which should be reachable in the $100\,\text{GPa}$ region [19]. According to Fig. 18.1, a MI transition should occur first, around $r_s \approx 2.6$, then the ionic phase could push to higher densities the structural transition.

At this stage, we should reverse our starting question, and we should ask instead: Why is sodium a metal? In a broken-symmetry ionic phase, the inter-ion Coulomb interactions add a considerable contribution to the total ground state energy. This very same term amounts to the almost entire cohesive energy of any ionic crystal. We cannot neglect such interactions, even at high densities, in comparison with the Fermi energy. On the other hand, the rôle played by the on-site Hubbard U and by the nearest neighbour effective interaction W is competitive, as it is evident from Eqs. (18.9) and (18.10). The former favours a SDW instability, whereas the latter is responsible for the onset of a CDW. In sodium, under normal thermodynamic conditions, such interaction terms compensate each other, precluding any instability. The ground state is a metal, the interactions are strongly screened by the Fermi liquid, and give only rise to a small renormalization of the band parameters. In other words, the symmetric

---

[2]Notice that for $r_s \sim 2.6$, V is almost independent of the atomic size, while U is slightly smaller for larger atomic orbitals, so that the critical r_s value could be even larger for the other alkali metals.

[3]From the minimal basis set of Clementi and Raimondi [22], the effective $3s$ atomic radius is $a_0 = 3/(z - \sigma)$ a.u., where $z = 11$ and the shield constant $\sigma = 8.4927$ for sodium, thus yielding $a_0 = 1.196$, to be compared with $a_0 \equiv 1$ for hydrogen.

metallic phase is based on the equilibrium between competing interactions. Whenever we alter such an equilibrium (e.g., increasing W by compressing the distances) the system falls into a broken-symmetry insulating phase, where the interactions are no longer screened and play an essential rôle. Such a phase is expected for sodium in the 100 GPa region, for $r_s \approx 2.6$. The ionic phase could explain the failure to reach hydrogen metallization, and would give a possible interpretation of the anomalous IR active vibron mode observed in the high pressure A phase. A very high density is required in order to reduce the rôle of the interactions in comparison with the kinetic electronic energy, and to restore symmetry. For a perfectly nested Fermi surface ($t_2 = 0$), such a density is infinite, since the integral in the gap Eq. (18.10) diverges for $\Gamma = 0$. Out of nesting ($t_2 > 0$), a metallic phase arises around the $U \approx 2\alpha_M V$ region (Fig. 18.1), which also explains the behaviour of all the alkali metals.

Note added in proof

After this work was completed (and released as preprint `arXiv:cond-mat/-9512011`), we received the information that the hypothesis of a CDW instability in hydrogen under pressure was first advanced by Ashcroft [24]. Furthermore, some months later we got news that Edwards and Ashcroft [25] performed a density functional calculation where they show that a CDW is present in hydrogen under high pressure, due to the exchange contribution to the total energy.

Quite recently, Hemley et al. [26], from an analysis of new measurements of the high infrared vibron mode of solid hydrogen above 150 GPa, have argued that the effective charge of hydrogen reaches $0.04e$ at 216 GPa (85 K). This rules out (up to this pressure) full ionization of the hydrogen molecules, but it is not in contradiction with a CDW instability.

References

1. E. Wigner, H.B. Huntington, J. Chem. Phys. **3**(12), 764 (1935). DOI 10.1063/1.1749590
2. J.H. Eggert, K.A. Goettel, I.F. Silvera, Europhys. Lett. **11**(8), 775 (1990). DOI 10.1209/0295-5075/11/8/014
3. J.H. Eggert, K.A. Goettel, I.F. Silvera, Europhys. Lett. **12**(4), 381 (1990). DOI 10.1209/0295-5075/12/4/017
4. H.K. Mao, R.J. Hemley, Science **244**(4911), 1462 (1989). DOI 10.1126/science.244.4911.1462
5. L. Cui, N.H. Chen, I.F. Silvera, Phys. Rev. Lett. **74**, 4011 (1995). DOI 10.1103/PhysRevLett.74.4011
6. M. Hanfland, R.J. Hemley, H. Mao, Phys. Rev. Lett. **70**, 3760 (1993). DOI 10.1103/PhysRevLett.70.3760
7. N.H. Chen, E. Sterer, I.F. Silvera, Phys. Rev. Lett. **76**, 1663 (1996). DOI 10.1103/PhysRevLett.76.1663
8. R. Pucci, F. Siringo, in *Molecular systems under high pressure: Proceedings of the II Archimedes Workshop on Molecular Solids Under Pressure, Catania, Italy, 28–31 May 1990*, ed. by R. Pucci, G. Piccitto (North Holland, Amsterdam, 1991), pp. 97–115

9. H. Fujihisa, Y. Fujii, K. Hase, Y. Ohishi, N. Hamaya, K. Tsuji, K. Takemura, O. Shimomura, H. Takahashi, T. Nakajima, High Press. Res. **4**(1-6), 330 (1990). DOI 10.1080/08957959008246112

10. B. Baranowski, Polish J. Chem. **66**, 1737 (1992)

11. T.M. Giebultowicz, A.W. Overhauser, S.A. Werner, Phys. Rev. Lett. **56**, 1485 (1986). DOI 10.1103/PhysRevLett.56.1485. Erratum Phys. Rev. Lett. **56**, 2228 (1986)

12. P.G. Coulter, W.R. Datars, Phys. Rev. B **34**, 2963 (1986). DOI 10.1103/PhysRevB.34.2963

13. A.W. Overhauser, in *Highlights of Condensed Matter Theory*, ed. by F. Bassani, F. Fumi, M.P. Tosi (North-Holland, Amsterdam, 1985), p. 194

14. N.W. Ashcroft, N.D. Mermin, *Solid State Physics* (Saunders College Publ., Fort Worth, 1976)

15. J.C. Slater, *Quantum theory of molecules and solids*, vol. 1 (McGraw-Hill, New York, 1963)

16. N.W. Ashcroft, Physics World **8**(7), 43 (1995). DOI 10.1088/2058-7058/8/7/30

17. N.W. Ashcroft, Phys. Rev. B **41**, 10963 (1990). DOI 10.1103/PhysRevB.41.10963

18. D. Ceperley, in *Simple Molecular Systems at Very High Density: Proceedings of a NATO Advanced Research and EPS Workshop, Les Houches, 1988*, *NATO ASI Series*, vol. B186, ed. by A. Polian, P. Loubeyre, N. Boccara (Plenum Press, New York, 1989)

19. H. Olijnyk, W. Holzapfel, Phys. Lett. A **99**(8), 381 (1983). DOI 10.1016/0375-9601(83)90298-0

20. B. Olinger, J.W. Shaner, Science **219**(4588), 1071 (1983). DOI 10.1126/science.219.4588.1071

21. I.V. Alexandrov, S.M. Stishov, V.N. Kachinsky, I. Makarenko, in *Proceedings of the XX EHPRG Meeting* (Stuttgart, 1982)

22. E. Clementi, D.L. Raimondi, J. Chem. Phys. **38**(11), 2686 (1963). DOI 10.1063/1.1733573

23. J.A. Moriarty, A.K. McMahan, Phys. Rev. Lett. **48**, 809 (1982). DOI 10.1103/PhysRevLett.48.809

24. N.W. Ashcroft, in *Elementary processes in dense plasmas*, ed. by S. Ichimaru, S. Ogata (Addison-Wesley, Reading, MA, 1995)

25. B. Edwards, N.W. Ashcroft (1995). Private communication

26. R.J. Hemley, I.I. Mazin, A.F. Goncharov, H.K. Mao, M. Hanfland, M. Li, in *Annual March Meeting of the American Physical Society* (APS, 1996). Abstract #N27.04

Part II
Molecular Chemistry

Chapter 19
Negative Condensed-to-Atom Fukui Functions: A Signature of Oxidation-Induced Reduction of Functional Groups

E. Echegaray, A. Toro-Labbe, K. Dikmenli, F. Heidar-Zadeh, N. Rabi, S. Rabi, P.W. Ayers, C. Cardenas, Robert G. Parr and J.S.M. Anderson

Abstract We show that the orbital relaxation contribution to the Fukui function, first derived by Yang, Parr, and Pucci in 1984, is decisive for predicting redox-induced electron transfer. Specifically, we study a dinuclear cobalt complex which has been shown to exhibit redox-induced electron transfer, and show that this effect is associated, computationally, with negative values of the condensed Fukui functions. This establishes that not only can the Fukui function be negative, but negative values of the Fukui function are chemically significant.

E. Echegaray · A. Toro-Labbe
Facultad de Química, Nucleus Millennium Chemical Processes and Catalysis (CPC),
Pontificia Universidad Católica de Chile, Casilla 306, Santiago, Chile
e-mail: atola@uc.cl

K. Dikmenli · F. Heidar-Zadeh · N. Rabi · S. Rabi · P.W. Ayers
Department of Chemistry and Chemical Biology, McMaster University,
Hamilton, Ontario L8S 4M1, Canada

P.W. Ayers (✉)
e-mail: ayers@mcmaster.ca

C. Cardenas
Departamento de Física, Facultad de Ciencias, Universidad de Chile,
Casilla 653, 7800024 Santiago, Chile
e-mail: cardena@macul.ciencias.uchile.cl

C. Cardenas
Centro para el Desarrollo de la Nanociencia y la Nanotecnología (CEDENNA),
Avda. Ecuador, Casilla 653, Santiago, Chile

R.G. Parr
Department of Chemistry University of North Carolina, Chapel Hill, NC 27759, USA
e-mail: rgparr@email.unc.edu

J.S.M. Anderson
Computational Materials Science Research Team, Minatojima-minami-machi,
Chuo-ku, Kobe, Hyogo 650-0047, Japan
e-mail: james.anderson@riken.jp

269

Background

The Fukui Function

In 1984, Yang, Parr, and Pucci defined the Fukui function as the change in electron density due to a change in the number of electrons, at fixed external potential [1, 2],

$$f(\mathbf{r}) = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_{v(\mathbf{r})}. \tag{19.1}$$

Owing to the derivative discontinuity of the energy, electron density, and other molecular properties [3, 4], this derivative does not exist for isolated molecules (which always have an integer number of electrons), and should be replaced by the derivative from above (often called the electrophilic Fukui function, or simply the Fukui function from above)

$$f^+(\mathbf{r}) = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_{v(\mathbf{r})}^+ = \lim_{h \to 0^+} \frac{\rho(\mathbf{r}; N+h) - \rho(\mathbf{r}; N)}{h} = \rho(\mathbf{r}; N+1) - \rho(\mathbf{r}; N) \tag{19.2}$$

and the derivative from below (often called the nucleophilic Fukui function, or simply the Fukui function from below)

$$f^-(\mathbf{r}) = \left(\frac{\partial \rho(\mathbf{r})}{\partial N} \right)_{v(\mathbf{r})}^- = \lim_{h \to 0^+} \frac{\rho(\mathbf{r}; N) - \rho(\mathbf{r}; N-h)}{h} = \rho(\mathbf{r}; N) - \rho(\mathbf{r}; N-1). \tag{19.3}$$

The final equality in Eqs. (19.2) and (19.3) is a consequence of the piecewise linear dependence of the energy and other properties on the number of electrons [5, 6]. The electrophilic Fukui function is large where it is most favorable for a molecule to accept electrons, and therefore serves as the fundamental density-based reactivity indicator for the location where the molecule is most susceptible to nucleophilic attack. The nucleophilic Fukui function is large where it is most favorable for a molecule to donate electrons, and therefore indicates where the molecule is most susceptible to attack by an electrophile. The Fukui function is one of the most important reactivity indicators in conceptual density functional theory [7–11]. It is the key indicator for regioselectivity in density-based approaches to chemical reactivity [6, 12].

The Fukui function was named because it is the density functional theoretic analog of the frontier molecular orbital reactivity indicators developed by Kenichi Fukui [13–17]. This link is made more clear when the Fukui function is rewritten in terms of the Kohn–Sham orbitals [2, 18],

$$f^-(\mathbf{r}) = |\phi_{\text{HOMO}}(\mathbf{r})|^2 + \sum_{i \leq \text{HOMO}} \left(\frac{\partial |\phi_i(\mathbf{r})|^2}{\partial N} \right)_{v(\mathbf{r})}, \qquad (19.4a)$$

$$f^+(\mathbf{r}) = |\phi_{\text{LUMO}}(\mathbf{r})|^2 + \sum_{i \geq \text{LUMO}} \left(\frac{\partial |\phi_i(\mathbf{r})|^2}{\partial N} \right)_{v(\mathbf{r})}, \qquad (19.4b)$$

where HOMO and LUMO denote the highest occupied molecular orbital and the lowest unoccupied molecular orbital, respectively. The first term in these expressions is the frontier molecular orbital contribution, while the second term includes the effects of orbital relaxation. While usually the frontier-orbital term is dominant (in which case, the Fukui function mirrors the frontier molecular orbital theory of Fukui), there are important cases where quantitative, and even qualitative, effects of orbital relaxation are important [19–29]. When orbital relaxation is important, the simple frontier molecular orbital approach fails, but the Fukui function is still a reliable descriptor of molecular electronic structure and chemical reactivity. Equation (19.4) are popularly known as the Yang–Parr–Pucci formulae [2]. They are the key working equations for most practical applications of the Fukui function.

Redox-Induced Electron Transfer (RIET)

One of the greatest successes of the Fukui function was its use in predicting the phenomenon of redox-induced electron transfer (RIET) [30] prior to the recent surge of experimental interest in this phenomenon [31–35]. In RIET, the oxidation (or reduction) of a molecule leads to the reduction (or oxidation) of one of the moieties within it. Such a phenomenon can happen only for compounds with at least three redox-active components like metal centers and noninnocent ligands [28]. Since adding (or subtracting) electrons from the molecule as a whole is associated with a decrease (or increase) of the electron population on one of the redox centers, RIET is associated with negative values of the Fukui function $f^+(\mathbf{r})$ (or $f^-(\mathbf{r})$). To quantify the change of the Fukui function on the molecular region associated with an atom in a molecule, one frequently uses the condensed Fukui function, which can be defined as the change in atomic population associated with the addition (or subtraction) of an electron [36]. Alternatively, in terms of the charges on the atoms in a molecule, q_α, one has

$$f_\alpha^+ = q_\alpha(N) - q_\alpha(N+1), \qquad (19.5a)$$
$$f_\alpha^- = q_\alpha(N-1) - q_\alpha(N). \qquad (19.5b)$$

RIET is associated with molecules where the condensed Fukui function is significantly less than zero on one or more atoms or functional groups within the molecule [30]. In this chapter, we present the first computational study of the condensed Fukui function for a molecule that is known to undergo RIET, [TPyA-Co$^{\text{II}}$-CA^{2-}-Co$^{\text{II}}$-

TPyA]$^{2+}$ [31]. The definition of the condensed Fukui function is somewhat ambiguous [37], and the entire concept of the Fukui function is problematic for molecules with degenerate or quasidegenerate ground states [38, 39]. In this paper, however, we will ignore these nuances and use Eq. (19.5) as working formulas for the Fukui function of a molecular site.

System and Method

Redox-Induced Electron Transfer in the TPyA-Co$^{\mathrm{II}}$-CA^{2-}-Co$^{\mathrm{II}}$-TPyA Dication

Min et al. synthesized and characterized the [TPyA-Co$^{\mathrm{II}}$-CA^{2-}-Co$^{\mathrm{II}}$-TPyA]$^{2+}$ dinuclear transition metal complex, where TPyA denotes the tetradentate Tris(2-pyridylmethyl)amine ligand and CA is the chloranilate ligand [31]. The chloranilate ligand is a noninnocent (i.e., redox-active) aromatic ligand that links the two antiferromagnetically coupled (d^7, high spin) Co$^{\mathrm{II}}$ metal centers. When the system is oxidized, one would expect an electron to be removed from the HOMO (which, based on our calculations is located on the Co$^{\mathrm{II}}$ centers), thereby obtaining a mixed-oxidation-state compound with one Co$^{\mathrm{II}}$ and one Co$^{\mathrm{III}}$ metal center. This is not what is observed experimentally: instead both metal centers become oxidized to (d^6, low-spin) Co$^{\mathrm{III}}$, while the chloranilate linker ligand is reduced to the CA^{3-} radical [31]. This complex therefore undergoes RIET, whereby oxidation of [TPyA-Co$^{\mathrm{II}}$-CA^{2-}-Co$^{\mathrm{II}}$-TPyA]$^{2+}$ gives [TPyA-Co$^{\mathrm{III}}$-CA$^{.3-}$-Co$^{\mathrm{III}}$-TPyA]$^{3+}$, a metal complex in which the chloranilate ligand is reduced. This presumably occurs because as the cobalt ions become more positive, they stabilize the orbitals on the chloranilate ligand between them, so that it is more favorable to have the "extra" electron on the chloranilate ligand than on one of the metal centers. That is, we can envision this as a two-step process. In the first step, removing an electron from the HOMO of [TPyA-Co$^{\mathrm{II}}$-CA^{2-}-Co$^{\mathrm{II}}$-TPyA]$^{2+}$ results in the [TPyA-Co$^{\mathrm{III}}$-CA^{2-}-Co$^{\mathrm{II}}$-TPyA]$^{3+}$ complex. However, the loss of antiferromagnetic coupling destabilizes the remaining Co$^{\mathrm{II}}$ center, and the loss of an electron from the coordinating cobalt center stabilizes the lowest unoccupied orbital on the chloranilate ligand. This makes it favorable for an electron to transfer from the remaining Co$^{\mathrm{II}}$ center to the chloranilate, giving the [TPyA-Co$^{\mathrm{III}}$-CA$^{.3-}$-Co$^{\mathrm{III}}$-TPyA]$^{3+}$ complex.

Because of the RIET effect, if one computes $f^-(\mathbf{r})$ for [TPyA-Co$^{\mathrm{II}}$-CA^{2-}-Co$^{\mathrm{II}}$-TPyA]$^{2+}$, one expects the Fukui function to have substantial positive values on the metal centers and significant negative values on the chloranilate ligand. The goal of this chapter is to test this hypothesis computationally.

Computational Approach

We performed calculations on the metal complex $[\text{TPyA-Co}^{\text{II}}\text{-CA}^{2-}\text{-Co}^{\text{II}}\text{-TPyA}]^{2+}$ in both its oxidized and reduced states using the BLYP density functional [40, 41]. The initial guesses for these calculations were constructed by combining fragments (ligands and metal ions) that had been precomputed with the charge and spin multiplicity suggested by the experiment [31]. The initial geometry was taken from the crystal structure; geometry optimization with the LANL2DZ basis set (for Cobalt) and the 6-31G(d) basis set (all other atoms) led to only very small changes in molecular geometry. The relaxed molecular geometry is shown in Fig. 19.1. Atomic charges and spin densities were computed at the optimized geometry using the larger 6-311+G(d) basis set for all atoms. The spin densities agree with what is expected from the experiment: $[\text{TPyA-Co}^{\text{II}}\text{-CA}^{2-}\text{-Co}^{\text{II}}\text{-TPyA}]^{2+}$ has high-spin antiferromagnetically coupled cobalt atoms (Fig. 19.2a), and the chloranilate ligand is essentially closed shell, with negligible spin density. After oxidation to $[\text{TPyA-Co}^{\text{III}}\text{-CA}^{.3-}\text{-Co}^{\text{III}}\text{-TPyA}]^{3+}$, an unpaired electron is localized on the chloranilate ligand and the cobalt ions become low spin (Fig. 19.2b). All calculations were performed using the *Gaussian09* program [42, 43].
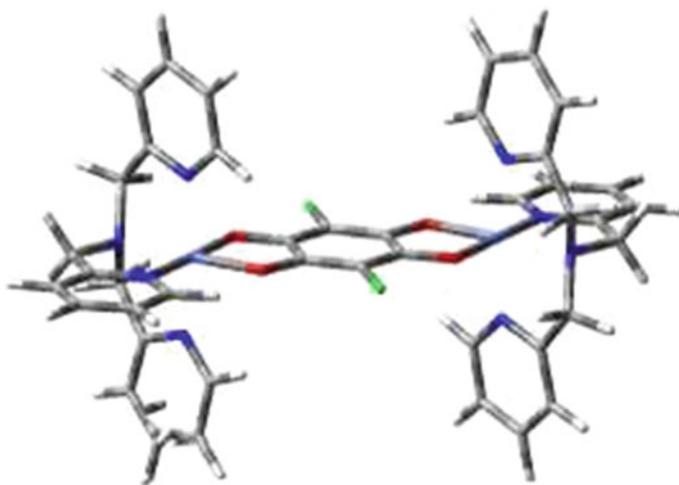
Fig. 19.1 The optimized structure of $[\text{TPyA-Co}^{\text{II}}\text{-CA}^{2-}\text{-Co}^{\text{II}}\text{-TPyA}]^{2+}$. The essentially planar chloranilate anion is in the *center* of the picture, linking the Cobalt ions. The TPyA ligands bind the *top*, *bottom*, and *back side* of the Cobalt ions
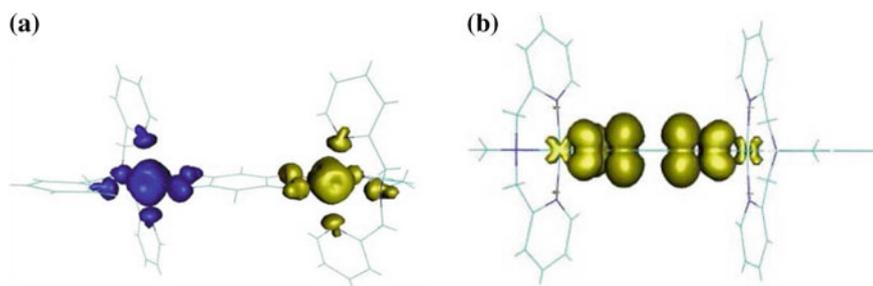
Fig. 19.2 The spin density on the complex α-spin (in *yellow*) and β-spin electrons (in *blue*) density of studied complex. **a** $[\text{TPyA-Co}^{\text{II}}\text{-CA}^{2-}\text{-Co}^{\text{II}}\text{-TPyA}]^{2+}$ has high-spin antiferromagnetically coupled Co^{III} centers. **b** $[\text{TPyA-Co}^{\text{III}}\text{-CA}^{3-}\text{-Co}^{\text{III}}\text{-TPyA}]^{3+}$ has low-spin Co^{III} centers and an unpaired electron on the chloranilate trianion radical (color figure online)

Results

We computed the atom-condensed Fukui function using the differences of atomic charges adhering to the fragment of molecular response approach, i.e.,

$$f_\alpha^- = q_\alpha \left(\left[\text{Co}^{\text{III}} - \text{CA}^{.3-} - \text{Co}^{\text{III}} \right]^{3+} \right) - q_\alpha \left(\left[\text{Co}^{\text{II}} - \text{CA}^{2-} - \text{Co}^{\text{II}} \right]^{2+} \right). \quad (19.6)$$

Because atomic charges are inherently ambiguous [44, 45] we selected two different approaches. The iterative Hirshfeld [46] approach of Bultinck et al. [47] is based on finding the atoms in a molecule whose densities maximally resemble those of the isolated atom/ion in an information-theoretic sense [44, 48–52], but with an additional step that ensures that the reference ion and the atom in a molecule have the same charge [47]. The electrostatic potential fitting method of Breneman and Wiberg, CHELPG [53], chooses the atomic charges so that the electrostatic potential on the exterior of the molecule is reproduced as accurately as possible.

As shown in Table 19.1, both methods of computing atomic charges give similar results, with the chloranilate ligand having an overall negative condensed Fukui function, with decisively negative values on the carbon atoms that are closest to the cobalt centers. The fact that the condensed Fukui function is not minus one on the ligand should not be surprising: upon oxidation of the metal centers and reduction of the chloranilate ligand, the cobalt ions are better electron acceptors and the chloranilate ligand is a better electron donor. Therefore, if one imagines the redox process as a two-step procedure, where at the end of the first step the chloranilate ligand has an additional electron, and then in the next step inductively donates electrons to the oxidized metal atoms, one expects that the total value of the condensed Fukui function on the chloranilate ligand should be negative, but far greater than minus one. This is consistent with Fig. 19.2, where the unpaired electrons from the metal are slightly donated to the chloranilate ligand in the reduced complex, but the unpaired electron from the chloranilate ligand is slightly donated to the cobalt

Table 19.1 The condensed Fukui functions on the chloranilate ligand, computed using iterative Hirshfeld (iH) and CHELPG charges. The numbering scheme of the atoms is shown in Fig. 19.3. The total condensed Fukui function (including the oxygens) of the chloranilate ligand is also included

Atom labels	Atomic charge method	
	iH	CHELPG
C 1	0.074	0.016
C 2	−0.096	−0.142
C 3	−0.077	−0.024
C 4	−0.077	−0.018
C 5	−0.096	−0.146
C 6	0.074	0.019
O 8	0.024	0.038
O 9	−0.020	0.095
O 11	0.024	0.037
O 12	−0.020	0.096
Cl 7	0.003	0.025
Cl 10	0.003	0.025
total	−0.192	−0.245

Fig. 19.3 The numbering scheme for the atoms in the chloranilate ion that is used in Table 19.1
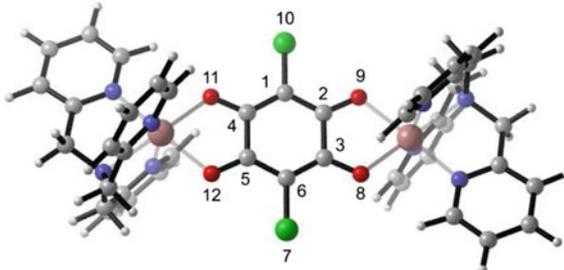
ions in the oxidized complex. It is also consistent with the computed value of about −0.2 for the condensed Fukui function of the chloranilate ring system.

Discussion

The Fukui function of Parr, Yang, and Pucci extends the frontier molecular orbital theory of Fukui by including the effects of electron correlation and orbital relaxation. This is especially important for chemical processes where orbital relaxation is crucial. Among these processes, perhaps the most interesting is the effect of redox-induced electron transfer (RIET). RIET, and its less extreme version called redox-

induced electron rearrangement, depends critically on orbital relaxation, and cannot be described using simple frontier molecular orbital theory.

To test the ability of the Fukui function to model RIET, we considered the [TPyA-Co^{II}-CA^{2-}-Co^{II}-TPyA]$^{2+}$ complex that was synthesized and studied by Min et al. [31]. When this compound is oxidized, one would predict that the electron should be removed from the highest occupied molecular orbital. The electron density would then decrease everywhere in space

$$\rho(\mathbf{r}) \rightarrow \rho(\mathbf{r}) - |\phi_{\mathrm{HOMO}}(\mathbf{r})|^2. \tag{19.7}$$

This is not what is observed either experimentally or theoretically. Experimentally, it is observed that oxidization of the metal complex is associated with a reduction of the noninnocent chloranilate ligand, which is confirmed by the negative values of the condensed Fukui function on this ligand, indicating that the chloranilate ligand gains electrons even though the overall complex loses an electron. This means that for RIET in this complex, and presumably also in other molecules, the orbital relaxation term in the Yang–Parr–Pucci equations (Eq. (19.4)) is decisive. Negative values of the condensed Fukui function are thereby established as the hallmark of RIET. More generally, this study resolves the question of whether the Fukui function can ever be negative: negative values of the Fukui function not only exist, but signify chemically important effects.

References

1. R.G. Parr, W. Yang, J. Am. Chem. Soc. **106**(14), 4049 (1984). DOI 10.1021/ja00326a036
2. W. Yang, R.G. Parr, R. Pucci, J. Chem. Phys. **81**(6), 2862 (1984). DOI 10.1063/1.447964. Reprinted in this volume, p. 307
3. J.P. Perdew, R.G. Parr, M. Levy, J.L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982). DOI 10.1103/PhysRevLett.49.1691
4. W. Yang, Y. Zhang, P.W. Ayers, Phys. Rev. Lett. **84**, 5172 (2000). DOI 10.1103/PhysRevLett.84.5172
5. W.P. Ayers, M. Levy, Theor. Chem. Acc. **103**(3), 353 (2000). DOI 10.1007/s002149900093
6. P.W. Ayers, J. Math. Chem. **43**(1), 285 (2008). DOI 10.1007/s10910-006-9195-5
7. P.A. Johnson, L.J. Bartolotti, P.W. Ayers, T. Fievez, P. Geerlings, in *Modern charge density analysis*, ed. by C. Gatti, P. Macchi (Springer, New York, 2012), pp. 715–764
8. P.W. Ayers, J.S.M. Anderson, L.J. Bartolotti, Int. J. Quantum Chem. **101**(5), 520 (2005). DOI 10.1002/qua.20307
9. P. Geerlings, F.D. Proft, W. Langenaeker, Chem. Rev. **103**(5), 1793 (2003). DOI 10.1021/cr990029p
10. J.L. Gázquez, J. Mex. Chem. Soc. **52**(1), 3 (2008)
11. S.B. Liu, Acta Physico-Chimico Sinica **52**(3), 590 (2009)

12. P. Ayers, W. Yang, L.J. Bartolotti, in *Chemical Reactivity Theory: A Density Functional View*, ed. by P.K. Chattaraj (CRC Press, Boca Raton, 2009)
13. K. Fukui, T. Yonezawa, H. Shingu, J. Chem. Phys. **20**(4), 722 (1952). DOI 10.1063/1.1700523
14. K. Fukui, T. Yonezawa, C. Nagata, J. Chem. Phys. **21**(1), 174 (1953). DOI 10.1063/1.1698592
15. K. Fukui, T. Yonezawa, C. Nagata, Bull. Chem. Soc. Jpn. **27**(7), 423 (1954). DOI 10.1246/bcsj.27.423
16. H. Fujimoto, K. Fukui, G. Klopman, in *Chemical reactivity and reaction paths*, ed. by G. Klopman (J. Wiley & Sons, New York, 1974), pp. 23–54
17. K. Fukui, *Theory of orientation and stereoselection* (Springer, Berlin, 1975)
18. M.H. Cohen, M.V. Ganduglia-Pirovano, J. Kudrnovský, J. Chem. Phys. **101**(10), 8988 (1994). DOI 10.1063/1.468026
19. W. Langenaeker, K. Demel, P. Geerlings, J. Mol. Struct.: Theochem **234**, 329 (1991). DOI 10.1016/0166-1280(91)89021-R
20. K. Flurchick, L. Bartolotti, J. Mol. Graph. **13**(1), 10 (1995). DOI 10.1016/0263-7855(94)00010-P
21. L.J. Bartolotti, P.W. Ayers, J. Phys. Chem. A **109**(6), 1146 (2005). DOI 10.1021/jp0462207
22. P.W. Ayers, R.C. Morrison, R.K. Roy, J. Chem. Phys. **116**(20), 8731 (2002). DOI 10.1063/1.1467338
23. J. Melin, P.W. Ayers, J.V. Ortiz, J. Phys. Chem. A **111**(40), 10017 (2007). DOI 10.1021/jp075573d
24. E. Echegaray, C. Cárdenas, S. Rabi, N. Rabi, S. Lee, F.H. Zadeh, A. Toro-Labbe, J.S.M. Anderson, P.W. Ayers, J. Mol. Model. **19**(7), 2779 (2013). DOI 10.1007/s00894-012-1637-3
25. E. Echegaray, S. Rabi, C. Cárdenas, F.H. Zadeh, N. Rabi, S. Lee, J.S.M. Anderson, A. Toro-Labbe, P.W. Ayers, J. Mol. Model. **20**(3), 2162 (2014). DOI 10.1007/s00894-014-2162-3
26. P. Bultinck, R. Carbó-Dorca, J. Math. Chem. **34**(1), 67 (2003). DOI 10.1023/A:1025136721324
27. P. Bultinck, R. Carbó-Dorca, W. Langenaeker, J. Chem. Phys. **118**(10), 4349 (2003). DOI 10.1063/1.1542875
28. P. Senet, M. Yang, J. Chem. Sci. **117**(5), 411 (2005). DOI 10.1007/BF02708344
29. D. Petrini, K. Larsson, J. Phys. Chem. C **112**(37), 14367 (2008). DOI 10.1021/jp711190r
30. P.W. Ayers, Phys. Chem. Chem. Phys. **8**, 3387 (2006). DOI 10.1039/B606167B
31. K.S. Min, A.G. DiPasquale, A.L. Rheingold, H.S. White, J.S. Miller, J. Am. Chem. Soc. **131**(17), 6229 (2009). DOI 10.1021/ja900909u
32. J.S. Miller, K.S. Min, Angewandte Chemie International Edition **48**(2), 262 (2009). DOI 10.1002/anie.200705138
33. N. Paul, S. Samanta, S. Goswami, Inorg. Chem. **49**(6), 2649 (2010). DOI 10.1021/ic9016195
34. W. Kaim, Inorg. Chem. **50**(20), 9752 (2011). DOI 10.1021/ic2003832
35. S. Samanta, P. Ghosh, S. Goswami, Dalton Trans. **41**, 2213 (2012). DOI 10.1039/C2DT10986G
36. W. Yang, W.J. Mortier, J. Am. Chem. Soc. **108**(19), 5708 (1986). DOI 10.1021/ja00279a008
37. P. Bultinck, S. Fias, C. Van Alsenoy, P.W. Ayers, R. Carbó-Dorca, J. Chem. Phys. **127**(3), 034102 (2007). DOI 10.1063/1.2749518
38. C. Cárdenas, P.W. Ayers, A. Cedillo, J. Chem. Phys. **134**(17), 174103 (2011). DOI 10.1063/1.3585610
39. P. Bultinck, C. Cardenas, P. Fuentealba, P.A. Johnson, P.W. Ayers, J. Chem. Theory Comput. **10**(1), 202 (2014). DOI 10.1021/ct400874d
40. A.D. Becke, Phys. Rev. A **38**, 3098 (1988). DOI 10.1103/PhysRevA.38.3098
41. C. Lee, W. Yang, R.G. Parr, Phys. Rev. B **37**, 785 (1988). DOI 10.1103/PhysRevB.37.785
42. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R.

Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Ö. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, *Gaussian 09, Revision A.1*. Gaussian, Inc., Wallingford, CT (2009)

43. A.D. Becke, J. Chem. Phys. **98**(7), 5648 (1993). DOI 10.1063/1.464913
44. R.G. Parr, P.W. Ayers, R.F. Nalewajski, J. Phys. Chem. A **109**(17), 3957 (2005). DOI 10.1021/jp0404596
45. P.L. Ayers, R.J. Boyd, P. Bultinck, M. Caffarel, R. Carbó-Dorca, M. Causá, J. Cioslowski, J. Contreras-Garcia, D.L. Cooper, P. Coppens, C. Gatti, S. Grabowsky, P. Lazzeretti, P. Macchi, Á.M. Pendás, P.L.A. Popelier, K. Ruedenberg, H. Rzepa, A. Savin, A. Sax, W.H.E. Schwarz, S. Shahbazian, B. Silvi, M. Solà, V. Tsirelson, Comput. Theor. Chem. **1053**, 2 (2015). DOI 10.1016/j.comptc.2014.09.028. Special Issue: Understanding structure and reactivity from topology and beyond
46. F.L. Hirshfeld, Theor. Chim. Acta **44**(2), 129 (1977). DOI 10.1007/BF00549096
47. P. Bultinck, C. Van Alsenoy, P.W. Ayers, R. Carbó-Dorca, J. Chem. Phys. **126**(14), 144111 (2007). DOI 10.1063/1.2715563
48. R.F. Nalewajski, R.G. Parr, Proc. Natl. Acad. Sci. USA **97**(16), 8879 (2000). DOI 10.1073/pnas.97.16.8879
49. R.F. Nalewajski, R.G. Parr, J. Phys. Chem. A **105**(31), 7391 (2001). DOI 10.1021/jp004414q
50. P.W. Ayers, J. Chem. Phys. **113**(24), 10886 (2000). DOI 10.1063/1.1327268
51. P.W. Ayers, Theor. Chem. Acc. **115**(5), 370 (2006). DOI 10.1007/s00214-006-0121-5
52. R.F. Nalewajski, Chem. Phys. Lett. **372**(1–2), 28 (2003). DOI 10.1016/S0009-2614(03)00335-X
53. C.M. Breneman, K.B. Wiberg, J. Comput. Chem. **11**(3), 361 (1990). DOI 10.1002/jcc.540110311

Chapter 20
Simple Approaches to Calculate Correlation Energy in Polyatomic Molecular Systems

A. Grassi, G.M. Lombardo and G. Forte

Abstract Ab initio calculation including electron correlation are still extremely costly, except for the smallest atoms and molecules. In this paper we present some simple semi-empirical methods to obtain correlation energy. These methods are based on the relation between energy and the off-diagonal density matrix elements, which represent the bonding between atoms in the molecule. The results of our previous studies are reported here and compared with the results obtained by using more accurate techniques.

Introduction

The terms 'electron correlation' and correlation energy was first used by Wigner [1] indicating an energy difference between the "exact" nonrelativistic molecular energy and energy obtained from a Hartree–Fock calculation, i.e.: $E_c = E_{\text{exact, nonrel}} - E_{\text{HF}}$. Independent particle models, such as Hartree and Hartree–Fock methods, neglect this correlation at various level of approximation. In the Hartree model, there is an unphysical finite probability to find two electron in the same place, whereas the Hartree–Fock model electrons with parallel spin are prevented from overlapping but not those with antiparallel spin. Indeed, in the previous definition E_c includes also the "basis set error" which is present in a simple HF calculation. For this reason some authors define the correlation energy E_c as $E_c = E_{\text{exact, nonrel}} - E_{\text{HF, limit}}$, i.e., the energy difference with respect to a HF calculation with an "infinite" basis set. The value of $E_{\text{HF, limit}}$ can be obtained as limit value of various HF calculations on increasing the basis set.

A. Grassi (✉) · G.M. Lombardo · G. Forte
Dipartimento di Scienze del Farmaco, Università di Catania,
Viale A. Doria 6, I-95125 Catania, Italy
e-mail: agrassi@unict.it

G.M. Lombardo
e-mail: glombardo@dipchi.unict.it

G. Forte
e-mail: gforte@unict.it

Now, to account for electron correlation, several post-HF analytical techniques have been proposed, such as the Møller–Plesset perturbation theory (MP) [2], Configuration Interaction (CI) [3] or the Multi-configurational self-consistent field (MCSCF) [4] techniques, where the correlation is included at various levels within the HF calculation. Unfortunately, these methods are computationally very expensive even for small molecular systems.

Therefore, in the last 20 years alternative approaches to calculate the correlation energy have been developed. In these methods the electron correlation energy E_c is related to some molecular properties, such as single- or two-particle electron density, bond order, bond distances, overlap populations, and so on. Some of these techniques are 'semi-analytical' methods, because they need an *a priori* empirical estimation of some parameters [5]. In general, the parameters are related to some atomic and atom–atom pair properties. It is worth to note how these methods are quite accurate, giving comparable results with respect to the analytical ones, and furthermore, they require lightweight computational efforts. The detailed exposition of these analytical or semi-analytical methods are beyond the scope of this paper. Therefore, in the following, we will expose in detail the methods that we have developed to calculate the correlation energy for simple or polyatomic molecules.

Methods

Correlation Energy versus Electron Number

Starting from a study of March and Wind [6], where the functional density theory (DFT) was used to explain the approximately linear variation in neutral atom correlation energy with atomic number Z, we have extended the model to treat some neutral diatomic molecules. As in Ref. [7], the starting point was to rewrite the correlation energy as:

$$E_c = \int \varepsilon_c(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}, \qquad (20.1)$$

where $\varepsilon_c(\mathbf{r})$ is the correlation energy per electron (CEPE) at $\mathbf{r}$ and $\rho(\mathbf{r})$ is the electron density. Following the classical LCAO (Linear Combination of Atomic Orbitals) procedure, we expand the molecular wavefunction in an atomic basis set $\phi_\mu(\mathbf{r})$. So that we rewrite Eq. (20.1) as:

$$E_c = \sum_\mu^A P_{\mu\mu} \int \varepsilon_c(\mathbf{r})\phi_\mu^2(\mathbf{r})d\mathbf{r} + \sum_\mu^B P_{\mu\mu} \int \varepsilon_c(\mathbf{r})\phi_\mu^2(\mathbf{r})d\mathbf{r}$$

$$+ \sum_{\mu \neq \nu} P_{\mu\nu} \int \varepsilon_c(\mathbf{r})\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})d\mathbf{r},$$

$$(20.2)$$

where $P_{\mu\nu}$ are density matrix elements partitioned into atomic (A e B) and overlap terms. Assuming that the function $\varepsilon_c(\mathbf{r})$ varies in $\mathbf{r}$-space slowly with respect to $\phi_\mu^2(\mathbf{r})$, we can rewrite Eq. (20.2) as a sum of atomic and overlap contribution, i.e.:

$$E_c = E_c^{\text{atomic}} + E_c^{\text{overlap}} = \varepsilon_A \sum_\mu^A P_{\mu\mu} + \varepsilon_B \sum_\mu^B P_{\mu\mu} + \varepsilon_0 \sum_{\mu\neq\nu} P_{\mu\nu} S_{\mu\nu}. \quad (20.3)$$

The first two terms refer to the atomic contribution, whereas the last refers to the interatomic one. In particular, ε_A, ε_B and ε_0 are the CEPE for atom A and B (taken from experimental tabulated values [8]) and for the overlap terms represented by $S_{\mu\nu}$, the overlap matrix obtained from ab initio calculations. The assumption of Eq. (20.3) is to treat the overlap term at same level as for atoms, i.e., that correlation energy per electron for the overlap part is constant. This assumption is partially supported by the work of Gombas [9], who shows that ε_c is not to strongly varying function of $\rho(\mathbf{r})$ which, in the overlap region, is a smoothly varying function of $\mathbf{r}$. Recalling the orthonormality of the atomic basis set, the term $\sum_{\mu\neq\nu} P_{\mu\nu} S_{\mu\nu}$ represents the total overlap population ($n^0(AB)$), and the term $\sum_\mu^A P_{\mu\mu}$ effective electron charge on the atom ($n^{\text{eff}}(A)$). Then, we can rewrite Eq. (20.3) as:

$$E_c = \varepsilon_A n^{\text{eff}}(A) + \varepsilon_B n^{\text{eff}}(B) + \varepsilon_0 n^0(AB). \quad (20.4)$$

Now, three different approaches have been taken in order to estimate the value of ε_0:

(i) $\varepsilon_0 \simeq k_{(i)} N$, where N is the total number of electrons in the molecules. This assumption is based on the work of March and Wind [6], which shows that for atoms with small atomic number Z, the CEPE is nearly proportional to the total number of electrons; $k_{(i)}$ is a constant, calculated by the least square minimization procedure.

(ii) $\varepsilon_0 \simeq (\varepsilon_A Z_A + \varepsilon_B Z_B)/2$. In this approach the overlap term of the CEPE is obtained as a mean value between total atomic correlation energy E_A and E_B.

(iii) $\varepsilon_0 \simeq k_{(iii)} n^0(AB)$. At variance of (i) the overlap CEPE is proportional to the total overlap population, i.e., to the bond electron density. A detailed test of his assumption is showed in [10]. The constant $k_{(iii)}$ was also calculated by the least square minimization procedure.

In Table 20.1 we have reported the experimental and calculated ones of the correlation energies for some diatomic molecules. Note that for method (i) the value of the constant $k_{(i)}$ was found to be equal to 12.5 mhartree/electron^2, whereas for method (iii) the value of the constant $k_{(iii)}$ is equal to 136.8 mhartree/electron^2.

In Fig. 20.1 we report the difference (in mhartree) between the experimental E_c values with respect to the various theoretical methods.

In particular, the top figure shows that for the hydride molecular systems better results are obtained with the more sophisticated methods as HS [5] or GCP. Whereas in homonuclear systems X_2 our methods appear to be more accurate.

Table 20.1 Experimental and calculated correlation energies for some diatomic molecules. In columns the experimental values, taken from [8], the calculated values with our methods (columns (i), (ii) and (iii)), Hollister and Sinanoglu (HS) [5], Savin et al. (SPP) [8], Langreth et al. (GCL) [8] and Perdew et al. (GCP) [8]

System	Exp	(i)	(ii)	(iii)	HS	SPP	GCL	GCP
LiH	83	72	52	136	79	113	98	93
BH	153	155	148	187	154	181	165	163
NH	243	254	251	278	239	249	251	255
FH	387	356	366	362	380	364	381	380
Li_2	122	140	115	191	112	171	139	134
C_2	514	471	479	456	422	398	405	399
N_2	546	566	584	518	522	489	511	506
F_2	746	691	714	670	713	674	684	697

Fig. 20.1 Correlation energy difference between the experimental E_c and the calculated values for the various theoretical methods. Top figure the hydride systems (XH). *Bottom* figure the homonuclear systems (X_2). All values are in mhartree
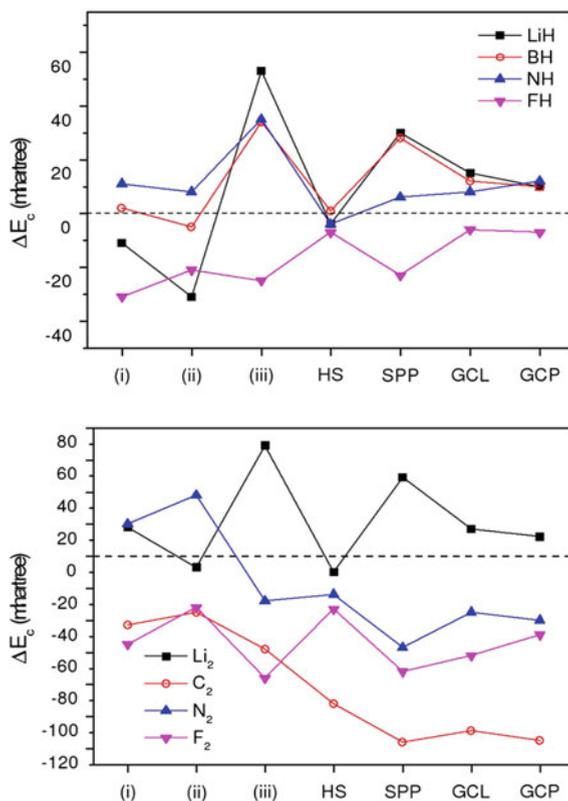
It is noteworthy that in our method there is a direct relation between bond correlation energy and electron number of the molecular system. On the basis of this hypothesis and according to the work of Kais et al. [11], which expands the atomic E_c in term of Z, we have shown [12] that this relation holds also for homonuclear diatomic systems.

Finally, we can infer that, despite the simplicity of the methods, our results reflect qualitatively the correct general trends of the correlation energy in diatomic molecules, and, for some systems, the difference found with respect to the experimental E_c is smaller than those obtained by using more sophisticated methods reported in Table 20.1.

Correlation Energy and Bond Order

In order to treat molecular systems containing more than two atoms, following Cremer's work [13], we define the correlation energy E_c as:

$$E_c = E_S - E_{HF}, \tag{20.5}$$

where E_{HF} is the molecular Hartree–Fock energy and E_S is the so called Schrödinger energy, which is obtained by the exact solution of the Schrödinger equation when vibrational, rotational and relativistic effects are excluded. This definition, as formerly stated, includes also than the "basis set error" which is present in a simple HF calculation.

Now, taking into account the formation reaction of a generic molecule,

$$A + B = C + \text{binding energy}, \tag{20.6}$$
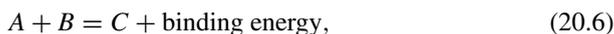
the molecular energy can be partitioned as follows:

$$E(C) = E(A) + E(B) - \text{binding energy} = E(A) + E(B) + E(AB). \tag{20.7}$$

The first step is to consider this Schrödinger molecular energy E_S partitioned as follows:

$$E_S = \sum_{A=1}^{N} E_S(A) + \sum_{\text{all } AB} E_S(AB), \tag{20.8}$$

where N is the number of atoms in the molecule, $E_S(A)$ is the Schrödinger energy of the atom A and $E_S(AB)$ is the Schrödinger binding energy of the two bonded atoms A and B. Likewise, the total HF molecular energy can be written as:

$$E_{HF} = \sum_{A=1}^{N} E_{HF}(A) + \sum_{\text{all } AB} E_{HF}(AB). \tag{20.9}$$

Molecular correlation energy is obtained subtracting Eq. (20.9) from Eq. (20.8), i.e.:

$$E_c = E_S - E_{HF} = \sum_{A=1}^{N} E_c(A) + \sum_{\text{all } AB} E_c(AB). \tag{20.10}$$

The Bond Order

In the framework of LCAO-MO theory (linear combination of atomic orbitals–molecular orbital), the definition of bond order in a polyatomic molecule was given first by Coulson [14] in the context of the Hückel MO approximation [15]. These bond orders are the off-diagonal elements $T_{\mu\nu}$ of the first-order density matrix T. More explicitly:

$$T_{\mu\nu} = \sum_{i=1}^{\text{occ}} n_i c_{i\mu} c_{i\nu}^*, \tag{20.11}$$

where n_i is the occupation number of the ith MO, and $c_{i\mu}$ is the coefficient of the ith atomic orbital in the μth MO. This definition of bond order is applicable when the AOs (atomic orbitals) are mutually orthogonal and only one AO is considered for each atom. Various definitions for bond order have been proposed, e.g. Löwdin definition [16], which gives an orthogonal first-order density matrix Q by using the transformation:

$$Q = S^{1/2} T S^{1/2}, \tag{20.12}$$

where S is the overlap matrix. In this framework the bond order P_{AB} between the atoms A and B in a closed shell molecule is defined as:

$$P_{AB} = \sum_{\mu}^{A} \sum_{\nu}^{B} Q_{\mu\nu} Q_{\nu\mu}. \tag{20.13}$$

More detailed treatments of the bond order for both closed shell and open shell systems is reviewed by Sannigrahi [17].

Bond Order Correlation Energy (BOCE)

The starting point is to relate the bond order population to the correlation energy, assuming that $E_c(AB)$ can be developed as series function of the bond order P_{AB}, we have:

$$E_c(AB) = \sum_{m} a_{m,AB} (P_{AB})^m. \tag{20.14}$$

Table 20.2 $a_{1,AB}$ (mhartree/electron) as defined by Eq. (20.14)

A—B	$a_{1,AB}$	A—B	$a_{1,AB}$
C—H	35.12	F—F	167.50
C—C	31.70	Cl—Cl	131.44
O—H	60.54	F—H	113.77
C—O	49.11	Cl—H	77.00
O—O	72.04	F—Si	127.33
H—H	39.55	Cl—Si	102.46
Si—H	62.38	Cl—F	158.62

The coefficient $a_{m,AB}$ are determined for a bond between two equal or different atom types (C−C, C−O, O−O, etc.), selecting a set of molecules containing the bond types reported in Table 20.2. Without loss of generality, we have retained only the first term of the expansion. Detailed procedure followed to obtain the bond parameters $a_{m,AB}$, and the selected molecular systems used in the best fit are reported in Refs. [18–20]. In Table 20.2 are reported the parameters $a_{1,AB}$ derived for the 6-31G** basis set.

In the first paper, Ref. [18], the BOCE methods were applied to calculate the correlation energies of 20 molecules containing C, O, and H atoms. In the next work, Ref. [19], the method was applied to calculate the molecular dissociation energies (D_0) and heats of formation (ΔH). Finally in Ref. [20], the procedure was extended to calculate correlation energies in polyatomic molecules containing Si, F, and Cl atoms.

In Table 20.3 we have reported the Schrödinger and the BOCE molecular energies, $E_{BOCE} = E_{HF} + E_c^{BOCE}$, for some molecules containing the atoms of Table 20.2.

Table 20.3 Schrödinger and BOCE molecular energies (hartree) for some polyatomic molecules

	E_S	E_{BOCE}		E_S	E_{BOCE}
C_6H_6	−232.24219	−232.24129	SiH_3F	−391.14909	−391.13615
C_2H_4	−78.58567	−78.58084	SiH_2F_2	−490.47468	−490.46072
C_2H_2	−77.33430	−77.31958	$SiHF_3$	−589.79953	−589.79277
H_2CO	−114.50344	−114.50278	SiH_3Cl	−751.37713	−751.37455
CH_3OH	−115.72137	−115.72340	SiH_2Cl_2	−1210.93407	−1210.92994
$H_2C=CO$	−152.60000	−152.59566	$SiHCl_3$	−1670.49132	−1670.48676
CO_2	−188.58504	−188.60229	SiH_2FCl	−850.697705	−850.69489
$(CH_3)CHO$	−153.82924	−153.82682	$SiHF_2Cl$	−950.024261	−950.02430
$(CH_2)CHOH$	−153.81314	−153.80838	$SiHFCl_2$	−1310.25583	−1310.25538
HCOOH	−189.76370	−189.77838	SiF_2Cl_2	−1409.57885	−1409.58379
CO	−113.31751	−113.32008	SiF_3Cl	−1049.35506	−1049.35384
Cyclopropane	−117.89065	−117.88280	$SiFCl_3$	−1769.81401	−1769.81308
Furane	−230.01980	−230.00899	SiH_2	−290.56195	−290.53928
t-Butene	−157.21942	−157.21007	SiF_2	−489.21821	−489.20621
c-Butene	−157.21848	−157.20760	$SiCl_2$	−1209.69956	−1209.67880
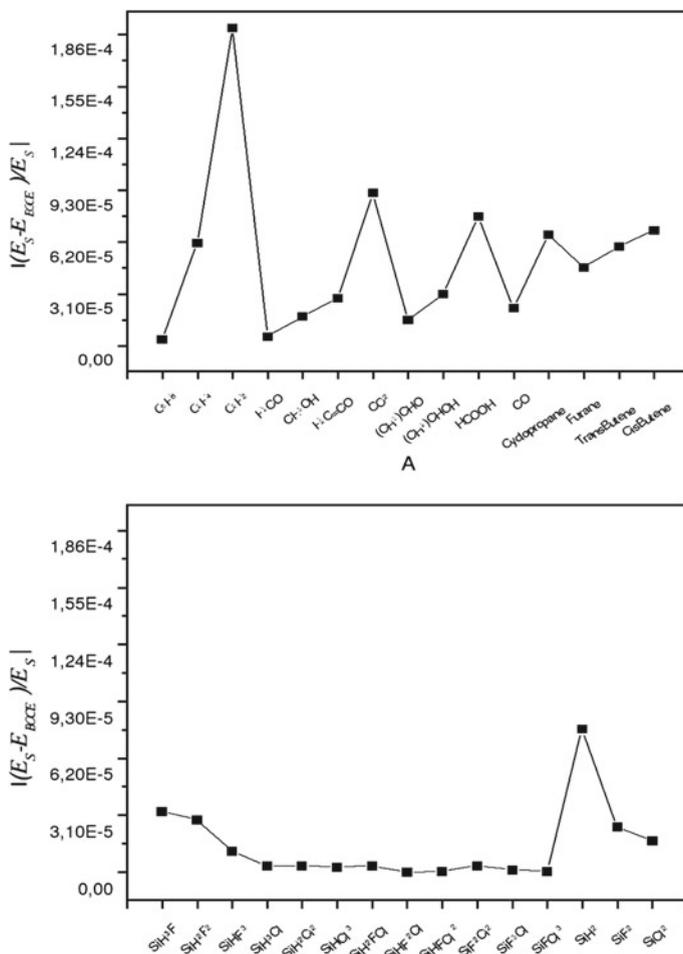
Fig. 20.2 Absolute relative energy difference between the Schrödinger and the calculated (BOCE) energy values for various molecular systems. In the *top figure*, molecular systems containing C, O and H atoms. In the *bottom figure*, molecular systems containing Si, F, Cl and H atoms

The very good agreement of the calculated energies using the BOCE techniques with respect to the Schrödinger ones, appears clearly form the values in Table 20.3. The highly accurate values of the correlation energy obtained using BOCE approach, for all molecules of the series, are confirmed from the calculated relative errors with respect to the Schrödinger energies, reported in Fig. 20.2.

However, it is important to add a comment: some calculated BOCE energies are lower than experimental (Schrödinger) energies. This is explained by taking into account that these values are obtained from various experimental data and any experimental measurement is subject to an error that, in general, is about $\pm 5\%$.

The BOCE results, for the molecular systems in the series, obtain values which are in the range of $\pm 0.007\%$ of the Schrödinger energy, which are much lower than the experimental error.

Finally, the good results obtained from BOCE technique in the molecular energy calculation is confirmed in the calculation of the molecular dissociation energy (D_0) as well as in the estimation of the molecular heat formation. Both molecular quantities were compared with experimental values and with the values calculated by using a very accurate, but computationally expensive, method (G2) [21], as shown in Ref. [20].

In conclusion, despite the simplicity of the BOCE approach, the results obtained by using this approach in the molecular energy calculation and other related molecular quantities are well comparable with the results obtained by using more accurate and very computationally expensive techniques.

References

1. E. Wigner, Phys. Rev. **46**, 1002 (1934). DOI 10.1103/PhysRev.46.1002
2. C. Møller, M.S. Plesset, Phys. Rev. **46**, 618 (1934). DOI 10.1103/PhysRev.46.618
3. J.A. Pople, R. Seeger, R. Krishnan, Int. J. Quantum Chem. **12**(S11), 149 (1977). DOI 10.1002/qua.560120820
4. D. Hegarty, M.A. Robb, Mol. Phys. **38**(6), 1795 (1979). DOI 10.1080/00268977900102871
5. C. Hollister, O. Sinanoglu, J. Am. Chem. Soc. **88**(1), 13 (1966). DOI 10.1021/ja00953a003
6. N.H. March, P. Wind, Mol. Phys. **77**(4), 791 (1992). DOI 10.1080/00268979200102771
7. A. Grassi, G.M. Lombardo, N.H. March, R. Pucci, Mol. Phys. **81**(5), 1265 (1994). DOI 10.1080/00268979400100851
8. A. Savin, H. Stoll, H. Preuss, Theor. Chim. Acta **70**(6), 407 (1986). DOI 10.1007/BF00531922
9. P. Gombás, *Pseudopotentials* (Springer, New York, 1967)
10. N. Spartà, A. Grassi, G.M. Lombardo, G. Piccitto, R. Pucci, N.H. March, Mol. Phys. **83**(6), 1047 (1994). DOI 10.1080/00268979400101771
11. S. Kais, S.M. Sung, D.R. Herschbach, Int. J. Quantum Chem. **49**(5), 657 (1994). DOI 10.1002/qua.560490511
12. A. Grassi, G.M. Lombardo, N.H. March, R. Pucci, Mol. Phys. **86**(5), 1229 (1995). DOI 10.1080/00268979500102691
13. D. Cremer, Journal of Computational Chemistry **3**(2), 154 (1982). DOI 10.1002/jcc.540030206
14. C.A. Coulson, Proc. R. Soc. Lond. A **169**(938), 413 (1939). DOI 10.1098/rspa.1939.0006
15. E. Hückel, Z. Physik **70**, 204 (1942)
16. P.O. Löwdin, J. Chem. Phys. **18**(3), 365 (1950). DOI 10.1063/1.1747632
17. A.B. Sannigrahi, in *Advances in Quantum Chemistry*, vol. 23, ed. by P.O. Löwdin, J.R. Sabin, M.C. Zerner (Academic Press, 1992), pp. 301–351. DOI 10.1016/S0065-3276(08)60032-5
18. A. Grassi, G.M. Lombardo, N.H. March, R. Pucci, Mol. Phys. **87**(3), 553 (1996). DOI 10.1080/00268979600100381
19. A. Grassi, G.M. Lombardo, G. Forte, G.G.N. Angilella, R. Pucci, N.H. March, Mol. Phys. **104**, 453 (2006). DOI 10.1080/00268970500404273
20. A. Grassi, G.M. Lombardo, G. Forte, G.G.N. Angilella, R. Pucci, N.H. March, Mol. Phys. **104**, 1447 (2006). DOI 10.1080/00268970500509899
21. L.A. Curtiss, K. Raghavachari, G.W. Trucks, J.A. Pople, J. Chem. Phys. **94**(11), 7221 (1991). DOI 10.1063/1.460205

Chapter 21
Novel Common Methodologies Between Physics and Theoretical Chemistry: Density Functional Theory

R. Pucci

Introduction

Over the last five decades, theoretical chemistry has often been considered a highly systematic discipline, which attempts at solving Schrödinger equation for increasingly complex systems, with greater accuracy. Since then, this research programme has been believed as largely successful. In this context, the following quote by Clementi [1] is clarifying: «We can calculate everything».

But 'to calculate' a molecule does not mean 'to understand' it [2], as the calculations do not unveil by themselves the fundamental laws of physics and chemistry. Nor even accurate calculations on the single systems can evidence regularities and universalities in atomic and molecular systems.

An effort toward the quest of new methodologies has been done by physicists,[1] who, despising brute-force ab initio calculations, devised simple models [3], which

---

[1]The distinction between molecular physicists and theoretical chemists is purely academic, and is here employed for the sake of mere comodity, essentially to indicate different methodologies, rather than different research fields. In this context, I would like to report a quote by C.A. Coulson: «There are no physicists and chemists, but problems and people who are able to solve them, and people who are not» (N.H. March, private communication).

R. Pucci (✉)
Dipartimento di Fisica e Astronomia, Università di Catania, Via S. Sofia, 64, 95123 Catania, Italy
e-mail: renato.pucci@ct.infn.it

R. Pucci
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci
IMM-CNR, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

introduce drastic approximations [4, 5]. However, these approximations are not always able to yield results with the required accuracy, from the 'chemical' point of view. Therefore, theoretical chemists have always considered the methods of molecular physics as interesting and 'extravagant,' but scarcely useful for practical purposes.

On the other hand, physicists too need ever more realistic models and accurate calculations.

During the Archimedes Workshop on Molecular Theory, held at the Department of Physics, University of Catania, Italy, in June 1985,[2] it was evidenced that new methodologies were emerging, which combine the techniques of both physics and chemistry. During that workshop, the right balance of physicists and chemists among the conveners enabled a lively debate and a proficuous exchange of experiences.

It would be beyond the scopes of the present article to enumerate all the contributions on the various topics covered by the workshop.[3] Therefore, I will only propose, as an example of what has been said previously, density functional theory. This theory allows to analyze the concept of electronegativity and several regularities in atoms and molecules.

Another important subject, which was covered by the workshop and where the collaboration among chemists and physicists appears essential, is gas chemisorption on surfaces. I will come back on this topic more in detail in the conclusions.

Density Functional Theory

The theory of the inhomogeneous electron gas, or density functional theory (DFT), was originated by the seminal works of Thomas [8] and Fermi [9]. The so-called Thomas–Fermi (TF) approximation already provides good results for the binding energy of atomic ions [10], but it is not able to reproduce the richness of the periodic table of the elements, which is connected with the orbital structure of atoms. This is why Hartree theory [11–13] enjoyed a wider diffusion.

The introduction within TF theory of the exchange term is due to Dirac [14], and in Hartree's theory to Fock [15]. Later, Slater [16] demonstrated that an approximation of the Hartree–Fock (HF) method was equivalent to the Thomas–Fermi–Dirac (TFD) approximation. One often ignores that the so-called 'X_α method' can be obtained as an approximation of density functional theory [17].

It was thanks to Gombas [18] and others that correlation effects were introduced in a consistent way within TFD theory. For some time several physicists were sceptical about the reliability of a theory based on the description of atomic and molecular properties in terms of the electronic density, but in 1964 Hohenberg and Kohn's

---

[2]A second Archimedes Workshop was held afterwards in 1990. See Ref. [6] [*Translator's Note*].

[3]Suffice it here to mention the contribution on *Electron density and response theory,* by R. McWeeny, that on *Nonadditive dispersion forces,* by E.A. Power, and that on *Cloud of virtual photons around a hydrogen atoms,* by F. Persico. See Ref. [7].

theorem [19] achieved the formal completion of the theory. Hohenberg and Kohn's theorem [19] demonstrates that the total energy of a system of N electrons is a unique functional of the electronic density. Evidently, Hohenberg and Kohn's theorem [19] is a proof of the existence of such a functional, but does not indicate its explicit form. Much progress, however, has been made toward the determination of such a functional, at least in some of its limiting forms. Nowadays, density functional theory is a well-established theory, whereby several 'reasonable' approximations have been derived. These have been largely successful in explaining several problems in molecular as well as solid-state physics, as witnessed by several books and review articles [5, 17, 20–31].

However, density functional theory found it hard to get established among chemists, because of the difficulties inherent to systems with a relatively small number of electrons.

The papers by March [28] and Parr [17] did however demonstrate that this theory can greatly contribute to the development of fundamental concepts in chemistry, and the number of its practical applications are ever increasing [22].

In the following sections, I will review the concept of electronegativity and some regularities in atomic and molecular systems.

Chemical Potential and Electronegativity

Electronegativity is a relevant quantity for chemistry. It is connected with the 'attitude' of an atom or a molecule to attract electrons. Several definitions are available, each yielding a different scale of electronegativity [32, 33], but none of them is derived from first principles. Quite a successful definition of the electronegativity χ is due to Mulliken [34]:

$$\chi_M = \frac{I + A}{2}, \tag{21.1}$$

where I is the potential of first ionization, and A the electronic affinity.

Electronegativity should fulfill two conditions: (i) it should have the same value in any point of the system under consideration, and (ii) if one merges two systems A and B having electronegativities χ_A and χ_B, the resulting system C will have a unique electronegativity χ_C. The fundamental physical quantity appearing in density functional theory which actually fulfills requirements (i) and (ii) is denoted by μ. It is introduced in the theory as a Lagrange multiplier, in order to take into account of the condition

$$\int \rho(\mathbf{r}) \, d\mathbf{r} = N, \tag{21.2}$$

where ρ is the electronic density and N the total number of electrons. The variational principle for the total energy E with respect to density ρ then takes the form

$$\frac{\delta(E - \mu N)}{\delta\rho} = 0. \tag{21.3}$$

Hohenberg and Kohn's theorem [19] then allows one to write the exact Euler equation as

$$\mu = \frac{\delta T}{\delta\rho} + V_{\mathrm{H}} + \frac{\delta\varepsilon_{\mathrm{xc}}}{\delta\rho}, \tag{21.4}$$

where T is the total kinetic energy, V_{H} is Hartree potential energy, and $\varepsilon_{\mathrm{xc}}$ is the exchange-correlation potential energy. Within the statistical theory, where energy is a continuous function of N, μ coincides with the chemical potential,

$$\mu = \frac{dE}{dN}. \tag{21.5}$$

However, for a system characterized by a finite, and often relatively small, value of N, it has not been demonstrated that E is a continuous function of N, and that it is possible to define E for even slightly noninteger values of N. This poses serious difficulties to the fundamentals of quantum mechanics in the study of systems with a finite N.

One way to overcome this difficulty has been proposed by March and the present author [35]. This suggests to adopt μ, as defined by Euler equation, Eq. (21.4), as a definition of electronegativity. Equation (21.4) is a formally exact expression, allowing for several approximations. Particularly interesting are those obtained in the limit $r \to \infty$, where

$$\mu = -I + \lim_{r \to \infty} \frac{\delta\varepsilon_{\mathrm{xc}}}{\delta\rho}. \tag{21.6}$$

We note at this point that Parr et al. [36] strongly favor Eq. (21.5) also for finite systems. Then they demonstrate that

$$\mu = -\chi_{\mathrm{M}}, \tag{21.7}$$

namely that the chemical potential equals the opposite of Mulliken's electronegativity. Equations (21.1), (21.6), and (21.7) provide evidence of the remarkable importance that the potential of first ionization I and electronic affinity A have toward the determination of the equilibrium conditions for an atomic or molecular system.

This suggests the possibility that density functional theory may serve to establish more firmly the theory of frontier orbitals [37, 38]. More precisely, March and Pucci [39] have demonstrated, within the HF approximation and making use of Koopmans theorem, that

$$\frac{\delta T}{\delta\rho} = -\frac{\hbar^2}{2m} \frac{\nabla^2 \psi_{\mathrm{homo}}}{\psi_{\mathrm{homo}}}, \tag{21.8}$$

where the quantity on the left-hand side is the same appearing in Eq. (21.4) (but here evaluated at the HF level), and ψ_{homo} is the eigenfunction of the highest occupied

molecular orbital. A generalization of Eq. (21.8) for Kohn–Sham orbitals [40] has been derived by Yang et al. [41].

The still open debate on the questions mentioned above concerns several fundamental aspects of density functional theory, thus demonstrating the relevance thereof toward the development of concepts and calculation methods in chemistry.

Regularities in Atoms

One of the most interesting aspects of density functional theory, which has been emphasized since its early stages, is that such a theory enables to evidence regularities in some atomic and molecular properties, that would be otherwise difficult to identify on the basis of other calculations, even very accurate, but valid on a 'case by case' basis.

I shall try to illustrate this aspect by considering the moments of the radial density, the atomic scattering factors, and the relation between total energy and diamagnetic susceptibility. In the following section, I will deal with diatomic molecules and quasi-spherically symmetric polyatomic molecules.

Since the energy E of the ground state is a functional of the electronic density ρ, let us consider a simple representation of ρ for the closed-shell atomic ions. It is convenient to employ the Fourier transform of ρ, namely the atomic scattering factor

$$f(\mathbf{K}) = \int_0^\infty 4\pi r^2 \rho(r) \frac{\sin Kr}{Kr} dr. \tag{21.9}$$

The series expansion at small K of this expression above is given by [42]

$$f(K) = N + a_2 K^2 \langle r^2 \rangle + \cdots a_{2n} K^{2n} \langle r^{2n} \rangle, \tag{21.10}$$

where use has been made of the normalization condition

$$\int_0^\infty D(r)dr \equiv \int_0^\infty 4\pi r^2 \rho(r)dr = N, \tag{21.11}$$

and the moments $\langle r^{2n} \rangle$ are defined by

$$\langle r^{2n} \rangle \equiv \int_0^\infty D(r) r^{2n} dr. \tag{21.12}$$

It is known [43] that the term in $\langle r^2 \rangle$ determines the Langevin–Pauli magnetic susceptibility, and possesses therefore a direct physical meaning. Using the exact asymptotic form of the electronic density [44]

$$\rho(r) = A r^\gamma \exp[-2(2I)^{1/2} r] \tag{21.13}$$

and the approximate relation $I \propto Z^{-1/3}$, Pucci and March [45] demonstrated that

$$\frac{\langle r^{2n} \rangle}{\langle r^{2n-2} \rangle} \propto \frac{1}{I}. \tag{21.14}$$

This relation allowed to derive, for neutral atoms with $N = Z$, the identity

$$\frac{f(K) - Z - a_2 K^2 \langle r^2 \rangle}{K^2 \langle r^4 \rangle} = G\left(\frac{K}{\sqrt{I}}\right), \tag{21.15}$$

where G is a universal function of $K/\sqrt{I}$. Although G is not known analytically, it has been reported for the neutral atoms Ar and Kr, using HF calculations [45]. The results obtained are in complete agreement with the predictions of Eq. (21.15).

Fig. 21.1 $\ln I$ versus $\ln\langle r^2 \rangle$ for the isoelectronic sequences with $N = 6, 7, 8, 9,$ and 11. All quantities are in a.u. Values of I are taken from Ref. [46], while $\langle r^2 \rangle$ has been derived from HF calculations by Mann [47, 48]
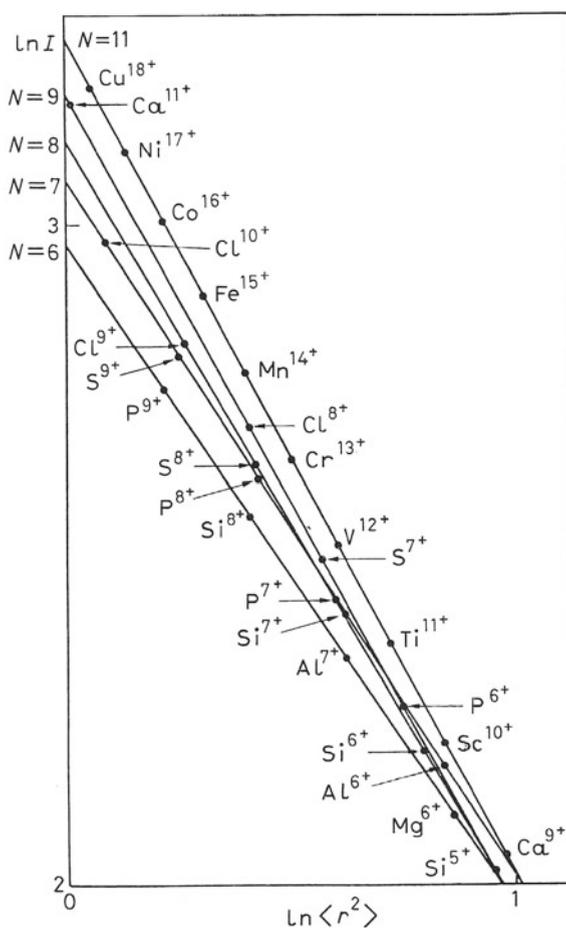
Equation (21.15) indicates that ρ depends considerably on I and $\langle r^2 \rangle$. Therefore, one may expect that $E = E[\rho]$ is also approximately a function of I and $\langle r^2 \rangle$ (Fig. 21.1).

It has been found that for numerous isoelectronic sequences the following relation between I and $\langle r^2 \rangle$ holds [45]

$$\ln I = \alpha(N) \ln \langle r^2 \rangle + \beta(N). \tag{21.16}$$

This is verified very well using experimental values for I and HF results for $\langle r^2 \rangle$, as is illustrated by Fig. 21.1. Equation (21.16) allows to write the total energy in the form

$$E(Z, N) = I + \sum_{n=1}^{N-1} \langle r^2 \rangle_{Z,n} \alpha(n) \exp[\beta(n)]. \tag{21.17}$$

Equations (21.15)–(21.17) unveil unexpected relations among independent physical quantities, and are indicative of some universal behaviors, at least within some classes of systems.

Regularities in Some Classes of Molecules

The combined use of Thomas–Fermi statistical theory and the $1/Z$-expansion of the energy [10] have been very useful to study some regularities in neutral atoms and atomic ions [45, 49]

The generalization of that approach to diatomic molecules allowed Pucci and March [50] to demonstrate that $(R_e^2/Z^{5/3})|E|$ is a universal function of ZR_e^3. Evidently, in this case it is the equilibrium internuclear distance R_e to play a fundamental role in the theory. For B_2, C_2, B_2, and F_2 the relation between $(R_e^2/Z^{5/3})|E|$ and ZR_e^3 is linear, therefore the equilibrium lengths can be evaluated from the relation

$$|E| = aZ^{8/3}R_e + bZ^{5/3}R_e^2, \tag{21.18}$$

with $a = 8.53$ and $b = 60.42$ in a.u. The same regularity is found for heavier homonuclear diatomic molecules.

The theoretical generalization of the arguments above to heteronuclear diatomic molecules is more complicated. For the halides of Li, Na, and K it is however possible to find some empirical regularities [51], similar to the ones above, between $E_e^2 E/N^{5/3}$ and $\bar{Z}R_e^3$, where N is the total number of electrons and $\bar{Z} = (Z_1 Z_2)^{1/2}$ is the geometric average between the nuclear charges Z_1 and Z_2 of the molecule.

An approximate justification of the scaling laws of the Hamiltonian, in the case of heteronuclear diatomic molecules, can be derived when $Z_1 \simeq Z_2$ [51].

In the opposite limit in which $Z_2 \to \infty$ and $Z_1 e$ is kept fixed, a better model is employed, in which the charge $Z_1 e$ is smeared over a spherical surface with radius

equal to the equilibrium distance [52]. Assuming such a model and making use of
TF theory, it is possible to demonstrate [51] that the bond length R_e tends to a finite
value, when $Z_2/Z_1 \to \infty$ (Fig. 21.2).

The above model [21] is particularly useful to treat quasi-spherical polyatomic
molecules, such as tetrahedral or octahedral molecules. In these cases, it is the nuclear
charge $Z_2 e$ of the heaviest central atom which is to infinity, while one uniformly
smears on aspherical surface the charge $N_p e$ of the external nuclei (e.g., the charge
of the 4 protons in the CH_4 molecule). Again making use of this model and of the
TF approximation, it is possible to find an analytical expression for the bond length
R_e [51]

$$R_e = \frac{1}{Z_1^{1/3}} \frac{12}{[3(1-d)]^{1/3}} \left(\frac{3}{32\pi^2}\right)^{2/3} \frac{h^2}{2me^2}, \quad (21.19)$$

where d is determined by the geometry of the molecule. Such a prediction of a finite
bond length as $Z_2 \to \infty$ is completely consistent [54] with experimental data not
only for the series $CH_4, \ldots PbH_4$, but also for the tetrahedral fluorides, chlorides, and
bromides, for the octahedral molecules, and for the diatomic hydrides. An example
of this behavior is depicted in Fig. 21.2.

The dominant role of geometry and dimensionality in determining the electronic
structure is a fundamental property also of the conjugated polymers. It is now agreed

[55] that polyacetylene is a quasi-linear system with a semiconducting band structure characterized by a gap caused by a Peierls distortion, i.e., alternating bond lengths. Actually, in this linear chain with one electron per atom, both the Peierls distortion and an antiferromagnetic staggering are present. Indeed, a nonalternating linear chain possesses an antiferromagnetic ground state [56], due to the Coulombic repulsion between electrons with opposite spins on each atomic site. Such a repulsion increases the dimerization [57]. Moreover, electronic correlations are essential to describe the excited states of finite polyenes [58].

A study of linear polyacenes demonstrated [59] that the corresponding linear polymer should be characterized again by a semiconductor-like ground state, but with a gap due to spin-waves, without any Peierls distortion.

The link between linear polyacenes and graphite (which is known as a bidimensional conductor) is provided by the nonlinear polycyclic hydrocarbons [60]. One has therefore a smooth crossover from one- to two-dimensional systems, which is accompanied by a semiconductor to conductor transition.

Conclusions and Perspective

It seems that a close interchange between the methods of molecular physics and the concepts of theoretical chemistry is not only proficuous, but also necessary.

This has been illustrated by analyzing density functional theory. The topics that have been treated more in detail have been (a) electronegativity and chemical potential, (b) theory of the frontier orbitals, (c) regularities in atoms and in some molecular systems. The concept of electronegativity is not only important for the problems of chemical recativity, but also for the study of polymers [61], as well as of gas chemisorption in metals [62]. In this case, the chemical potential of one of the two systems under consideration, and precisely that of the metal surface, evidently coincides with the Fermi energy E_F.

Both in the case of polymers, and in particular of polyacetylene, and in the case of chemisorption, Seel [63] showed how two complementary descriptions, namely the cluster model (which emphasizes local effects) and the Green function formalism (which includes the effects of the solid), can be combined to maximize the information.

Chemisorption is at the basis of several physico-chemical processes. In particular, chemisorption of atoms on transition metals is the first step toward many catalysis processes. However, even the simplest case of hydrogen chemisorption on transition metals is characterized by remarkable theoretical difficulties.

Recent experimental techniques allowed to measure the chemisorption energies, the corresponding vibrational frequencies, and the bond lengths [64]. In the following, we shall refer in particular to chemisorption of H on Ni(100), Ni(111), and W(110).

Almost every method available in solid state and molecular physics have been applied to the problem of hydrogen chemisorption on transition metals. Among

these, particularly successful have been (1) the Anderson-Newns (AN) model [65, 66], (2) the 'immersion' (ES) method [67, 68], and (3) the cluster calculations (CC) [69].

The AN method emphasizes the role of d electrons in the H–metal bond, and of electronic correlations on the adatom. Recently [70], a generalization of this model has been presented, which allows to treat also the effects of s-p electrons, and to study the various observable quantities, for different surface geometries, and for different positions of the adatom with respect to the surface.

The ES method originates from the works of Lang and Kohn [71] on the jellium model of the surface. It was developed by Inglesfield [67, 72] and several other authors [73, 74]. A common feature of all these works is to emphasize the almost complete screening on the adatom due to the s electrons.

The cluster calculations employ several models, ranging from the extended Hückel method to ab initio methods. These are particularly indicated to clarify the nature of the chemical bond between the H adatom and the metal surface.

All these theories account sufficiently well for the experimental data, but they sometimes provide quite different physical descriptions. In particular, according to some ES schemes, the charge distribution around the adatom is essentially made by an H⁻ ion, surrounded by a localized positive hole, induced by the electron gas on the surface. On the contrary, within the AN model, the adatom remains almost neutral, with a small excess of negative charge. Naturally, it is difficult to resolve experimentally such a discrepancy, since the total charge is similar in the various models. Also in this case, like in numerous other problems in molecular physics, it is necessary to recourse to methods which compensate physical concepts with those of chemistry (Fig. 21.3).
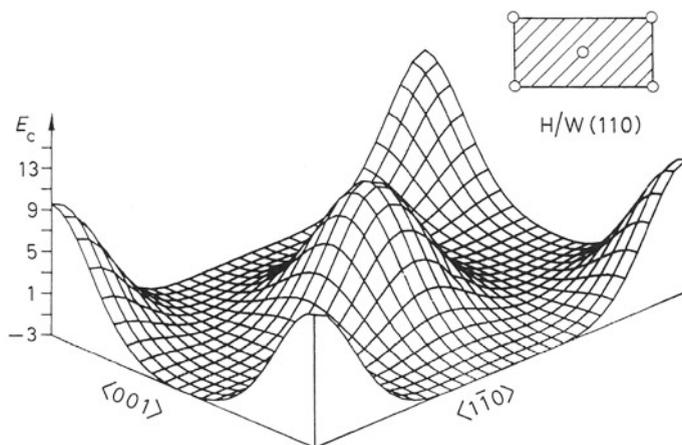
Fig. 21.3 Energy E_c (eV) of hydrogen chemisorption in a plane parallel to the W(110) surface at the equilibrium distance $d_e = 1.7$ a.u. [Figure 21.3, in false colors, was used as the cover image of *Giornale di fisica* **27**(4), and of several of its issues after the appearance of the present paper (*Translator's Note*)]

Some work in this direction has been achieved by our group. We developed a model [66], based on the AN Hamiltonian, which takes into account of the correlations on the adatom, and allows to obtain relatively simple expressions for the quantities of interest for chemisorption. An extension of this model [70] allowed to evaluate entirely the energy surface of an hydrogen atom in the proximity of a metal surface. This has been possible through the extended Hückel approximation for the estimate of the hopping integral, and with the inclusion of core–core repulsion. The model displays the correct asymptotic behavior at large distances, and was applied to H/Ni(100), H/Ni(111), and H/W(110). The results show that it is possible to obtain an accurate description of all the observable physical quantities involved in the process of chemisorption. In the case of H/W(110), our results demonstrate a clear competition between the trigonal and 'long bridge' positions, and therefore a vibration parallel to the surface, as shown in Fig. 21.3. Such a vibration has been recently observed experimentally [75], and the theoretical value of the vibrational frequency is close to the value observed experimentally.

References

1. E. Clementi, in *Proceedings of the Robert A. Welch Foundation Conferences on Chemical Research*, *Theoretical Chemistry*, vol. 16 (Welch Foundation, Houston, 1973), p. 117
2. R.G. Parr, in *Proceedings of the NATO Advanced Study Institute on density functional methods in physics* (Alcabideche, Portugal, 1983)
3. P.W. Anderson, Rev. Mod. Phys. **50**, 191 (1978). DOI 10.1103/RevModPhys.50.191
4. P.W. Anderson, Phys. Rev. **124**, 41 (1961). DOI 10.1103/PhysRev.124.41
5. N.H. March, Adv. Phys. **6**(21), 1 (1957). DOI 10.1080/00018735700101156
6. R. Pucci, G. Piccitto (eds.), *Molecular systems under high pressure: Proceedings of the II Archimedes Workshop on Molecular Solids Under Pressure, Catania, Italy, 28–31 May 1990* (North Holland, Amsterdam, 1991)
7. R. Pucci (ed.), *Abstracts of the Archimedes Workshop on Molecular Theory: Catania, Italy, June 3–6, 1985* (Dipartimento di Fisica dell'Università, Catania, 1985)
8. L.H. Thomas, Math. Proc. Cambridge Phil. Soc. **23**, 542 (1926). DOI 10.1017/S0305004100011683
9. E. Fermi, Rendiconti dell'Accademia Nazionale dei Lincei **6**, 602 (1927)
10. N.H. March, *Self-Consistent Fields in Atoms* (Pergamon Press, Oxford, 1975)
11. D.R. Hartree, Proc. Cambridge Philos. Soc. **24**, 89 (1928). DOI 10.1017/S0305004100011919
12. D.R. Hartree, Proc. Cambridge Philos. Soc. **24**, 111 (1928). DOI 10.1017/S0305004100011920
13. D.R. Hartree, Proc. Cambridge Philos. Soc. **24**, 426 (1928). DOI 10.1017/S0305004100015954
14. P.A.M. Dirac, Proc. Cambridge Phil. Soc. **26**, 376 (1930). DOI 10.1017/S0305004100016108
15. V. Fock, Z. Physik **61**(1), 126 (1930). DOI 10.1007/BF01340294
16. J.C. Slater, Phys. Rev. **81**, 385 (1951). DOI 10.1103/PhysRev.81.385
17. R.G. Parr, Annu. Rev. Phys. Chem. **34**(1), 631 (1983). DOI 10.1146/annurev.pc.34.100183.003215
18. P. Gombás, *Die statistische Theorie des Atoms und ihre Anwendungen* (Springer-Verlag, Vienna, 1949)
19. P.C. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964). DOI 10.1103/PhysRev.136.B864
20. S. Lundqvist, N.H. March (eds.). *Theory of the Inhomogeneous Electron Gas* (Plenum Press, New York, 1983)
21. B.M. Deb (ed.), *The force concept in chemistry* (Van Norstrand Reinhold, New York, 1981)

22. J. Avery, J.P. Dahl (eds.), *Local density approximations in quantum chemistry and solid state physics* (Plenum Press, New York, 1984)
23. R.M. Erdahl, V.H. Smith (eds.), *Density matrices and density functionals: proceedings of the A. John Coleman symposium* (Reidel, Dordrecht, 1987)
24. A.K. Rajagopal, in *Advances in Chemical Physics*, vol. 41, ed. by I. Prigogine, S.A. Rice (J. Wiley & Sons, 1979), pp. 59–193. DOI 10.1002/9780470142608.ch2
25. R.G. Parr, in *Electron distributions and the chemical bond*, ed. by M.B. Hall, P. Coppens (Plenum Press, New York, 1982)
26. R.G. Parr, in *Horizons of quantum chemistry*, ed. by K. Fukui, B. Pullman (Reidel, Dordrecht, 1980)
27. N.H. March, in *Orbital theories of molecules and solids*, ed. by N.H. March (Clarendon Press, Oxford, 1974)
28. N.H. March, Specialist Periodical Reports: Theoretical Chemistry **4**, 92 (1981)
29. E.H. Lieb, Rev. Mod. Phys. **53**, 603 (1981). DOI 10.1103/RevModPhys.53.603
30. S.K. Ghosh, B.M. Deb, Phys. Rep. **92**(1), 1 (1982). DOI 10.1016/0370-1573(82)90134-X
31. N.H. March, Contemp. Phys. **24**, 373 (1983)
32. R.T. Sanderson, *Chemical bonds and bond energy* (Academic Press, New York, 1971)
33. R. Ferreira, in *Advances in Chemical Physics*, vol. 13, ed. by I. Prigogine (J. Wiley & Sons, 1967), pp. 55–84. DOI 10.1002/9780470140154.ch4
34. R.S. Mulliken, J. Chem. Phys. **2**(11), 782 (1934). DOI 10.1063/1.1749394
35. N.H. March, R. Pucci, in *Local density approximations in quantum chemistry and solid state physics*, ed. by J. Avery, J.P. Dahl (Plenum Press, New York, 1984), p. 64
36. R.G. Parr, R.A. Donnelly, M. Levy, W.E. Palke, J. Chem. Phys. **68**(8), 3801 (1978). DOI 10.1063/1.436185
37. K. Fukui, *Theory of orientation and stereoselection* (Springer, Berlin, 1975)
38. R.B. Woodward, R. Hoffmann, *The conservation of orbital symmetry* (Verlag Chemie, Weinheim, 1970)
39. N.H. March, R. Pucci, J. Chem. Phys. **78**(5), 2480 (1983). DOI 10.1063/1.445054
40. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965). DOI 10.1103/PhysRev.140.A1133
41. W. Yang, R.G. Parr, R. Pucci, J. Chem. Phys. **81**(6), 2862 (1984). DOI 10.1063/1.447964. Reprinted in this volume, p. 307
42. A.J. Thakkar, V.H. Smith, J. Phys. B: At. Mol. Phys. **11**(22), 3803 (1978). DOI 10.1088/0022-3700/11/22/009
43. S. Fraga, G. Malli, *Many-electron systems: properties and interactions* (Sanders, Philadelphia, 1968)
44. M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, Phys. Rev. A **16**, 1782 (1977). DOI 10.1103/PhysRevA.16.1782
45. R. Pucci, N.H. March, J. Chem. Phys. **76**(8), 4089 (1982). DOI 10.1063/1.443483
46. C.A. Moore, National Bureau of Standards Circular 647 **1**, 34 (1958)
47. J.B. Mann, Atomic structure calculations. I. Hartree-Fock energy results for the elements hydrogen to lawrencium (1967). Los Alamos Scientific Laboratory Report LA-3690
48. J.B. Mann, Atomic structure calculations. II. Hartree-Fock wavefunctions and radial expectation values: hydrogen to lawrencium (1968). Los Alamos Scientific Laboratory Report LA-3691
49. R. Pucci, N.H. March, J. Chem. Phys. **76**(12), 6091 (1982). DOI 10.1063/1.443013
50. R. Pucci, N.H. March, J. Chem. Phys. **78**(5), 2466 (1983). DOI 10.1063/1.445051
51. R. Pucci, N.H. March, Int. J. Quantum Chem. **29**(4), 949 (1986). DOI 10.1002/qua.560290431
52. N.H. March, Proc. Cambridge Philos. Soc. **48**, 665 (1952). DOI 10.1017/S0305004100076441. Reprinted in Ref. [76]
53. K.H. Hellwege (ed.), *Structure data of free polyatomic molecules* (Springer, Berlin, 1976)
54. R. Pucci, N.H. March, Phys. Rev. A **33**, 3511 (1986). DOI 10.1103/PhysRevA.33.3511
55. W. Su, J. Schrieffer, A. Heeger, Phys. Rev. B **22**(4), 2099 (1980). DOI 10.1103/PhysRevB.22.2099
56. E. Lieb, F. Wu, Phys. Rev. Lett. **20**(25), 1445 (1968). DOI 10.1103/PhysRevLett.20.1445
57. S. Mazumdar, S.N. Dixit, Phys. Rev. Lett. **51**, 292 (1983). DOI 10.1103/PhysRevLett.51.292

58. M. Baldo, R. Pucci, P. Tomasello, Int. J. Quantum Chem. **23**(3), 1111 (1983). DOI 10.1002/qua.560230335
59. M. Baldo, G. Piccitto, R. Pucci, P. Tomasello, Phys. Lett. A **95**(3–4), 201 (1983). DOI 10.1016/0375-9601(83)90835-6
60. R. Pucci, M. Baldo, A. Martín-Rodero, G. Piccitto, P. Tomasello, Int. J. Quantum Chem. **26**(5), 783 (1984). DOI 10.1002/qua.560260518
61. G. Del Re, in *Quantum chemistry of polymers: solid state aspects*, ed. by J. Ladik, J.M. André, M. Seel (Reidel, Dordrecht, 1984)
62. G. Del Re, in Pucci [7]
63. M. Seel, in Pucci [7]
64. F.C. Tompkins, *Chemisorption of gases on metals* (Academic Press, London, 1978)
65. D.M. Newns, Phys. Rev. **178**, 1123 (1969). DOI 10.1103/PhysRev.178.1123
66. M. Baldo, R. Pucci, F. Flores, G. Piccitto, A. Martin-Rodero, Phys. Rev. B **28**, 6640 (1983). DOI 10.1103/PhysRevB.28.6640
67. J.E. Inglesfield, J. Phys. C: Solid State Phys. **14**(26), 3795 (1981). DOI 10.1088/0022-3719/14/26/015
68. J.K. Nørskov, N.D. Lang, Phys. Rev. B **21**, 2131 (1980). DOI 10.1103/PhysRevB.21.2131
69. K.H. Johnson, in Avery and Dahl [22]
70. G. Piccitto, F. Siringo, M. Baldo, R. Pucci, Surf. Sci. **167**(2-133), 437 (1986). DOI 10.1016/0039-6028(86)90716-8
71. N.D. Lang, W. Kohn, Phys. Rev. B **1**, 4555 (1970). DOI 10.1103/PhysRevB.1.4555
72. G.A. Benesh, J.E. Inglesfield, J. Phys. C: Solid State Phys. **17**(9), 1595 (1984). DOI 10.1088/0022-3719/17/9/016
73. P. Nordlander, S. Holloway, J.K. Nørskov, Surf. Sci. **136**(1), 59 (1984). DOI 10.1016/0039-6028(84)90655-1
74. O. Gunnarsson, H. Hjelmberg, B.I. Lundqvist, Surf. Sci. **63**, 348 (1977). DOI 10.1016/0039-6028(77)90350-8
75. G.B. Blanchet, N. Dinardo, E.W. Plummer, Surf. Sci. **118**(3), 496 (1982). DOI 10.1016/0039-6028(82)90201-1
76. N.H. March, G.G.N. Angilella (eds.), *Many-body Theory of Molecules, Clusters, and Condensed Phases* (World Scientific, Singapore, 2009)

Chapter 22
Electron Density, Kohn–Sham Frontier Orbitals, and Fukui Functions

Weitao Yang, Robert G. Parr and R. Pucci

In this note we shall show that the ground-state electron density $\rho(\mathbf{r})$ is a functional of the highest occupied orbital in Kohn–Sham [1] theory, $\psi_{\max}$. The functionals $\rho[\psi_{\max}]$ for an $(M + \delta)$-electron system are resolved into three cases and connected to three Fukui functions defined by Parr and Yang [2].

We follow the formalism of Janak [3] and define a self-consistent problem for a system with continuous occupation number,

$$\left[-\frac{1}{2}\nabla^2 + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}) \right] \psi_i = \varepsilon_i \psi_i, \tag{22.1}$$

$$\rho(\mathbf{r}) = \sum_i n_i |\psi_i(\mathbf{r})|^2, \tag{22.2}$$

also

$$E[\rho] = T_s[\rho] + U[\rho] + E_{xc}[\rho]. \tag{22.3}$$

W. Yang · R.G. Parr
Department of Chemistry, University of North Carolina, Chapel Hill, NC 27514, USA

R. Pucci (✉)
Dipartimento di Fisica e Astronomia, Università di Catania, Via S. Sofia, 64,
95123 Catania, Italy
e-mail: renato.pucci@ct.infn.it

R. Pucci
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci
IMM-CNR, UdR, Via S. Sofia, 64, 95123 Catania, Italy

In these equations, $v_H(\mathbf{r}) = \delta U[\rho]/\delta\rho$ is the classical electrostatic potential and $v_{xc}(\mathbf{r}) = \delta E_{xc}[\rho]/\delta\rho$ is the exchange-correlation potential;

$$T_s[\rho] = \sum_i n_i \int \psi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\psi_i(\mathbf{r})d\tau$$

is the Kohn–Sham noninteracting kinetic energy functional.

Recent results of Perdew et al. [4] show that the highest occupied orbital energy of the Kohn–Sham formalism is identical with the chemical potential $\mu = \varepsilon_{max}$. We have therefore for the ground state of an $(M + \delta)$-electron system [5]: $\mu^+ = \varepsilon_{M+1}$, for $\delta > 0$; $\mu^- = \varepsilon_M$, for $\delta < 0$; $\mu^0 = (\varepsilon_{M+1} + \varepsilon_M)/2$, for $\delta = 0$. In the same way that March and Pucci [6] proved that $\rho(\mathbf{r})$ is a functional of ψ_{HOMO} in Hartree–Fock theory, we demonstrate that $\rho(\mathbf{r}) = \rho[\psi_{max}]$ for Kohn–Sham theory.

For a slightly negative ion ($\delta > 0$), Eq. (22.1) for the highest occupied orbital becomes

$$\left(-\frac{1}{2}\nabla^2 + v_H^+(\mathbf{r}) + v_{xc}^+(\mathbf{r})\right)\psi_{M+1}^+ = \varepsilon_{M+1}^+\psi_{M+1}^+, \tag{22.4}$$

where ψ_{M+1}^+ is the $(M + 1)$th Kohn–Sham orbital with eigenvalue ε_{M+1}^+ for the $(M + \delta)$-electron system. As $\delta \to 0$, ψ_{M+1}^+ approaches ψ_{M+1}, the $(M + 1)$th orbital for the M-electron system. The Euler equation for $E[\rho]$ in Eq. (22.3) to be minimum, under the constraint that ρ integrates to $M + \delta$, is

$$\frac{\delta T_s[\rho^+]}{\delta\rho} + v_H^+(\mathbf{r}) + v_{xc}^+(\mathbf{r}) = \mu^+. \tag{22.5}$$

From Eqs. (22.4) and (22.5), we find for a slightly negative ion,

$$\frac{\delta T_s[\rho^+]}{\delta\rho(\mathbf{r})} = -\frac{1}{2}\frac{\nabla^2\psi_{M+1}^+}{\psi_{M+1}^+}. \tag{22.6}$$

We conclude that $\rho^+(\mathbf{r})$ is a functional $\rho[\psi_{M+1}^+]$ for a slightly negative ion, since the left-hand side of Eq. (22.6) is a functional of ρ, while the right-hand side is an explicit functional of ψ_{M+1}^+.

For a slightly positive ion ($\delta < 0$), Eq. (22.1) for the highest occupied orbital is

$$\left(-\frac{1}{2}\nabla^2 + v_H^-(\mathbf{r}) + v_{xc}^-(\mathbf{r})\right)\psi_M^- = \varepsilon_M^-\psi_M^-. \tag{22.7}$$

The same procedure as above leads to

$$\frac{\delta T_s[\rho^-]}{\delta\rho(\mathbf{r})} = -\frac{1}{2}\frac{\nabla^2\psi_M^-}{\psi_M^-}. \tag{22.8}$$

Thus $\rho^-(\mathbf{r})$ is $\rho[\psi_M^-]$.

Now we consider a neutral system. The Euler equation is

$$\mu^0 = \frac{\delta T_s[\rho^0]}{\delta \rho} + v_H^0(\mathbf{r}) + v_{xc}^0(\mathbf{r}). \tag{22.9}$$

From Eqs. (22.4) and (22.7) we have

$$
\frac{\varepsilon_M^- + \varepsilon_{M+1}^+}{2} = -\frac{1}{4}\left(\frac{\nabla^2 \psi_M^-}{\psi_M^-} + \frac{\nabla^2 \psi_{M+1}^+}{\psi_{M+1}^+}\right)
$$
$$
+ \frac{1}{2}\left(v_H^+(\mathbf{r}) + v_H^-(\mathbf{r})\right) + \frac{1}{2}\left(v_{xc}^+(\mathbf{r}) + v_{xc}^-(\mathbf{r})\right). \tag{22.10}
$$

Since the classical electrostatic potential is continuous as $\delta \to 0$, and $v_{xc}^0(\mathbf{r}) = \frac{1}{2}\lim_{\delta\to 0}[v_{xc}^+(\mathbf{r}) + v_{xc}^-(\mathbf{r})]$, even though $v_{xc}(\mathbf{r})$ may contain a discontinuity as $\delta \to 0$ [7], comparison of Eqs. (22.9) and (22.10) at limit $\delta \to 0$ gives

$$\frac{\delta T_s[\rho^0]}{\delta \rho} = -\frac{1}{4}\left(\frac{\nabla^2 \psi_M}{\psi_M} + \frac{\nabla^2 \psi_{M+1}}{\psi_{M+1}}\right), \tag{22.11}$$

and we conclude that $\rho^0[\psi_M, \psi_{M+1}]$.

Another way to obtain these results is to note that the Kohn–Sham effective potential $v_H + v_{xc}$ is determined (up to a constant) by any one of the solutions of Eq. (22.1); hence M and one of the orbitals determine the total ρ. Equation (22.6) contains specific information about this functional relation.

In summary, we have proved in Janak's generalization of Kohn–Sham theory for a system of $(M + \delta)$ electrons, the general functional relations, in the limit $\delta \to 0$: $\rho^+ = \rho[\psi_{M+1}] = \rho[\psi_{\text{LUMO}}]$ for $\delta > 0$, $\rho^- = \rho[\psi_M] = \rho[\psi_{\text{HOMO}}]$ for $\delta < 0$, $\rho^0 = \rho^0[\psi_{M+1}, \psi_M] = \rho^0[\psi_{\text{LUMO}}, \psi_{\text{HOMO}}]$ for $\delta = 0$. The total electron density is a functional of the appropriate frontier orbital or orbitals.

It follows immediately that the Fukui functions f^+ and f^-, defined by Parr and Yang [2] as the right and left limits of the function $\partial \rho(\mathbf{r})/\partial N$, and f^0 as the average of f^+ and f^-, to characterize chemical reactivity of molecules, are functionals of the appropriate frontier orbital(s).

We may derive an explicit expression for the Fukui functions in terms of Kohn–Sham orbitals. Let us consider the ground-state density of an $(M + \delta)$-electron system with $\delta > 0$. Its occupation numbers have to satisfy [3] $n_i = 1$ for $i \le M$, and $n_{M+1} = \delta$. That is to say,

$$\rho_{M+\delta}(\mathbf{r}) = \sum_{i=1}^{M} |\psi_i^+(\mathbf{r})|^2 + \delta |\psi_{M+1}^+(\mathbf{r})|^2, \tag{22.12}$$

so that

$$f^+(\mathbf{r}) = \lim_{\delta \to 0^+} \frac{\partial \rho_{M+\delta}(\mathbf{r})}{\partial N} = |\psi_{M+1}(\mathbf{r})|^2 + \sum_{i=1}^{M} \frac{\partial}{\partial N}|\psi_i(\mathbf{r})|^2. \tag{22.13}$$

Similarly, for an $(M + \delta)$-electron ground state with $\delta < 0$, $n_i = 1$ for $i \leq M - 1$, and $n_M = 1 + \delta$. Thus,

$$f^-(\mathbf{r}) = \lim_{\delta \to 0^-} \frac{\partial \rho_{M+\delta}(\mathbf{r})}{\partial N} = |\psi_M(\mathbf{r})|^2 + \sum_{i=1}^{M-1} \frac{\partial}{\partial N}|\psi_i(\mathbf{r})|^2. \tag{22.14}$$

All the above results put the 'frozen-core' approximations [2] of equating the Fukui functions to the corresponding frontier densities on a sound mathematical basis: they are well-defined first approximations. The importance of this is that $f(\mathbf{r})$ is identically the quantity $\delta\mu/\delta v(\mathbf{r})$ in density functional theory [5].

References

1. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965). DOI 10.1103/PhysRev.140.A1133
2. R.G. Parr, W. Yang, J. Am. Chem. Soc. **106**(14), 4049 (1984). DOI 10.1021/ja00326a036
3. J.F. Janak, Phys. Rev. B **18**, 7165 (1978). DOI 10.1103/PhysRevB.18.7165
4. J.P. Perdew, R.G. Parr, M. Levy, J.L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982). DOI 10.1103/PhysRevLett.49.1691
5. R.G. Parr, L.J. Bartolotti, J. Phys. Chem. **87**(15), 2810 (1983). DOI 10.1021/j100238a023
6. N.H. March, R. Pucci, J. Chem. Phys. **78**(5), 2480 (1983). DOI 10.1063/1.445054
7. J.P. Perdew, M. Levy, Phys. Rev. Lett. **51**, 1884 (1983). DOI 10.1103/PhysRevLett.51.1884

Part III
Theoretical Physics

Chapter 23
From Condensed Matter to QCD: A Journey Through Gauge Theories on Board of a Variational Tool

Fabio Siringo

Abstract Starting with a review of the thermal fluctuations in superconductors, the Gaussian Effective Potential is shown to be a powerful variational tool for the study of the breaking of symmetry in gauge theories. A novel rederivation of the massive expansion for QCD is presented, showing its variational nature and its origin from the Gaussian potential that also provides a variational proof for chiral symmetry breaking and dynamical generation of a gluon mass.

Introduction

Since 1873, when Lord Rayleigh [1] described a variational method for calculating the frequencies of mechanical systems, the Rayleigh–Ritz method has become an important tool for the approximate solution of physical problems in quantum mechanics and quantum field theory. My personal experience on variational methods dates back to 1985, when I was a graduate student of Professor Renato Pucci's. He proposed to put an hydrogen molecule inside a rigid box and evaluate the energy. His key idea was the insertion of a dielectric constant for simulating the effects of the other molecules as if it were in a very dense phase under high pressure [2]. That very physical idea was successful and since then Renato Pucci has been contributing to the physics of solids under pressure with many model calculations based on the remarkable physical ideas. Hydrogen was believed to become a superconductor in its solid phase under pressure and the fascinating Anderson-Higgs mechanism of gauge symmetry breaking was one of the milestones in Professor Pucci's teaching. That is where my personal journey has begun, going from the scalar $U(1)$ gauge theory of superconductivity [3, 4], through the $SU(2) \times U(1)$ theory of weak interactions [5–9], up to $SU(3)$ theory and QCD [10–16]. Still collaborating with Renato Pucci in 2003, we found that a variational tool like the Gaussian Effective Potential (GEP) can describe the thermal fluctuations of a superconductor in its broken-symmetry

F. Siringo (✉)

Dipartimento di Fisica e Astronomia, Università di Catania, and INFN
Sezione di Catania, Via S.Sofia 64, 95123 Catania, Italy
e-mail: fabio.siringo@ct.infn.it

phase [3, 4]. While the same variational tool had been very successful for describing the breaking of symmetry in a scalar theory [17], its potentiality in the study of gauge theories were not fully explored yet. The idea was then developed through several papers attempting to enlarge the gauge group [5, 6], introduce fermions [18, 19] and eventually describe other mechanisms of symmetry breaking, like the chiral symmetry breaking of QCD [16, 20–23] where the gluon and quark masses emerge without any breaking of the gauge symmetry.

In this contribution, after reviewing the use of the GEP for the study of superconductivity [3, 4], the *massive expansion* [16, 21, 22] is rederived from the GEP and shown to be a powerful variational tool for addressing the problem of mass generation in Yang–Mills theories and QCD, even when the gauge symmetry is not broken. While the massive expansion provides an analytical description of the propagators of QCD from first principles [16, 21] and is in remarkable agreement with the data of lattice simulations [22, 23], its variational nature is hided and disguised to look like a perturbative method. The present novel and alternative derivation of the massive expansion illustrate its direct origin from the GEP. Moreover, in chiral QCD the GEP provides a variational proof of chiral symmetry breaking and dynamical generation of the gluon mass.

GEP and Superconductivity

The standard Ginzburg–Landau (GL) effective Lagrangian still provides the best framework for a general description of the phenomenology of $U(1)$ gauge symmetry breaking and superconductivity (the Anderson-Higgs mechanism). The GL action can be seen as a power expansion of the exact action around the critical point and is recovered by any microscopic theory around the transition. The Gaussian fluctuations can be studied by the GEP for $U(1)$ scalar electrodynamics in three space dimensions [3, 4], where it represents the standard static GL effective model of superconductivity.

At variance with the approach of Ibañez–Meier et al. [24] who computed the GEP by use of a general covariant gauge, we work in unitary gauge, in order to make the physical content of the theory more evident. It has been shown to be formally equivalent to a full gauge-invariant method once all the gauge degrees of freedom have been integrated out [25]. The variational method provides a way to evaluate both the correlation length ξ and the penetration depth λ as a solution of coupled equations. The GL parameter $\kappa_{GL} = \lambda/\xi$ is found to be temperature dependent in contrast to the simple mean-field description and its behavior turns out to be in perfect agreement with many experimental data [3, 4].

Let us consider the standard static GL action (see, e.g., [26])

$$S = \int d^3x \left[\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2}(D_\mu \phi)^*(D^\mu \phi) + \frac{1}{2} m_B^2 \phi^* \phi + \lambda_B (\phi^* \phi)^2 \right]. \quad (23.1)$$

where ϕ is a complex (charged) scalar field, its covariant derivative is defined according to

$$D_\mu \phi = \partial_\mu + i e_B A_\mu \tag{23.2}$$

and $\mu, \nu = 1, 2, 3$ run over the three space dimensions. The magnetic field components are $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. We may assume a transverse gauge $\nabla \cdot \mathbf{A} = 0$, and then switch to unitarity gauge in order to make ϕ real.

By a pure variational argument [4] the longitudinal gauge field can be integrated out yielding the effective action

$$S = \int d^3x \left[\frac{1}{2}(\nabla\phi)^2 + \frac{m_B^2}{2}\phi^2 + \lambda_B\phi^4 + \frac{e_B^2\phi^2 A^2}{2} + \frac{1}{2}(\nabla \times \mathbf{A})^2 + \frac{(\nabla \cdot \mathbf{A})^2}{2\varepsilon} \right]. \tag{23.3}$$

The partition function is expressed as a functional integral over the real scalar field ϕ and the generic three-dimensional vector field $\mathbf{A}$, with the extra prescription that the parameter ε is set to zero at the end of the calculation. As usual, the free energy (effective potential) follows by inserting a source term and by a Legendre transformation [3, 4].

The GEP may be evaluated by the δ expansion method [24, 27] and is a variational estimate of the exact free energy. We introduce a shifted field

$$\tilde{\phi} = \phi - \varphi \tag{23.4}$$

then we split the Lagrangian into two parts

$$\mathscr{L} = \mathscr{L}_0 + \mathscr{L}_{int} \tag{23.5}$$

where $\mathscr{L}_0$ is the sum of two free-field terms describing a vector field A_μ with mass Δ and a real scalar field $\tilde{\phi}$ with mass Ω:

$$\mathscr{L}_0 = \left[+\frac{1}{2}(\nabla \times \mathbf{A})^2 + \frac{1}{2}\Delta^2 A_\mu A^\mu + \frac{(\nabla \cdot \mathbf{A})^2}{2\varepsilon} \right] + \left[\frac{1}{2}(\nabla\tilde{\phi})^2 + \frac{1}{2}\Omega^2\tilde{\phi}^2 \right]. \tag{23.6}$$

The interaction then reads

$$\mathscr{L}_{int} = v_0 + v_1\tilde{\phi} + v_2\tilde{\phi}^2 + v_3\tilde{\phi}^3 + v_4\tilde{\phi}^4 +$$
$$+ \frac{1}{2}\left(e_B^2\varphi^2 - \Delta^2\right) A_\mu A^\mu + e_B^2\varphi A_\mu A^\mu \tilde{\phi} + \frac{1}{2}e_B^2 A_\mu A^\mu \tilde{\phi}^2 \tag{23.7}$$

where

$$v_0 = \frac{1}{2}m_B^2\varphi^2 + \lambda_B\varphi^4, \quad v_1 = m_B^2\varphi + 4\lambda_B\varphi^3,$$
$$v_2 = \frac{1}{2}m_B^2 + 6\lambda_B\varphi^2 - \frac{1}{2}\Omega^2, \quad v_3 = 4\lambda_B\varphi, \quad v_4 = \lambda_B. \tag{23.8}$$

The nonconventional splitting of the Lagrangian has two important effects: arbitrary mass parameters are inserted in the *free* part; mass counterterms are inserted in the interaction in order to leave the Lagrangian unmodified. Then the standard perturbation theory is used for determining the first-order effective potential. The sum of vacuum graphs up to first-order yields the free energy density

$$
\begin{aligned}
V_{\text{eff}}[\varphi] = {} & I_1(\Omega) + 2I_1(\Delta) + \\
& + \left[\lambda_B \varphi^4 + \frac{1}{2} m_B^2 \varphi^2 + \frac{1}{2} \left\{ m_B^2 - \Omega^2 + 12\lambda_B \varphi^2 + 6\lambda_B I_0(\Omega) \right\} I_0(\Omega) \right] \\
& + \left(e_B^2 \varphi^2 + e_B^2 I_0(\Omega) - \Delta^2 \right) I_0(\Delta)
\end{aligned}
\tag{23.9}
$$

where the divergent integrals I_n are defined according to

$$
I_0(M) = \int \frac{d^3k}{(2\pi)^3} \frac{1}{M^2 + k^2}, \quad I_1(M) = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} \ln(M^2 + k^2) \tag{23.10}
$$

and must be regularized somehow.

The free energy (23.9) now depends on the mass parameters Ω and Δ. Since none of them was present in the original GL action of Eq. (23.3), the free energy should not depend on them, and the *minimum sensitivity* method [28] can be adopted in order to fix the masses: the free energy is required to be stationary for variations of Ω and Δ. On the other hand the stationary point can be shown to be a minimum for the free energy and the method is equivalent to a pure variational method [24]. At the stationary point the masses give the inverse correlation lengths for the fields, the so-called coherence length $\xi = 1/\Omega$ and penetration depth $\lambda = 1/\Delta$.

The stationary conditions

$$
\frac{\partial V_{\text{eff}}}{\partial \Omega^2} = 0, \qquad \frac{\partial V_{\text{eff}}}{\partial \Delta^2} = 0 \tag{23.11}
$$

give two coupled gap equations:

$$
\Omega^2 = 12\lambda_B I_0(\Omega) + m_B^2 + 12\lambda_B \varphi^2 + 2e_B^2 I_0(\Delta) \tag{23.12}
$$

$$
\Delta^2 = e_B^2 \varphi^2 + e_B^2 I_0(\Omega). \tag{23.13}
$$

For any value of φ, the equations must be solved numerically and the minimum-point values Ω and Δ must be inserted back into Eq. (23.9) in order to get the Gaussian free energy $V_{\text{eff}}(\varphi)$ as a function of the order parameter φ. For a negative and small enough m_B^2, we find that V_{eff} has a minimum at a non zero value of $\varphi = \varphi_{\min} > 0$, thus indicating that the system is in the broken-symmetry superconducting phase. Of course the masses Ω, Δ only take their physical value at the minimum of the free energy $\varphi_{\min}$. That point may be found by requiring that

$$\frac{\partial V_{\text{eff}}}{\partial \varphi^2} = 0 \qquad (23.14)$$

where as usual the partial derivative is allowed as far as the gap equations (23.12), (23.13) are satisfied [17]. The condition (23.14) combined with the gap equation (23.12) yields the very simple result

$$\varphi_{\min}^2 = \frac{\Omega^2}{8\lambda_B}. \qquad (23.15)$$

However, we notice that here the mass Ω must be found by solution of the coupled gap equations. Thus Eqs. (23.15), (23.12) and (23.13) must be regarded as a set of coupled equations and must be solved together in order to find the physical values for the correlation lengths and the order parameter.

Insertion of Eq. (23.15) into Eq. (23.13) yields a simple relation for the GL parameter κ_{GL}

$$\kappa_{GL}^2 = \left(\frac{\lambda}{\xi}\right)^2 = \kappa_0 \frac{1}{1 + \dfrac{I_0(\Omega)}{\varphi_{\min}^2}} \qquad (23.16)$$

where $\kappa_0 = e_B^2/(8\lambda_B)$ is the mean-field GL parameter which does not depend on temperature. Equation (23.16) shows that the GL parameter is predicted to be temperature dependent through the nontrivial dependence of Ω and $\varphi_{\min}$. At low temperature, where the order parameter $\varphi_{\min}$ is large, the deviation from the mean-field value κ_0 is negligible. Conversely, close to the critical point, where the order parameter is vanishing, the correction factor in Eq. (23.16) becomes very important [3, 4].

It is instructive to look at the effective potential in the limit $\varphi \to 0$ of the unbroken-symmetry phase and at the *chiral* point $m_B = 0$ where the original Lagrangian is scaleless. In that limit Eq. (23.9) reads

$$V_{\text{eff}}[0] = [I_1(\Omega) + 2I_1(\Delta)] - \frac{1}{2}\left[\Omega^2 I_0(\Omega) + 2\Delta^2 I_0(\Delta)\right]$$
$$+ 3\lambda_B [I_0(\Omega)]^2 + e_B^2 I_0(\Omega) I_0(\Delta) \qquad (23.17)$$

and is a function of the mass parameters. Its minimum might fall at a finite set of masses Δ_0, Ω_0 yielding a generation of mass from a scaleless Lagrangian. That property turns out to be useful for addressing the problem of mass generation in chiral QCD, where the gauge symmetry is not broken. Moreover, we observe that all the terms in Eq. (23.17) arise from the sum of the vacuum graphs up to first order, as shown in Fig. 23.1, where the internal lines are the massive propagators that can be read from the free-particle Lagrangian $\mathscr{L}_0$ of Eq. (23.6). We obtain a *massive* expansion, with massive free-particle propagators in the loops, from a massless Lagrangian.

The Gaussian Effective Potential Revisited

The massive expansion can be seen as an expansion around the vacuum of massive particles. The search for the best vacuum is the aim of the GEP that has been studied by several authors, mainly in the context of spontaneous symmetry breaking and scalar theories. While the GEP is a genuine variational method, several extensions to higher orders have been proposed. However, being a first-order approximation, the GEP fails to predict any useful result for the fermions of the standard model, because of the minimal gauge interaction that requires a second-order graph at least [12, 19]. Even the idea of an expansion around the optimized vacuum of the GEP is not new [29] but has not been developed further.

In the next section, the pure variational nature of the GEP is used as a tool for demonstrating that the standard vacuum of QCD is unstable toward the vacuum of massive gluons and quarks. Expanding around the optimized vacuum we recover the massive expansion that has been recently developed for pure Yang–Mills theory [16, 20, 21]. Thus, the unconventional massive perturbative expansion can be seen to emerge from the GEP formalism in a natural way.

One of the important merits of the GEP is its paradox of being a pure variational method disguised as a perturbative calculation, making use of the standard graphs of perturbation theory. In this section, we set the formalism of the expansion, starting with the simple scalar theory and then moving toward Yang–Mills theory and QCD in the next section.

Let us revise briefly the method for the simple case of a self-interacting scalar theory [17] where the effective potential is given by three vacuum graphs as shown in Fig. 23.1 (to be compared with the six graphs of the scalar electrodynamics of section "GEP and Superconductivity"). The Lagrangian density reads

$$\mathscr{L} = \frac{1}{2}\phi\left(-\partial^2 - m_B^2\right)\phi - \frac{\lambda}{4!}\phi^4 \tag{23.18}$$

where m_B is a bare mass. We can then split the total Lagrangian as $\mathscr{L} = \mathscr{L}_0 + \mathscr{L}_{int}$ where the trial free part is
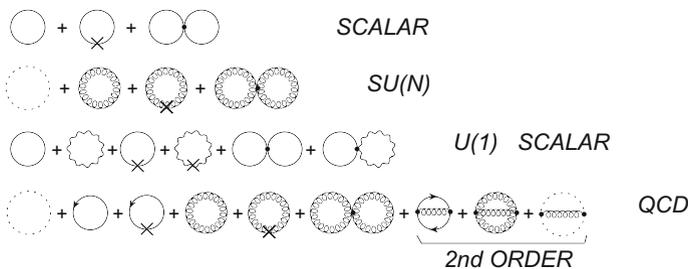
Fig. 23.1 Vacuum graphs contributing to the GEP for different theories

$$\mathscr{L}_0 = \frac{1}{2}\phi\left(-\partial^2 - m^2\right)\phi \tag{23.19}$$

and describes a free scalar particle with a trial mass $m \neq m_B$. The interaction follows as

$$\mathscr{L}_{int} = -\frac{\lambda}{4!}\phi^4 - \frac{1}{2}\left(m_B^2 - m^2\right)\phi^2 \tag{23.20}$$

so that the total Lagrangian has not been changed. If we neglect the interaction, then a free Hamiltonian $\mathscr{H}_0$ is derived from $\mathscr{L}_0$ and its ground state $|m\rangle$ satisfies

$$\mathscr{H}_0|m\rangle = E_0(m)|m\rangle \tag{23.21}$$

and depends on the trial mass m. Restoring the interaction $\mathscr{L}_{int}$, the full Hamiltonian reads $\mathscr{H} = \mathscr{H}_0 + \mathscr{H}_{int}$ and by standard perturbation theory, the first-order energy of the ground state reads

$$E_1(m) = E_0(m) + \langle m|\mathscr{H}_{int}|m\rangle = \langle m|\mathscr{H}|m\rangle \tag{23.22}$$

and is equivalent to the first-order effective potential $V_1(m)$ evaluated by standard perturbation theory with the interaction $\mathscr{L}_{int}$. Thus, the stationary condition

$$\frac{\partial V_1(m)}{\partial m} = \frac{\partial E_1(m)}{\partial m} = 0 \tag{23.23}$$

gives the best value of m that minimizes the vacuum energy of the ground state $|m\rangle$. While being a pure variational method, the first-order effective potential $V_1(m) = E_1(m)$ can be evaluated by the sum of all the vacuum graphs up to first order (the three loop graphs in Fig. 23.1). The resulting optimized effective potential is the GEP. Usually, the effective potential is evaluated for any value of the average $\varphi = \langle\phi\rangle$ and the best m also depends on that average. If the symmetry is not broken, then the minimum of the effective potential is at $\varphi = 0$ where $V_1(m)$ is a function of the trial mass, to be fixed by the stationary condition Eq. (23.23). We assume that the gauge symmetry is not broken in QCD so that $V_1(m)$ at $\varphi = 0$ is the effective potential we are interested in. The variational nature of the method ensures that the true vacuum energy is smaller than the minimum of $V_1(m)$. At the minimum, $|m\rangle$ provides an approximation for the vacuum and is given by the vacuum of a free massive scalar particle with mass equal to the optimized mass parameter $m \neq m_B$. Of course, the optimal state $|m\rangle$ is just a first approximation and the actual vacuum is much richer. However, we expect that a perturbative expansion around that approximate vacuum would be the best choice for the Lagrangian $\mathscr{L}$, prompting toward an expansion with an interaction $\mathscr{L}_{int}$ and a free part $\mathscr{L}_0$ that depend on m and can be optimized by a clever choice of the parameter m. Different strategies have been proposed for the optimization, ranging from the stationary condition of the GEP, Eq. (23.23), to Stevenson's principle of minimal sensitivity [28]. A method based on the minimal variance has been recently proposed for QCD and other gauge theories [10–14]. In all

those approaches, the underlying idea is that an optimal choice of m could minimize the effect of higher orders in the expansion. Since the total Lagrangian does not depend on m, the physical observables are expected to be stationary at the optimal m, thus suggesting the use of stationary conditions for determining the free parameter. As a matter of fact, if all graphs were summed up exactly, then the dependence on m would cancel in the final result, so that the strength of that dependence measures the weight of the neglected graphs at any order.

Leaving aside the problem of the best choice of m, we observe that at $\varphi = 0$ the calculation of the first-order effective potential $V_1(m)$ is quite straightforward and follows from the first-order expansion of the effective action $\Gamma(\varphi)$

$$e^{i\Gamma(\varphi)} = \int_{1PI} \mathscr{D}_\phi e^{iS_0(\phi+\varphi)+iS_{int}(\phi+\varphi)} \tag{23.24}$$

where the functional integral is the sum of all one-particle irreducible (1PI) graphs and $S = S_0 + S_{int}$ is the action. The effective potential then follows as $V_1(m) = -\Gamma(0)/\mathscr{V}_4$ where $\mathscr{V}_4$ is a total space-time volume. Moreover, being interested in the *chiral* limit, let set $m_B = 0$ in the interaction Eq. (23.20) and study a massless scalar theory.

The vertices of the theory can be read from $\mathscr{L}_{int}$ in Eq. (23.20) where we set $m_B = 0$ and are used in Fig. 23.1 in the vacuum graphs up to first order. The usual four-point vertex $-i\lambda$ is accompanied by the counterterm im^2 that is denoted by a cross in the graphs. This counterterm must be regarded as part of the interaction so that the expansion in not loopwise and we find one-loop and two-loop graphs summed together in the first-order effective potential. That is where the non-perturbative nature of the method emerges since the expansion in not in powers of λ but of the whole interaction $\mathscr{L}_{int}$. The zeroth-order (massive) propagator $i\Delta_m$ follows from $\mathscr{L}_0$

$$i\Delta_m(p) = \frac{i}{p^2 - m^2} \tag{23.25}$$

and is shown as a straight line in the vacuum graphs.

The tree term is the classical potential and vanishes in the limit $\varphi \to 0$. The first one-loop graph in Fig. 23.1 gives the standard one-loop effective potential, containing some effects of quantum fluctuations. It must be added to the second one-loop graph in Fig. 23.1, the crossed graph containing one insertion of the counterterm. It is instructive to see that the exact sum of all one-loop graphs with n insertions of the counterterm gives the standard vacuum energy of a massless particle. In other words, if we sum all the crossed one-loop graphs the dependence on m disappears and we are left with the standard one-loop effective potential of Coleman and Weinberg [30] $V_{1L}^0 = -\Gamma_{1L}^0/\mathscr{V}_4$ where Γ_{1L}^0 is the standard one-loop effective action at $\varphi = 0$

$$e^{i\Gamma_{1L}^0} = \int \mathscr{D}_\phi e^{i \int \frac{1}{2}\phi(-\partial^2)\phi d^4x} \sim \left[\mathrm{Det}(\Delta_0^{-1})\right]^{-1/2} \tag{23.26}$$
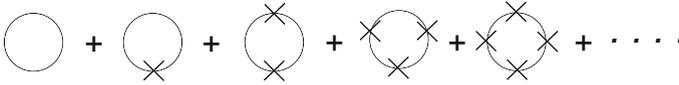
Fig. 23.2 Pictorial display of the *right* hand side of Eq. (23.28)

and $\Delta_0^{-1} = p^2$ is the free-particle propagator of a massless scalar particle.

Up to an additive constant, not depending on m, Eq. (23.26) can be written as

$$V_{1L}^0 = \frac{-i}{2\mathcal{V}_4} \operatorname{Tr} \log(\Delta_m^{-1} + m^2) \tag{23.27}$$

then expanding the log we obtain a massive expansion

$$V_{1L}^0 = \frac{-i}{2\mathcal{V}_4} \operatorname{Tr} \left\{ \log(\Delta_m^{-1}) + m^2 \Delta_m - \frac{1}{2} m^2 \Delta_m m^2 \Delta_m + \cdots \right\} \tag{23.28}$$

that is shown pictorially in Fig. 23.2 as a sum of crossed one-loop vacuum graphs. While the sum cannot depend on m, if we truncate the expansion at any finite order we obtain a function of the mass parameter. As a test of consistency, one can easily check that, once renormalized, the sum of all the crossed one-loop vacuum graphs in Fig. 23.2 gives zero exactly.

The calculation of the GEP requires the sum of only the first two terms of Eq. (23.28), the two one-loop graphs in Fig. 23.1. We cannot add higher order terms without spoiling the variational method since the average value of the Hamiltonian in the trial state $|m\rangle$ is $E_1(m) = V_1(m)$, according to Eq. (23.22). Using the identity

$$\Delta_m = -\frac{\partial}{\partial m^2} \log(\Delta_m^{-1}) \tag{23.29}$$

the sum of one-loop graphs in Fig. 23.1 can be written as

$$V_{1L}(m) = \left(1 - \frac{\partial}{\partial m^2}\right) I_1(m) = I_1(m) - \frac{1}{2} m^2 I_0(m) \tag{23.30}$$

where the diverging integrals $I_1(m)$, $I_0(m)$ generalize Eq. (23.10) and are defined as

$$I_1(m) = \frac{1}{2i} \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \log(-p^2 + m^2), \quad I_0(m) = -i \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \frac{1}{-p^2 + m^2} \tag{23.31}$$

so that

$$\frac{\partial I_1(m)}{\partial m^2} = \frac{1}{2} I_0(m). \tag{23.32}$$

We recognize $I_1(m)$ as the standard one-loop effective potential of Weinberg and Coleman for a massive scalar particle in the limit $\varphi \to 0$. This term contains the

quantum fluctuations at one-loop. The second term is a correction coming from the counterterm and arises because the exact Lagrangian was massless. The calculation of the GEP also requires the two-loop graph in Fig. 23.1 that is first-order in λ. A lazy way to evaluate it is by substituting the vertex im^2 in the crossed one-loop graph with the seagull one-loop self energy graph $-i\Sigma_{1L}$ that reads [12]

$$\Sigma_{1L} = \frac{\lambda}{2}I_0(m) \tag{23.33}$$

and adding a $1/2$ symmetry factor. The resulting two-loop term is

$$V_{2L}(m) = \frac{\lambda}{8}[I_0(m)]^2. \tag{23.34}$$

The GEP follows as the sum $V_{1L} + V_{2L}$

$$V_{\text{GEP}}(m) = I_1(m) - \frac{1}{2}m^2 I_0(m) + \frac{\lambda}{8}[I_0(m)]^2. \tag{23.35}$$

At this stage we just recovered the GEP in the limit $\varphi \to 0$ and Eq. (23.35) agrees with the well-known GEP in that limit [12, 17, 27, 31] (also compare to Eq. (23.17) by setting $4!\lambda_B = \lambda$, $\Omega = m$ and neglecting gauge field loops).

More precisely, V_{GEP} is the GEP when m is optimized by the stationary condition Eq. (23.23) that reads

$$\frac{\partial V_{\text{GEP}}(m)}{\partial m^2} = \frac{1}{2}\left(\frac{\partial I_0(m)}{\partial m^2}\right)\left[\frac{\lambda I_0(m)}{2} - m^2\right] = 0 \tag{23.36}$$

yielding the usual gap equation of the GEP

$$m^2 = \frac{\lambda I_0(m)}{2}. \tag{23.37}$$

From a mere formal point of view, the GEP predicts the existence of a mass for the massless scalar theory. That is of special interest because for $m_B = 0$ the Lagrangian in Eq. (23.18) has no energy scale, just like Yang–Mills theory and QCD in the chiral limit. Thus, it can be regarded as a toy model for the more general problem of mass generation and chiral symmetry breaking.

Actually, the integrals I_0, I_1 are badly diverging, and a mass scale arises from the regulator that must be inserted in order to get a meaningful theory. We can see that, in dimensional regularization, by setting $d = 4 - \varepsilon$, the integral I_0 is

$$I_0(m) = -\frac{m^2}{16\pi^2}\left[\frac{2}{\varepsilon} + \log\frac{\bar{\mu}^2}{m^2} + 1 + \mathcal{O}(\varepsilon)\right] \tag{23.38}$$

where $\bar{\mu} = (2\sqrt{\pi}\mu)\exp(-\gamma/2)$ is an arbitrary scale. Integrating Eq. (23.32) and neglecting an integration constant (that does not depend on m)

$$I_1(m) = -\frac{m^4}{64\pi^2}\left[\frac{2}{\varepsilon} + \log\frac{\bar{\mu}^2}{m^2} + \frac{3}{2} + \mathscr{O}(\varepsilon)\right]. \tag{23.39}$$

If we follow the usual approach of Coleman and Weinberg [30], the divergences must be absorbed by the physical renormalized parameters. Thus, let us define a *physical* renormalized energy scale Λ as

$$\log\Lambda^2 = \log\bar{\mu}^2 + \frac{2}{\varepsilon} + 1 \tag{23.40}$$

and write the integrals I_1, I_0 as simply as

$$I_0(m) = \frac{m^2}{16\pi^2}\log\frac{m^2}{\Lambda^2}$$
$$I_1(m) = \frac{m^4}{64\pi^2}\left[\log\frac{m^2}{\Lambda^2} - \frac{1}{2}\right]. \tag{23.41}$$

This approach is the same that is usually followed in lattice simulations of QCD: the lattice provides a scale that can be changed without affecting the physical scale which remains fixed at a phenomenological value. We assume that when $\varepsilon \to 0$ the scale $\bar{\mu}$ also changes, keeping Λ fixed at a physical value which cannot be predicted by the theory, but must come from the phenomenology.

First of all, we observe that by our renormalization scheme the standard one-loop effective potential is recovered, since that is equal to $I_1(m)$ in Eq. (23.41) and can be recognized as the mass-dependent term of the standard one-loop effective potential in the limit $\varphi \to 0$. That term has a relative maximum at $m = 0$, is negative for $m < \Lambda\exp(1/4)$ and has the absolute minimum at $m = \Lambda$. Thus, the one-loop effective potential would predict a massive vacuum if the symmetry were not broken and the physical vacuum were at $\varphi = 0$.

The full renormalized GEP is finite in terms of the physical scale Λ and can be written as

$$V_{\text{GEP}}(m) = \frac{\Lambda^4}{128\pi^2}U(\alpha, m^2/\Lambda^2) \tag{23.42}$$

where the adimensional potential $U(\alpha, x)$ is

$$U(\alpha, x) = x^2\left[\alpha(\log x)^2 - 2\log x - 1\right] \tag{23.43}$$

and α is the effective coupling $\alpha = \lambda/(16\pi^2)$.

The behavior of the potential $U(\alpha, x)$ is shown in Fig. 23.3. For any coupling α the point $x = 0$ is a relative minimum while the potential has a relative maximum at $x = 1/e$. The absolute minimum is at $x_0 = \exp(2/\alpha)$ where $U(\alpha, x_0) = -x_0^2 <$

Fig. 23.3 The adimensional potential $U(\alpha, x)$ of Eq. (23.43) is shown for different values of the effective coupling α
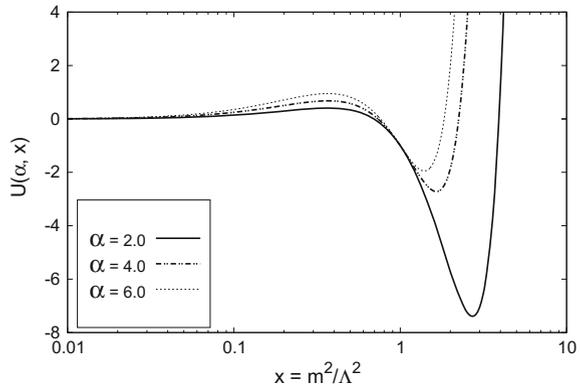
0. The two stationary points $x = 1/e$ and $x = x_0$ are the points, where the first or second factor in Eq. (23.36) is zero, respectively. Thus, the absolute minimum $m^2/\Lambda^2 = x_0 = \exp(2/\alpha)$ is the solution of the gap equation, Eq. (23.37). However, since the original theory has no scale, the quantitative value of m remains arbitrary as it depends on the unknown scale Λ. We can only predict that, since the GEP provides a genuine variational approximation for the vacuum energy, the massless vacuum must be unstable toward the vacuum of a massive scalar particle with an exact effective potential $V_{exact}(m) \leq V_{GEP}(m) < 0$.

In the chiral limit, the GEP can be easily extended to more complex theories by just adding up the graphs of Fig. 23.1. It is instructive to see how Eq. (23.17) can be recovered by the graphs for scalar $U(1)$ electrodynamics. In the next section, the GEP is evaluated for the chiral limit of QCD.

QCD in the Chiral Limit

The full Lagrangian of QCD, including N_f massless chiral Quarks, can be written as

$$\mathcal{L}_{QCD} = \mathcal{L}_{YM} + \sum_{i=1}^{N_f} \bar{\Psi}_i \left[i\partial\!\!\!/ - g\,A\!\!\!/_a \hat{T}_a \right] \Psi_i \tag{23.44}$$

where $\mathcal{L}_{YM}$ is the full $SU(N)$ Yang–Mills Lagrangian, including a covariant gauge-fixing term and the ghost terms arising from the Faddeev–Popov determinant. The generators of $SU(N)$ satisfy the algebra

$$\left[\hat{T}_a, \hat{T}_b \right] = i f_{abc} \hat{T}_c \tag{23.45}$$

with the structure constants normalized according to

$$f_{abc} f_{dbc} = N \delta_{ad}. \tag{23.46}$$

In a background field $\tilde{A}_a^{\mu}$ the effective action $\Gamma(\tilde{A})$ is the sum of 1PI graphs that can be formally given by the functional integral

$$e^{i\Gamma(\tilde{A})} = \int_{1PI} \mathscr{D}_{\Psi,A,\omega} e^{iS[A+\tilde{A},\Psi,\omega]} \tag{23.47}$$

where ω_a are the ghost fields. Assuming that the gauge symmetry is not broken, we are interested in the study of the limit $\tilde{A} \to 0$ and write the effective action as

$$e^{i\Gamma} = \int_{1PI} \mathscr{D}_{A,\omega} e^{iS_{YM}[A,\omega]+i\Gamma_{\Psi}[A]} \tag{23.48}$$

where S_{YM} is the action of pure Yang–Mills theory and the effective action Γ_{Ψ} is given by a functional integral over quark fields

$$e^{i\Gamma_{\Psi}[A]} = \int \mathscr{D}_{\Psi} e^{i \int \bar{\Psi} \hat{D}(A)\Psi d^4x} \tag{23.49}$$

with the operator $\hat{D}(A)$ that is given by

$$\hat{D}(A) = i\slashed{\partial} - g\slashed{A}_a \hat{T}_a. \tag{23.50}$$

The quark fields can be integrated exactly, yielding, up to a constant,

$$i\Gamma_{\Psi}(A) = \log \operatorname{Det} \hat{D}(A). \tag{23.51}$$

While S_{YM} contains the vertices of pure Yang–Mills theory, the expansion of Γ_{Ψ} in powers of gA_a^{μ} provides the standard insertions of quark–gluon vertices, yielding the usual Feynman rules of QCD. Some vacuum graphs, up to second order and two loops, are shown in Fig. 23.1.

As already noticed for the scalar theory, the calculation of the GEP requires the first-order effective potential that results from the sum of connected vacuum graphs up to first order. Thus, we may focus on the one-loop graphs in Fig. 23.1 and on the only first-order two-loop graph (the fourth for $SU(N)$ in Fig. 23.1). All other graphs are second order at least, starting from the other two-loop graphs of Fig. 23.1. Thus, at first order, the effective potential V_1 is just the sum of independent ghost, gluon and quark terms. This is an important limit of the GEP that cannot take in due account the second-order graphs, leaving us with a decoupled description of quarks, gluons, and ghosts. We can write the first-order effective potential as

$$V_1 = V_{YM} + V_{\Psi} \tag{23.52}$$

where the quark term contains only the one-loop zeroth-order vacuum graph that arises from Eq. (23.51) at $g = 0$

$$V_\Psi = \frac{i}{\mathscr{V}_4} \log \operatorname{Det} \hat{D}_0 \tag{23.53}$$

having defined the zeroth-order operator $\hat{D}_0 = i\hat{\partial}$ The Yang–Mills term V_{YM} is the first-order effective potential of pure Yang–Mills theory and can be written as

$$V_{\mathrm{YM}} = \frac{i}{\mathscr{V}_4} \log \int_{\mathrm{1st-order}} \mathscr{D}_{A,\omega} e^{i S_{\mathrm{YM}}[A,\omega]} \tag{23.54}$$

and is given by the one-loop ghost graph plus the one-loop and two-loop gluon graphs in Fig. 23.1.

At this stage, the whole calculation might seem to give trivial constant terms. However, we are interested in the *change* of these terms when a massive zeroth-order propagator is taken from the beginning for gluons and quarks. As already seen for the scalar theory, we have the freedom of adding a mass term in the zeroth-order Lagrangian provided that we subtract the same mass term in the interaction. The resulting massive expansion contains new two-point vertices (the mass counterterms) and their insertion in a graph does not change the number of loops but increases the order of the graph. Moreover, the first-order vacuum graphs in Fig. 23.1 remain uncoupled when any number of counterterms is inserted, so that we can study the change induced by the masses on V_Ψ and V_{YM} separately. It is instructive to see how the massive expansion [16, 21] of Yang–Mills theory emerges naturally in the calculation of the GEP and can be extended to chiral quarks.

Pure Yang–Mills Theory

In a generic linear covariant ξ-gauge, the first-order effective potential V_{YM} can be written as the sum of the second and fourth graph in Fig. 23.1, namely the zeroth-order gluon loop and the first-order two-loop graph which contains one insertion of the four-gluon vertex. We may drop the decoupled ghost loop that only gives an additive constant to the effective potential.

By the same notation of section "The Gaussian Effective Potential Revisited", we denote by V_{1L}^0 the one-loop graph that gives the standard one-loop effective potential in the limit of a vanishing background field

$$V_{1L}^0 = \frac{i}{\mathscr{V}_4} \log \int \mathscr{D}_A e^{i \int A_{a\mu} \Delta_0^{-1\,\mu\nu} A_{a\nu} \mathrm{d}^4 x} \tag{23.55}$$

containing the quadratic part of S_{YM} in Eq. (23.54) written in terms of the gluon propagator

$$\Delta_0^{\mu\nu}(p) = \Delta_0^T(p)t^{\mu\nu}(p) + \Delta_0^L(p)l^{\mu\nu}(p) \tag{23.56}$$

where $t^{\mu\nu}$, $l^{\mu\nu}$ are the transversal and longitudinal Lorentz projectors

$$t^{\mu\nu}(p) = g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2}, \qquad l^{\mu\nu}(p) = \frac{p^\mu p^\nu}{p^2} \tag{23.57}$$

and the corresponding free-particle scalar functions are

$$\Delta_0^T(p) = \frac{1}{-p^2}, \qquad \Delta_0^L(p) = \frac{\xi}{-p^2}. \tag{23.58}$$

The determinant of Δ_0 can then be written as a product of determinants in the orthogonal Lorentz subspaces, $\mathrm{Det}\,\Delta_0 = (\mathrm{Det}\,\Delta_0^T)(\mathrm{Det}\,\Delta_0^L)$, yielding

$$V_{1L}^0 = \frac{i}{2\mathcal{V}_4}\left[\mathrm{Tr}\log\Delta_0^T + \mathrm{Tr}\log\Delta_0^L\right]. \tag{23.59}$$

From now on, we work in the Landau gauge and take the limit $\xi \to 0$. In that limit $\Delta_0^L \to 0$ and the longitudinal part gives an (infinite) additive constant that we drop. The relevant part we will focus on reads

$$V_{1L}^0 = \frac{N_A}{2i}\int \frac{\mathrm{d}^4 p}{(2\pi)^4}\log\left(\Delta_0^{T-1}\right) \tag{23.60}$$

where N_A is a factor arising from the trace over color and Lorentz indices.

From now on, Following the same steps that lead to the GEP for a scalar theory, we may modify the quadratic part of the Lagrangian, i.e., Δ_0^{-1} in S_{YM}, provided that we add a counterterm to the total Lagrangian in order to leave it unchanged. Thus, we add a mass term to the transversal part Δ_0^T, leaving the longitudinal part unmodified. That would be a reasonable choice in any gauge since the longitudinal part Δ_0^L is left unmodified by the interaction at any order of perturbation theory. We define a new massive zeroth-order propagator Δ_m^T as

$$\Delta_m^{T-1} = \Delta_0^{T-1} + m^2 = -p^2 + m^2 \tag{23.61}$$

and insert the counterterm

$$\delta\mathcal{L}_c = m^2 t^{\mu\nu}A_{a\mu}A_{a\nu} \tag{23.62}$$

in the Lagrangian density. Then we look at the change of the first-order effective potential as a function of the mass parameter m, including the counterterm as a vertex of the theory. The result is formally equivalent to that obtained for the massless scalar theory in Eq. (23.28) and Fig. 23.2. By insertion of the counterterm, the one-loop gluon loop gives rise to an infinite sum of crossed loops where the straight line in Fig. 23.2 is now given by the massive propagator of Eq. (23.61) and the crosses

denote the insertion of a two-point vertex $-im^2 t^{\mu\nu}$. Even in a generic ξ-gauge the longitudinal part of the gluon propagator would not add any higher order contribution because of the transversal projector in the counterterm. Since everything is transversal in the Landau gauge, from now on we drop any projector $t^{\mu\nu}$ and the superscript T in the transverse propagator. Writing $\log(\Delta_0^{-1}) = \log(\Delta_m^{-1} - m^2)$ in Eq. (23.60) and expanding the log, the one-loop graph V_{1L}^0, that does not depend on m, reads

$$V_{1L}^0 = \frac{N_A}{2i} \int \frac{d^4 p}{(2\pi)^4} \left\{ \log \Delta_m^{-1} - \sum_{n=1}^{\infty} \frac{(m^2 \Delta_m)^n}{n} \right\}. \tag{23.63}$$

As before, in order to evaluate the GEP we must truncate the expansion and retain terms up to the first order, namely the zeroth-order gluon loop and the first-order crossed loop that are the first two graphs in Fig. 23.2. Then, at first-order, the one-loop effective potential is

$$V_{1L}(m) = N_A \left(1 - \frac{\partial}{\partial m^2} \right) I_1(m) \tag{23.64}$$

that is the same result of Eq. (23.30) scaled by the trace factor N_A.

The GEP also includes the two-loop first-order graph, the fourth graph in Fig. 23.1 with the propagator replaced by the massive propagator Δ_m and no insertions of the counterterm that would raise the order of the graph. By the same argument that leads to Eq. (23.34), the two-loop graph is easily evaluated by substituting the vertex $-im^2$ in the crossed one-loop graph with the one-loop seagull self energy graph $-i\Pi_{1L}$ that reads [15]

$$\Pi_{1L} = -\frac{9Ng^2}{4} I_0(m) \tag{23.65}$$

and adding a $1/2$ symmetry factor. The resulting two-loop term is

$$V_{2L}(m) = \frac{9N_A N g^2}{16} [I_0(m)]^2. \tag{23.66}$$

Adding the one-loop terms the GEP reads

$$V_{GEP}(m) = N_A \left\{ I_1(m) - \frac{1}{2} m^2 I_0(m) + \frac{9Ng^2}{16} [I_0(m)]^2 \right\} \tag{23.67}$$

which is exactly the same result of Eq. (23.35) for a scalar theory with an effective coupling $\lambda = 9Ng^2/2$, scaled by the trace factor N_A. Then by dimensional regularization, in the same scheme of section "The Gaussian Effective Potential Revisited", the GEP of pure Yang–Mills theory can be written as

$$V_{GEP}(m) = \frac{\Lambda^4 N_A}{128\pi^2} U(\alpha, m^2/\Lambda^2) \tag{23.68}$$

where the effective coupling $\alpha = \lambda/(16\pi^2) = 9N\alpha_s/(8\pi)$, $\alpha_s = g^2/(4\pi)$ and Λ is an unknown scale that must be fixed by the phenomenology. The a dimensional potential $U(\alpha, x)$ was defined in Eq. (23.43) and shown in Fig. 23.3.

Including Chiral Fermions

The inclusion of a set of chiral quarks is straightforward. As shown in Fig. 23.1, up to first order, the fermions are decoupled in the effective potential and we must just add the two one-loop graphs for the quarks. Let us derive them by the same method of section "Pure Yang–Mills Theory". For fermions, the standard one-loop effective potential V_ψ of Eq. (23.53) can be written as

$$V_\psi = \frac{i}{\mathscr{V}_4} \log \mathrm{Det}\left(\hat{D}_M + M\right) \tag{23.69}$$

where the *massive* inverse propagator $\hat{D}_M = \hat{D}_0 - M$ and the parameter M is an arbitrary trial quark mass. The exact expansion of Fig. 23.2 is recovered again as

$$V_\psi = \frac{i}{\mathscr{V}_4} \mathrm{Tr}\left[\log \hat{D}_M\right] + \frac{i}{\mathscr{V}_4} \mathrm{Tr}\left[\sum_{n=1}^{\infty} \frac{\left(\hat{D}_M^{-1}M\right)^n}{n}(-1)^{n+1}\right]. \tag{23.70}$$

yielding a massive expansion for the fermions. The GEP contains only graphs up to first order and is given by the first two terms, the two fermion loops in Fig. 23.1. The first term in the expansion, the zeroth-order loop, is

$$V_\psi^{(0)} = i\,\mathrm{Tr}\int \frac{\mathrm{d}^4 p}{(2\pi)^4} \log(\not{p}-M) = -4I_1(M) \tag{23.71}$$

while the second term, the crossed first-order loop, by Eq. (23.32) reads

$$V_\psi^{(1)} = -M\frac{\partial}{\partial M}V_\psi^{(0)} = 4M^2 I_0(M). \tag{23.72}$$

We observe that, without the crossed graph, the one-loop vacuum energy would be given by $V_\psi^{(0)} = -4I_1(M)$ which is unstable and unbounded from below according to Eq. (23.41). On the other hand, with one counterterm insertion, the first-order crossed graph makes the GEP bounded and yields the total first-order effective potential

$$V_\psi = -4\left[I_1(M) - M^2 I_0(M)\right] \tag{23.73}$$

which is exactly the GEP found in Ref. [32] by a direct variational method, provided that we take the chiral limit and set the external gluon field to zero. By dimensional

regularization, inserting Eq. (23.41), the quark contribution to the GEP reads

$$V_\psi(M) = \frac{3M^4}{16\pi^2}\left[\log\frac{M^2}{\Lambda^2} + \frac{1}{6}\right] \tag{23.74}$$

and has a minimum at $M_0^2 = \Lambda^2 e^{-2/3}$ where $V_\psi(M_0) = -3M_0^4/(32\pi^2) < 0$.

Discussion

Let us summarize the main findings of the previous sections.

The GEP for the GL model of superconductivity, namely $U(1)$ scalar electrodynamics, is recovered by a more general analysis based on a massive expansion, yielding a mass generation even when the original model is scaleless. The derivation of the GEP for pure $SU(N)$ Yang–Mills theory and chiral QCD also gives an original independent way to introduce the massive expansion: a change of the expansion point with massive propagators in the internal lines of the loops. The expansion acquires an evident variational meaning and emerges from the same variational argument that leads to the GEP. However, while the GEP is limited because of its first-order nature that leaves the fermions decoupled, in the massive expansion higher order terms can be easily included, yielding a powerful analytical tool for the study of QCD in the infrared and providing two-point functions that are in very good agreement with the results of lattice simulations [16, 21–23].

That said, the GEP gives a variational proof for chiral symmetry breaking and dynamical mass generation. Even if the actual values of the masses cannot be trusted because the quarks are decoupled, the variational nature of the calculation gives a proof that the vacuum of massless gluons and quarks is not stable. The Yang–Mills effective potential is given by the function $U(\alpha, x)$ of Eq. (23.43) and is shown in Fig. 23.3. An interesting feature is the occurrence of an unstable relative minimum at $m = 0$ and a stable minimum at $m > 0$. We could speculate and see an analogy with the double solution that occurs in the Dyson–Schwinger formalism: an unphysical massless scaling solution and a physical massive gluon propagator. Even if decoupled, the quark term of the GEP has an absolute minimum at a finite $M > 0$ according to Eq. (23.74), predicting the breaking of chiral symmetry of QCD.

We can see the absolute minimum of the GEP as a best expansion point for the massive expansion. In that sense, it is relevant to note that, once the crossed graph is included, the quark term of the GEP is also bounded from below. In fact, the counterterm keeps trace of the scaleless nature of the original Lagrangian and is needed for imposing that the Lagrangian is not modified in the expansion.

We are left with two independent mass parameters, m and M, that must be determined by the phenomenology since their explicit expressions depend on the unknown renormalized scale Λ. Assuming that the scale Λ is the same in the gluon and quark sector, which is not obvious, the minimum of the GEP would give a best ratio of

masses by Eqs. (23.68), (23.74)

$$\frac{M_0}{m_0} = e^{\left(\frac{1}{3} + \frac{2}{3\alpha}\right)}$$

(23.75)

linking together the dynamical generation of the gluon mass with the chiral symmetry breaking. While highly non-perturbative and nonanalytic in the limit $\alpha \to 0$, the suggested ratio of Eq. (23.75) suffers the limitations of the quark-gluon decoupling in the GEP and can be only regarded as a starting point for more refined calculations.

References

1. J.W.S. Rayleigh, *Theory of sound* (MacMillan, London, 1937)
2. R. Pucci, N.H. March, F. Siringo, J. Phys. Chem. Solids **47**(2), 231 (1986). DOI 10.1016/0022-3697(86)90134-4
3. M. Camarda, G.G.N. Angilella, R. Pucci, F. Siringo, Eur. Phys. J. B **33**(3), 273 (2003). DOI 10.1140/epjb/e2003-00166-7
4. L. Marotta, M. Camarda, G.G.N. Angilella, F. Siringo, Phys. Rev. B **73**, 104517 (2006). DOI 10.1103/PhysRevB.73.104517
5. F. Siringo, L. Marotta, Phys. Rev. D **78**, 016003 (2008). DOI 10.1103/PhysRevD.78.016003
6. F. Siringo, L. Marotta, Phys. Rev. D **74**, 115001 (2006). DOI 10.1103/PhysRevD.74.115001
7. F. Siringo, Phys. Rev. D **62**, 116009 (2000). DOI 10.1103/PhysRevD.62.116009
8. F. Siringo, Europhys. Lett. **59**(6), 820 (2002). DOI 10.1209/epl/i2002-00116-1
9. F. Siringo, L. Marotta, Int. J. Mod. Phys. A **25**(32), 5865 (2010). DOI 10.1142/S0217751X10051232
10. F. Siringo, L. Marotta, Eur. Phys. J. C **44**(2), 293 (2005). DOI 10.1140/epjc/s2005-02358-x
11. F. Siringo, Mod. Phys. Lett. A **29**(05), 1450026 (2014). DOI 10.1142/S0217732314500266
12. F. Siringo, Phys. Rev. D **88**, 056020 (2013). DOI 10.1103/PhysRevD.88.056020
13. F. Siringo, Phys. Rev. D **89**, 025005 (2014). DOI 10.1103/PhysRevD.89.025005
14. F. Siringo, Phys. Rev. D **90**, 094021 (2014). DOI 10.1103/PhysRevD.90.094021
15. F. Siringo, Phys. Rev. D **92**, 074034 (2015). DOI 10.1103/PhysRevD.92.074034
16. F. Siringo, Nucl. Phys. B **907**, 572 (2016). DOI 10.1016/j.nuclphysb.2016.04.028
17. P.M. Stevenson, Phys. Rev. D **32**, 1389 (1985). DOI 10.1103/PhysRevD.32.1389
18. L. Marotta, F. Siringo, Mod. Phys. Lett. B **26**(20), 1250130 (2012). DOI 10.1142/S0217984912501308
19. F. Siringo, Phys. Rev. D **86**, 076016 (2012). DOI 10.1103/PhysRevD.86.076016
20. F. Siringo, Perturbation theory of non-perturbative QCD (2015). Preprint arXiv:1507.05543
21. F. Siringo, Perturbative study of Yang-Mills theory in the infrared (2015). Preprint arXiv:1509.05891
22. F. Siringo, Analytic structure of QCD propagators in Minkowski space (2016). Preprint arXiv:1605.07357
23. F. Siringo, Universal scaling of gluon and ghost propagators in the infrared (2016). Preprint arXiv:1607.02040
24. R. Ibañez-Meier, I. Stancu, P.M. Stevenson, Z. Phys. C **70**(2), 307 (1996). DOI 10.1007/s002880050108
25. P. Mansfield, Nucl. Phys. B **267**(3), 575 (1986). DOI 10.1016/0550-3213(86)90133-1
26. H. Kleinert, *Gauge Fields in Condensed Matter* (World Scientific, Singapore, 1989)
27. I. Stancu, Phys. Rev. D **43**, 1283 (1991). DOI 10.1103/PhysRevD.43.1283

28. P.M. Stevenson, Phys. Rev. D **23**, 2916 (1981). DOI 10.1103/PhysRevD.23.2916
29. P. Cea, L. Tedesco, Phys. Rev. D **55**, 4967 (1997). DOI 10.1103/PhysRevD.55.4967
30. S. Coleman, E. Weinberg, Phys. Rev. D **7**, 1888 (1973). DOI 10.1103/PhysRevD.7.1888
31. I. Stancu, P.M. Stevenson, Phys. Rev. D **42**, 2710 (1990). DOI 10.1103/PhysRevD.42.2710
32. P.M. Stevenson, G.A. Hajj, J.F. Reed, Phys. Rev. D **34**, 3117 (1986). DOI 10.1103/PhysRevD.34.3117

Part IV
Philosophy and History of Science

Chapter 24
The Languages of Science, Religion, and Theology

G. Ruggieri

To Renato Pucci.

Introduction

At the end of his well-known conference on ethics, Ludwig Wittgenstein wrote [1]

«Ethics so far as it springs from the desire to say something about the ultimate meaning of life, the absolute good, the absolute valuable, can be no science. What it says does not add to our knowledge in any sense. But it is a document of a tendency in the human mind which I personally cannot help respecting deeply and I would not for my life ridicule it.»

Wittgenstein thus admitted that, in addition to those languages which are functional to the 'knowledge of things', there are also 'tendencies in the human mind' that are better expressed by means of something different than scientific language. The desire to express the ultimate meaning of life 'cannot be a science', but should be 'profoundly respected'. This statement by Ludwig Wittgenstein will be the starting point of my reflections on the diversity and the relationship among three different kinds of language, which nonetheless are tightly intertwined: the language of science, the language of religion, and the language of theology. Concerning the latter, let me hasten to clarify that I plan to speak exclusively of the language of theology that has developed within the Christian tradition. In fact, I am not sufficiently competent in other theological traditions (e.g. Buddhism, Islam, etc.), which I am acquainted with only superficially.

G. Ruggieri (✉)
Istituto per le Scienze Religiose Bologna, Bologna, Italy
e-mail: rotger@tin.it

The Language of Science

My own knowledge of the scientific language is only indirect. Indeed, it does not belong to my direct experience. I am only acquainted with it via the effects that it has and continues to have in the practice of the Christian faith. And I deem those effects quite positive, notwithstanding the temporary conflicts and the difficulty to understand the reach of their respective statements. For instance, the results of palaeontology on the origin of the species have modified the Christian view of the beginnings of human history [2], as well as the interpretation of the dogma of the so-called 'original sin' [3]. Even without denying that the existing universe is the result of the incessant creation of God, no theologian nowadays thinks that a 'specific and direct' intervention of God is required on the origin of the human species. And assertions which were earlier considered as absolutely necessary for the dogma of the original sin, such as the descendance from a single human couple, today are not considered so any longer, since a deeper understanding of the biblical language, induced by the very true scientific results, is now open even to a polygenistic or polyphyletic origin of man.

On the other hand, it is beyond doubt that also theology had an influence on scientific practice. There exists in this regard a study which is to me a 'classic', but which limits itself only to the seventeenth century [4]. On the contrary, there is no study on the theological assumptions of contemporary scientific practice. Indeed, I would be curious to understand whether the debate between relativity and quantum mechanics ultimately descends from some sort of metaphysical or religious assumptions [5]. However, the main conclusion is that for an actually existing man there is never a 'separate' language from all his other fundamental attitudes and their respective languages, whereas there does exist a contamination among the various languages. Einstein's famous motto «God does not play dice» is not a scientific statement, and I believe that as such it is not even the expression of a religious attitude, but it only signifies the contamination between scientific language and assumptions which are often unconscious—a contamination basically induced by the unity of the human subject, who 'says himself' through the different languages in the various approaches to the reality that he puts into being.

At any rate, the theologian, even without using the scientific language directly as his own, if he wishes to reflect on the necessary conditions to communicate the faith, then he must understand closely the logic surmised by such a language. This is the logic of the ἐπιστήμη (episteme), of the true knowledge. But the meaning of ἐπιστήμη is no longer that of Aristotle, who identified it with the knowledge of the causes of the phenomena (*cognitio rei per causas*). In my personal case, in order to get acquainted with the logic of the scientific practice, I relied on Karl Popper [6]. And my choice was motivated by my job as a theologian. Indeed, a follower of Popper's, Karl Albert, undertook a lively debate not only with the theological hermeneutics of Gerhard Ebeling, but also against a theology open towards the scientific practice,

such as that of Wolfhart Pannenberg.[1] In order to gain more insight, I have then read Popper's book [6], and I was both surprised by his criticism to the Circle of Vienna, and by the 'humble', albeit vigorous, character of his thinking. Scientific knowledge does not constitute the only understanding of reality, but distinguishes itself from all others by its criterion of 'truth', intended as the best possible approximation of the various theories to the actually existing, or objective, reality. Any scientific theory is, by its very nature and as long as it wishes to remain such, subject to the criteria of universal intersubjective control. Popper then speaks of the criterion of falsifiability (i.e. the possibility of establishing the false character of a given statement) in order to distinguish the 'true' language of science, i.e. the one which approaches the knowledge of reality through the contributions of everyone, from all the other languages, which cannot claim to be 'true'.

The Language of Religion

At variance with scientific language, which looks for a true knowledge of reality through ever more adequate approximations, the language of 'religion' is character- ized by its claim for 'absolute' statements on the sense of reality, thus necessitating a behaviour coherent with such statements. This is why religion is considered funda- mentally equal to moral in Kant's philosophy. During the first decades of the twentieth century, the positivism of the school of Vienna considered the assertions of religion as nonsense. On the other hand Wittgenstein, in his maturity, correctly noticed that it is not possible to provide an interpretation of language as a whole, since human language is a collection of expressions with quite different functions, depending on their respective contexts, and obeying rules which cannot be fixed once and for all (the so-called *Sprachspiele,* or 'linguistic games').

First of all one should therefore identify those real situations whose sense is conveyed by the language of religion. Such situations, as Ian T. Ramsey has in my opinion clarified, are those in which an illumination takes place, which breaks the rind of everyday observations and at the same time urges to act [9]. He proposes the following example: when one sees a child who is about to drown in a river, one perceives that such a fact would not be merely a consequence of human errors, whereto another observer would be indifferent, but that something deeper is going to take place, something which involves the fate and the happiness of several people, so that one is urged to immediately dive into the water to save the child, even at the risk of his own life. In such a situation there would be no certitude, however one's involvement, which is due to one's free will, is not senseless. Several centuries

---

[1]Cf. Ref. [7]. W. Pannenberg takes sides against Albert in *Wissenschaftstheorie und Theologie* (1973), only partly accepting his assumptions, and even conceding that God be only a hypothesis to be verified. However, Albert does not agree with Pannenberg's interpretation on the conjectural character of any science; cf. Ref. [8].

earlier, and in a completely different context, Blaise Pascal would have said: my bet is that the stake on which I am effectively compelled to wager is worth the bet even if I should lose, since placing such a bet would change anyway my life for better.[2]

The language which conveys what happens in these situations of discernment-engagement is the ethical-religious language, where one can see 'objects and behind', where 'behind' stands for the illumination, the discovery of a sense of reality which is observable not from the physical point of view, but within an inner illumination. To use an example belonging to the Christian context, while reading the Gospel's tales of the death of Jesus of Nazareth, through the facts described by the evangelical narrators, it becomes extremely clear to me the depth of Jesus' attitude without opposing any resistance, in order to be faithful to the mission that he had earlier recognized as God's will. This discernment, to the extent that it touches my conscience, induces my existence to adequate my behaviour to itself, and my whole life gets changed by it. The language used to convey such an experience is the one that my cultural environment makes available to me, however it is changed by some 'qualifiers' (in Ramsey's words), which modify it so as to make it adequate to express what I have experienced. Common language offers the 'templates' wherewith to think and communicate to the others my own experience (God as 'cause', God as 'love', etc.), whereas the 'qualifiers' indicate the difference, the specificity of the sense that I have caught (an 'ultimate' and not immediate cause, an 'infinite' and 'absolute', unconditioned love, etc.).

The ethical-religious language, in the above sense, is not peculiar to any specific religious practice, be it Christian or of other religions, from Buddhism to Islam, from animism to any refined theism. It is also present in 'nonreligious' forms, even in atheism, to the extent in which they are characterized by an 'absolute' engagement. The Greek philosopher Heraclitus said «Ἦϑος ἀνϑρώπῳ δαίμων» (fr. 119). In any way one translated the Greek term ἦϑος (*ethos*)—habit, abode (Heidegger), character (Pasquinelli)—it is sure that Heraclitus gathers exactly the constitutive way in which the human character is anchored to the divine. Man cannot avoid founding his choice between the good to be done, and the evil to be averted, in something that overhangs him as a maxim of which man cannot dispose (cf. Kant himself, in *Die Religion innerhalb der Grenzen der bloßen Vernuft,* 1793) [11].

---

[2]In Blaise Pascal's own words (*Pensées,* ed. Brunschvicg, n. 233): «Yes; but you must wager. It is not optional. You are embarked. Which will you choose then? Let us see. Since you must choose, let us see which interests you least. You have two things to lose, the true and the good; and two things to stake, your reason and your will, your knowledge and your happiness; and your nature has two things to shun, error and misery. Your reason is no more shocked in choosing one rather than the other, since you must of necessity choose. This is one point settled. But your happiness? Let us weigh the gain and the loss in wagering that God is. Let us estimate these two chances. If you gain, you gain all; if you lose, you lose nothing. Wager, then, without hesitation that He is.»[10].

The Language of Theology

As I have stated at the outstart, when speaking about the language of theology, I will mainly keep within the scopes of the Christian faith. Contrary to Ramsey, the ethical-religious language should not be confused with the language of theology. The word theology, in the current meaning of 'scientific' understanding of the experience of the Christian faith, was born at the beginning of the thirteenth century in connection with the foundation of the universities and under the influence of the Aristotelianism. However, a 'reflection', i.e. a second-order, reflected understanding, has always been accompanying the Christian experience, starting from the first writer of the New Testament, Paul of Tarsus, who made use of the rabbinic hermeneutics towards the interpretation of the Old Testament, and applied it to the Christian faith. Indeed, Gerhard Ebeling notices that theology develops from the encounter of faith with thought, because «faith by itself induces to comprehension, in a manner adequate to the situation of comprehension.» [12] It is not that faith does not involve some sort of thinking, but just because it has always involved a comprehension of reality, since it is an 'understanding' of God and of man and of the things, this comprehension, comparing itself with the other ones which characterize human existence in its specific historical substance, drives towards a further comprehension, thereby creating a new inner equilibrium of the human spirit.

The language of theology is subjected to the language of the believer (which, as a language, stays on the ethical-religious level), but at the same time modifies it. However, modifying it means to 'shift' it from the ethical-religious level, to the 'rational' level. This is a fact already acknowledged by mediaeval theologians, exactly when the language of the Aristotelian ἐπιστήμη begins to settle in the European universities. Suffice it to quote here Bonaventura da Bagnoregio (ca. 1217/1221–1274). Bonaventura develops his conception of theological reason both in the first four issues introducing his comment of the first book of the *Sentences* (*Liber Sententiarum,* by Peter Lombard), and in his own *Breviloquium.* What strikes most, in the foreword to the comment to the *Sentences* is the qualification of theological reason as simple *ratio probabilitatis.* Indeed, whereas what is believed has in itself the *ratio primae veritatis* and pertains to the faith, and whereas the *auctoritates* have in themselves their own *ratio* (*ratio auctoritatis*), which pertains to the Christian doctrine on the Scriptures, since, with Saint Augustine, its authority is greater than the perspicacity of the intellect, theology indeed adds only a *ratio probabilitatis* (cf. I Sent. Prœm. Q. 1 concl. 5.6). But this implies that, for Bonaventura, theology is part of a constellation centred in the act of faith, around which everything revolves, whose object contains in itself the *ratio primae veritatis.* The reason of theology's weakness is best understood starting from what he writes about the *determinatio distrahens.* With this category, again in the comment to the *Sentences* by Peter Lombard (Prœm. Q. 2), he explained the use of the *ratio* in the comprehension of the Scriptures. Indeed, the Scriptures are related to what is believed, the 'credible', in the parlance of Scholasticism. On the other hand, theology, by introducing reason in the consideration of the Scriptures, *distrahit,* i.e. effects a distraction, a displacement that reduces to the

Sacred Scriptures the principles of reason in the fashion of some sort of subordination. Indeed, Scriptures *sunt de credibili ut credibili,* while Scriptures in theology *sunt de credibili ut facto intelligibili.* And this *determinatio distrahit.* Indeed, what we believe, we owe it to the authority (of God), and what we understand, we owe it to reason (here Bonaventura quotes Augustine's *De utilitate credendi*). Thus the degree of certitude that one extracts from the Scriptures is superior to that of theology, which derives its certitudes from reason.

In the attempt to translate Bonaventura's thinking using a terminology closer to ourselves, I shall say that theology, as a rigorous reflection on faith and its history by believers, 'shifts' (dis–tracts) the attention from the object of the faith itself, which is an existential answer to a message, whose origin is perceived as divine, to concentrate such an attention on the ways in which such a message is perceived: e.g. on the peculiar signifiers employed to convey a meaning, on the historical character of the message, on the context within which the message must be placed, on the relationship between such a message and other analogous messages, etc. This distraction or shift of the attention takes place through 'reason', and this 'reason' is historically well-determined. The 'reason' employed to analyze the faith by an author of the fourth century AD was influenced by Stoicism and/or Neoplatonism. The reason employed by protestant theologians in the second half of the nineteenth century was mainly the Kantian-idealistic one on one side, and the critical-positivistic one in the analysis of literary documents on the other side. The reason employed by the first Catholic theologians of the so-called Tübingen school was mainly that of idealistic Romanticism, while in the second half of the nineteenth century it is the Neoscholastic reason to prevail among Catholics. But this implies that all rational use in theology, as 'human', should not overcome the threshold of what is probable.

This observation constitutes the main difficulty of every theological practice. Indeed, every theologian has the moral duty to use reason, since faith, albeit of divine origin, remains a totally human act, and therefore always risks to become devious, corrupt. Thus the task of a theologian is that of using all rational tools at his disposal in order to correct such deviations, to help other believers to let their own faith become ever more authentic. However, he is not alone in accomplishing this task. The faith of the believers is helped and gets stimulated by others than theologians: by prophets, by those who witness in an exemplary way faith itself in their own lives, by those who have the responsibility of keeping in unity all the believers, and so forth. The theological practice is thus just an element of a collection of practices that have always marked the history of faith. Therefore, conflicts are inevitable. But they are useful, rather than dangerous, for the growth of everyone. There remains the possibility, however, that, without the use of reason, the community of the believers risks the fundamentalism, the fanaticism, the intolerance.

But, above all, theological practice is necessary for the communication of faith. Actually, the problem remains of the communicability of this language or, to use E. Jüngel's words, of the communicability of the 'Gospel as an analogous discourse on God' [12], within the other human languages, wherewith we articulate our fundamental approaches to reality, ranging from science to art, from law to ethics

(which should of course be distinguished from the Gospel of Jesus of Nazareth),[3] and so forth. Or, if you wish, the problem is that of the analogy of languages in their reciprocal irreducibility.

Is Communication Possible Among the Various Languages?

First of all one should stress again that only secondarily is the theologian the one who communicates the faith. The subject of such a communication is primarily the believer as such, as it is clear from the history of Christianism and its diffusion since the origins. With respect to such a communication having the believer as the main subject, the theologian plays a completely secondary and subordinate role. It is probably better to say that he can theologize only in harmony with the other subjects of the faith [14].

Theology should therefore respect the intimate reason of the faith, its original λόγος, the reason that Bonaventura calls *ratio primae veritatis,* which we might translate as the 'self-evidence of the truth which is God'. Faith indeed possesses its 'own' reason, with its own evidence, as appearing in the immediate language of faith, in whose original narrative structure a truthful claim is already present. Even in its simplest forms, in fact, the faith constitutes a linguistic event with its own λόγος, presumed by the very act of theologizing. And obviously I am not referring to the descriptions, albeit authoritative, of faith, such as that of the constitution *Dei Filius* of Vatican I, all oriented towards the intellectual adhesion to the revealed truths, or that of the *Dei Verbum,* with a different attention towards the personalistic dimension. It is actually the very experience of faith in itself that represents already a linguistic event, the saying (λέγειν) of a λόγος, i.e. the articulation of a 'reason' constituted by the event of Jesus of Nazareth, seen—whichever way one poses the emphasis, according to sensibility, period of time, culture—as a constitutive reference to one's own existence. Not any intervening description of faith can constitute the linguistic event of faith itself. The saying (λέγειν) of the Christian λόγος rather is all in the original experience, in the act of faith, since this is the kindling of a relationship with an event that precedes me and arrives to me only in its narrations, be they made explicit, as in preaching, in the announcement, in the catechesis, be they simply lived without words, as in the witnessing of the deed, however made intelligible thanks to their link with a horizon of otherwise known signs.

To this observation one should add another, namely that theology as such, characterized by the addition to the *fides* of the *ratio* commonly agreed upon by men of their own time, finds its own place in an order, a τάξις, wherein only it does receive its depth and specific function. Should it for instance transform its own *ratio* in that of faith itself, then both theology and faith would be corrupted at the same time.

---

[3]Recently, I have expressed my own view on the relationship between Gospel and ethics in my essay on *Gospel, moral, and civil law* [13].

How does one communicate the *ratio primae veritatis*? That is, how does the believer communicate this *ratio* to all the other *rationes*? The place of communication is language, and only an analysis of the intimate nature of language can answer this question. In this context, I have always found illuminating two early writings by Walter Benjamin [15, 16].

To Benjamin, language is the communicability itself of spiritual beings. In the spiritual being which is man, the communication of everything, of all reality, takes place through the language. More than that, the linguistic essence of things is their language. Therefore, the linguistic essence of man is his language. But the language of man speaks through words. Therefore man communicates his spiritual essence (since it is communicable as such) saying all the things through words, giving names to things. In this 'giving names to things', indeed, there is not only the communication of things to man, but also the communication of man to himself and to the others. Within every language one should find indeed the pure language, the *reine Sprache.* This, which is inseparably found in every language, is the assumption of any possible 'translation' and 'communication' from one language to another. This pure language is to Benjamin a relationship, a *Nennen,* it amounts to saying the name, and not just a definition of the object, a circumscription of its meaning. Now, man does not communicate his spiritual being through the names that he gives to things, but rather in themselves, in the act by which he gives those names. The conception according to which man communicates an object to another man via words is considered secondary by Benjamin. This conception can only be applied to the language of things, which are communicated through the saying of man, through his giving names to things. However, in the pure language of man there is no means, no object, no recipient of the communication, since 'in the name is the spiritual being communicated to God' [16]. It seems to me that, to Benjamin, God as absolute Spirit, and therefore original power of any spirituality/communication, is within the communication of man.

From Benjamin's observation we can draw the main conclusion that there is a common root to the various languages, even the most distant from one another, such as that objectifying of science and that invoking of faith, which is aware of tending to an ever greater Truth. This common root, Benjamin's 'pure language', is the basis of their analogy. And this basis resides in the fact that in all linguistic acts, man expresses himself as opening to the other by himself and ultimately in his 'saying himself to God' (albeit unconsciously, and even when negating the name of God). The believer does not communicate to the other the object of his own faith, but in witnessing his own hope he discloses to the other the λόγος that moves his opening, that same λόγος inducing the other towards the search for knowledge, as adequate as possible, for the law which moves everything, or for the beauty that shines in the world, or for the right that could eliminate violence from any relationship, and so forth.

To gain a deeper insight into Benjamin's thought, a further note is appropriate at this point. The double dimension of the language (the objectifying one, and the spiritual one of the 'pure language') is not equally 'loud'. We will always perceive the sound of the objectifying dimension which is present in the act of *Nennen,* of giving a name to reality, but we do not always perceive the sound of the pure

language. Such a perception is not an obvious event. There exists a language which opens, and one which closes, depending on the circumstances and the ways in which it is employed. In other words, not all linguistic acts in its actual happening allow the pure language to be unveiled.

It is on this particular point that Western theology is playing a big match. This, in my opinion, is linked with the very birth of the 'scientific' language of Scholastic theology. I would like to illustrate this point with a comparison that may seem incorrect, at a first sight: that between Anselm's *Proslogion* and one of the questions in Saint Thomas Aquinas' *De potentia*. Thus, Anselm in his *Proslogion* takes a definitive turn, when he overcomes the register of the objectifying knowledge and addresses the reader beyond this objectifying knowledge (cf. Ref. [17], pp. 90–94). This is not always noticed in the 'philosophical' interpretations of his writing. In this objectifying knowledge, indeed, after the break of Chap. 14, the ensuing chapter makes a leap forward. One abandons the register of understanding or, better, the register of understanding appears only as a step to be passed, coherently and without the *sacrificium intellectus,* to a completely different register. Thus, Chap. 15 reads:

> «Therefore, O Lord, not only are You that than which a greater cannot be thought, but You are also something greater than can be thought.»

Thus, he recovers the non objectifying dimension of language that had been exalted especially in the works of Dionysius the Aeropagite. At the beginning of the sixth century, the latter wrote indeed a pamphlet, entitled *De mystica theologia,* which was going to have a great influence not only on the forthcoming Eastern theology, but also on the Latin Middle Ages. There, it was stated that our understanding is necessary to ascend to God, but that our conceptions, even the highest and the most divine, are not God, but only the acme of our intellectual labour, overwhelmed by a presence which transcends them.[4] However Thomas transmitted to the West a distorted vision of the Aeropagite's thought. In fact, the Aquinas interprets the three dionysian ways (the negative way: one should negate any attribute to God which involves imperfection; the positive way: one can assert any attribute to God which implies perfection; the way of eminence: even in the positive assertions, one should speak in such a way that everything is realized in Him to the highest level, higher than for any other creature) as different knowledge methodologies of the mystery of God, which find their realization in an effective correspondence between the concept and its actual objective reference in the divine substance. Thus, to Thomas the question is whether *nomina* such as good, right and wise *significant divinam substantiam* or not.[5] On the other hand, to Dionysius, and to the great majority of Eastern tradition, the function of the theological concept, i.e. of the *nomina* used in theology, is different.

In Thomas one can also find examples of a less 'objectifying' conception of the theological knowledge. Thomas himself, in fact, while quoting Dionysius, asserts that theology proceeds «not only by learning (μαθ ὤν), but also by suffering (παθ ὤν)

---

[4]PG 3, p. 997 ff. (Ref. [18]).

[5]*De potentia,* q. 7 a. 5.

divine realities.»[6] But it cannot be denied that the occidental equilibrium in theology will be diverted ever more towards *discere,* i.e. learning, and ever less towards *pati,* i.e. the experience of harbouring the presence of God. But it is the sound of the 'pure language' which emits suitable vibrations to be in harmony with God and with the others, and not the objectifying one of the *Nennen,* of the *nomina* which we give to reality.

These observations tell us then that every language, as far as together with the objectifying of naming one is given with the 'pure language', may become a place for Christian communication. This is at least the believer's conviction, and therefore also of the theologian. Indeed, the believer is aware that Jesus Christ represents the definitive Yea of God to man, not abstractedly, but in its well-defined historicity, even though plagued by sin:

> «For while we were still weak, at the right time Christ died for the ungodly. Indeed, rarely will anyone die for a righteous person—though perhaps for a good person someone might actually dare to die. But God proves his love for us in that while we still were sinners Christ died for us.» (Rom 5, 6–8).

Thanks to the absolute character of the relationship set by Christ with the historically determined man, every human situation is 'called' by God, every historically determined language, and not only in its abstract essence, can then become a place for the communication of the Christian truth, and is therefore suitable for a translation of the Gospel of Jesus of Nazareth.

Due to the final and ultimate (eschatological) character of the relationship set by Christ as God's and man's truth, the assertion that all languages may be assumed as a place for the translation of the Christian truth, is equivalent to saying implicitly that the 'ultimate' truth of every experience, expressed in its specific language, is somehow in its 'possible' relationship with the λόγος of the Christological event, and therefore with its being 'called' by God. This last and profound attitude of every language is the one which enables the translation (and therefore the interpretation) of the experience of Jesus of Nazareth in the experience and in the language of faith (from the first disciples, up to ourselves, in the historical succession of testimony), thanks to the action of the Spirit. This translation does not belong as such to theology, and is substantially equivalent to conversion, in the Christian-biblical parlance. However it is also a linguistic conversion, effected according to different rhythms and timings (substantially identical to the rhythms and the seasons of existence), into the proper language of each man, from everyday language to the most specialistic one, in which he expresses his fundamental relationship with things. For this reason it is the first perception (*Wahr-nehmung*) of Christ's truth. This perception of truth takes place in the conscious experience of the fundamental acceptance of one's alterity in Jesus Christ and, from here, in the experience of the 'fruits' that this new acceptance generates in the believer, like the ability of new relationships with the other, in that *permission* of new relationships that de Certeau indentified with faith's establishing breaking, that sets ourselves forever in a relationship with the other

---

[6]*Summa theologica,* 1 q. 1 a. 6; cf. *De divinis nominibus,* PG 3, 648B (Ref. [18]).

[19]. It is a conversion that is condition for the knowledge of the truth of Christ. Paraphrasing Paul in Fil. 3, 10: when man conforms himself to the death of Christ, i.e. he translates into his own experience the absolute Yea pronounced on the Cross, it is then that he 'knows the Christ'; he thus knows 'the energy of his resurrection and the communion with his sufferings.' Indeed, through the conversion of his own language, made homogeneous to the logic of the acceptance of the other as disclosed on the Cross, this language becomes itself accepting, and participates to the energy of the Christological event: it acquires the ability of 'universal' communication.

In the Christian conversion, however, languages 'conform' to the Christian narration, they translate the truth of the Christological event, according to several irreducible paths, just because it is not 'immediately' a conversion of signifiers and meanings in which every language is articulated, but of the linguistic essence of man who says himself, communicates himself in them. Thus, truth received in conversion does not determine by itself the truth of the meanings expressed in every language. To dare a well-defined example, then the translation of the language of faith in the language of the palaeontologist Teilhard de Chardin does not imply a change of the logic implant of the scientific language of Teilhard, but of what can be communicated in it (thus not through it, to employ Benjamin's distinction). This is equivalent to say that the translation of the language of faith in the language of science of Teilhard does not tell anything yet on the 'scientific' truth of the language itself, of the language as the language through which one says the things. However, history shows also other translations, more unobtrusive, with different modalities, and even those which express themselves as the rule of the silence about anything which cannot be spoken of, about anything which in fact cannot be told by words.

The Role of Theology

The 'creativity' of translation does not belong to the theologian as such. Theology analyzes the conditions of possibility of the translation by analyzing the narrations present in the tradition of the Christian language, revealing its profound grammar by means of the reason, criticizing incoherences, opening new possibilities. However, it does not offer the criteria of 'truth' for the translation. Its criteria, those of the *ratio,* are essentially two: that of coherence with the history of narration itself (historical reason), and that of internal coherence of the assumed language (critical reason, be it 'philosophical' or 'scientific', etc.). The language of theology is therefore by its own nature 'contaminated' by other languages, including the language of science.

If the creativity of translation does not belong to the theologian as such, it is nonetheless true that the theologian is a believer by himself, he experienced and daily experiences the 'Yea' of the Father, he belongs to the confessing community and operates its translations. In the creativity of faith, the believer theologian thus assumes languages and *rationes* that reveal the fecundity of Christian truth even within the reflected rationality, and not only at the level of the immediate rationality or of the 'surface grammar'. The creativity of the great theologians, from Origen to

Augustine, from Thomas to Nicholas of Cusa, from Möhler to Newman or Barth, resides in this happy symbiosis between faith and *ratio*.

What does then theology add to the *ratio primae veritatis?* Both through the knowledge of the history of the tradition of the believers, and through the assimilation of the critical instruments accumulated by contemporary knowledge, and through his personal genius and the intensity of the his personal conversion experience, the theologian adds to the *ratio primae veritatis* accepted and lived by the common believer (and firstly by himself in person) the knowledge of the grammar subtended by faith itself and its enrichment, its exaltation, above all when he becomes able of a new reading of beauty, but also of the contradictions and troubles of human existence. In this context, suffice it to call forth a few names, in order to let these assertions become clear: from Augustine to Thomas, from Eckhart to Nicholas of Cusa, from Luther to Kirkegaard, from Rosmini to Newman, from Barth to Bonhoeffer, from Przywara to von Balthasar, etc. but also those humble workers who have contributed so much to the critical understanding of the Christian tradition. In the latter case, rather than individual names, we should rather speak of the various choirs: that of the critical exegesis, that of the recovery of the Fathers, that of the recovery of the Middle Ages, from the renovated knowledge of the mediaeval schisms and of the Reformation, up to ourselves.

References

1. L. Wittgenstein, *Lectures & conversations on aesthetics, psychology, and religious belief* (University of California Press, Berkeley, 1966)
2. P. Overhage, K. Rahner, *Das Problem der Hominisation: über den biologischen Ursprung des Menschen* (Herder, Freiburg im Breisgau, 1961)
3. S. Wiedenhofer, Theologische Revue **83**, 353 (1987)
4. A. Funkenstein, *Theology and the scientific imagination from the Middle Ages to the 17th century* (Princeton University Press, Princeton, 1989)
5. C. Rovelli, *La realtà non è come ci appare: la struttura elementare delle cose* (Raffaello Cortina, Milano, 2014)
6. K.R. Popper, *Logik der Forschung: zur Erkenntnistheorie der modernen Naturwissenschaft* (Springer, Wien, 1935)
7. G. Ebeling, *Kritischer Rationalismus? Zu Hans Alberts Traktat über kritische Vernunft* (Mohr, Tübingen, 1973)
8. H. Albert, *Die Wissenschaft und die Fehlbarkeit der Vernunft* (Mohr, Tübingen, 1982)
9. I.T. Ramsey, *Religious language: an empirical placing of theological phrases* (SCM Press, London, 1957)
10. B. Pascal, *Thoughts (Pensées)* (Penguin, Harmondsworth, 1966)
11. P. Ricoeur, Royal Institute of Philosophy Lectures **2** (Talk of God), 100 (1969)
12. G. Ebeling, in *Religion in Geschichte und Gegenwart*, vol. 6, ed. by K. Galling, 3rd edn. (Mohr, Tübingen, 1962), p. 760
13. G. Ruggieri, in *La sapienza del cuore. Omaggio a Enzo Bianchi* (Einaudi, Torino, 2013), pp. 406–427
14. G. Ruggieri, in *Una teologia in comunità*, ed. by D. Paoletti (Messaggero, Padova, 2015), pp. 41–59
15. W. Benjamin, in *Gesammelte Schriften, vol. II.1*, vol. 2.1 (Suhrkamp, Frankfurt, 1977), pp. 140–157

16. W. Benjamin, in *Gesammelte Schriften, vol. IV.1*, vol. 4.1 (Suhrkamp, Frankfurt, 1981), pp. 9–21
17. G. Ruggieri, *Della fede: la certezza, il dubbio, la lotta* (Carocci, Roma, 2014)
18. J.P. Migne, *Patrologiae cursus completus. Series graeca (1857–1936)* (Garnier, Paris, 1957)
19. M. de Certeau, *La faiblesse de croire* (Seuil, Paris, 1987)

Chapter 25
Symmetries and Physics

R. Pucci

Opening Address

MAGNIFICENT RECTOR, AUTHORITIES, COLLEAGUES, STUDENTS, TECHNICAL AND ADMINISTRATIVE PERSONNEL, LADIES AND GENTLEMEN - - -

I feel honoured and science symmetry privileged to hold the opening lecture of the academic year 2002–2003. The subject of my lecture, symmetries and physics, is a "classic" theme in physics and is, at the same time, at the basis of frontier research. Physics is, I believe, quite distant from the experience of the majority of the audience. Therefore, I will make my best efforts to employ not a too technical language. I will also make use of both examples and historical information.

Introduction

By symmetry, physicists mean that different "points of view" exist, from which a system appears to be the same. Take, for instance, a sphere and a cylinder (Fig. 25.1).

R. Pucci (✉)
Dipartimento di Fisica e Astronomia, Università di Catania,
Via S. Sofia, 64, 95123 Catania, Italy
e-mail: renato.pucci@ct.infn.it

R. Pucci
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci
IMM-CNR, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

345

Fig. 25.1 Different points of
view from which a sphere
and a cylinder can be looked
at (color figure online)
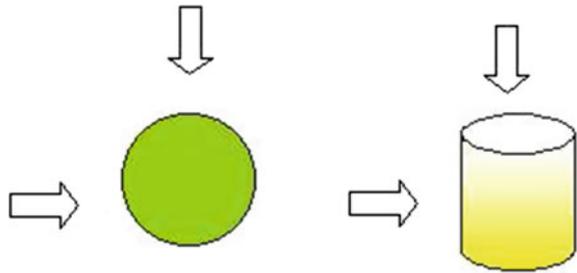
The projection of a sphere onto a plane will be seen as a circle by any observer, regardless of the direction she looks at it.

As far as a cylinder is concerned, observers looking at it perpendicularly to its axis will see a rectangle, while those looking at it according to a direction parallel to the axis will see a circle. This implies that the cylinder is endowed by less symmetry than the sphere, and that an observer looking at the cylinder from a single "point of view" cannot realize all the symmetries of the system. There may be hidden symmetries! And that's exactly one of physicists' most exciting tasks [1]: uncovering the hidden symmetries of nature. It is in this sense that physics may be considered one of men's creative activities, which drives them towards an intellectual adventure unique in its genre.

> «It is by making use of all the tools of our logical, mathematical, and technical apparatus, that we try to demonstrate that our hypotheses were false. Instead of the latter, we surmise new unjustified and unjustifiable hypotheses, new rough and premature prejudices.» [2]

What is important is that symmetries are closely connected with the laws of nature. Indeed, the discovery of the laws of physics is based on the research of the symmetries of space, time, and matter.

As I have already anticipated, I shall often make reference to the history of science. I shall start from Aristotle's physics, and I shall then emphasize how Galilean relativity paved the way towards contemporary physics. I shall indicate how the energy and momentum conservation laws may be derived from the symmetry of space and time, and how from such conservation laws the laws of motion may be derived. Guided by the history of science, I shall show how symmetries played a fundamental role for two important scientific revolutions: Einstein's relativity and quantum mechanics. The importance in physics of simple "models" will be shown by the introduction of the homogeneous electron gas due to Thomas and Fermi and by the relevance of this theory in the study of atoms, molecules and solids. I shall also introduce the theory of "symmetry breaking", and I shall show how the latter gives rise to the various laws that govern nature. I shall eventually bring two instances of what I have said, namely the solids under very high pressure, and the high critical temperature superconductors.

Galilean Relativity

What is taken for granted nowadays, was not always so. Indeed, for many centuries people had a completely different conception of the physical laws. Aristotle's theory of motion [3] was based upon the existence of natural places, i.e. privileged places in the Universe. The natural place of an heavy body was, for example, the centre of the Earth, while that of celestial bodies was above the Moon's orbit.

This amounts to say that the laws of physics vary depending on whether one performs an experiment at home, or at the centre of the Earth, or on the Moon. According to Aristotle, a stone falls towards the ground if I am at home, it stays still at the centre of the Earth, while it cannot reach the Moon anyhow.

The scientific revolution [4] which will lead towards classic physics, especially through the contributions of Galileo and Newton, begins with the work of Copernicus.

The approximations introduced by Copernicus' theory, with respect to Ptolemy's theory, were very important to Copernicus, since they reflected that peculiar mathematical harmony that, according to him, Universe "must possess". It was not conceivable that God created such a complicated Universe, as the one described by Ptolemy's theory. Copernicus' theory of the Universe is endowed with an extraordinary amount of symmetry: the Universe could not have had any other form, than that described by him.

A new theory of motion was however required, in order for Copernicus' mathematical model to become a new cosmology. Indeed, if the Copernican theory was not a mere mathematical invention, it should explain why towns were not uprooted from the Earth because of the motion thereof around the Sun.

Galileo afforded a new theory of motion via the principle of inertia. He made use of the approach known nowadays as "Galilean relativity".

The principle of inertia states that a body stays at rest or keeps on moving uniformly along a straight line, provided there are not external agents acting upon it. Within his "experimental method", he faces the problem of relativity in the following way:

> « Close you up with some friends in the largest room below the deck of some large vessel. [*A detailed description of several experiments follow...*] After you carefully observed all these things, although there is no doubt that, while the vessel is at rest they should not take place in a different way, let the vessel move with an arbitrary speed. Then, provided the motion be uniform and not fluctuating at all, you will not recognize any change whatsoever in all the aforementioned effects, nor you will be able, from any of them, to understand whether the vessel moves or is rather still.» [5]

In other words, there is no way, by means of any experiment of classical mechanics, to conclude whether a body is at rest or is moving with uniform rectilinear motion. Every observer, moving with uniform rectilinear motion one with respect to the other, will see and describe the motion of a given body *in the same way*.

Space is homogeneous (i.e. it is the same everywhere) and isotropic (i.e. it is the same in all directions); time is uniform (i.e. every clock always beats at the same rhythm). The homogeneity of space implies that the laws of physics must be independent of the place where they are measured, and vice versa. The uniformity

of time implies that the laws of nature are the same yesterday, today and tomorrow, i.e. they are the same regardless of when they are measured.

Galileo's theory was completed by Newton's axiomatic work. Newton's theory of gravitation is one of the highest achievements of human intellect: it can be applied to two arbitrary bodies both on the Earth and in the Universe. Together with the laws of motion, the theory of gravitation allows to predict the trajectory of a projectile, as well as that of the Earth orbiting around the Sun.

Galileo and Newton were the first who attempted at deriving some consequences from the fact that the laws of nature must be independent from space and time. It was demonstrated that from space homogeneity it is possible to derive the law of momentum conservation and that, from the uniformity of time, one may derive the law of energy conservation (i.e. energy does not vary with time). Conversely, from these two conservation laws, all Newton's laws of motion can be derived. In other words, the symmetries of space and time completely determine the laws of classical mechanics.

Einstein's Relativity

Newton was quite concerned about how the force of gravity could be transmitted across the apparently "vacuum" space. Therefore, he made the hypothesis that space was actually filled with a material medium, *viz.* the ether. Moreover, he introduced the concepts of absolute space and time that, as Kant stated later, are the a priori forms which guarantee the order, the understanding, and the very condition of our perception of the inertial reference systems.

Einstein upset in a fundamental way the concepts of absolute space and time, as well as the existence of the ether.

All of Einstein's three fundamental works of 1905 begin by the statement of formal asymmetries or other inconsistencies, of basically aesthetic nature, in the previous theories [6]. In order to eliminate such asymmetries, he then suggests a general principle and some experimental proofs thereof.

It is not without meaning that Einstein's paper *"On the electrodynamics of moving bodies"* [7], where he actually introduces special relativity, begins as:

> « It is known that Maxwell's electrodynamics —as usually understood at the present time— conceived at present— when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena.» [7]

As Galileo stated that there was no way to establish whether a ship is still or is moving uniformly along a straight line, by means of experiments of classical mechanics, in the same way Einstein states that this is impossible also if one performs electromagnetic experiments.

The consequences of this quest for new symmetries are amazing: not only the concepts of absolute space and time were abandoned, but a new structure of space

and time was discovered. Space and time cannot be put apart from each other, but they build up a single entity, called spacetime.

Einstein demonstrated that, if one wants to preserve both the laws of classical mechanics and those of classical electromagnetism, every observer must measure the *same* velocity of light. The only way by which this can be obtained is that space and time are measured differently by different observers. An observer moving with respect to another, will see the other's rule contract, and the other's clock accelerate.

Quantum Mechanics

Within classical mechanics there were two well-separated worlds: the world of waves, and that of particles. For instance, the phenomenon of interference took place in the former, while it did not in the latter. It was Einstein himself who, in his first work of 1905 [8], the one by which he was awarded the Nobel prize, demonstrated that the photoelectric effect can be explained only if one admits that electromagnetic waves behave like particles: the photons.

At the beginning of this work, Einstein writes:

« A profound formal difference exists between the theoretical concepts that physicists have formed about gases and other ponderable bodies, and Maxwell's theory of electromagnetic processes in the so-called empty space.» [8]

The discrepancy that Einstein puts forward, and that on the contrary seemed "natural" and acceptable by all the other physicists, is between the discontinuous or discrete character of particles and their energy, and the continuity of the functions relative to the electromagnetic waves and their energy per unit surface.

Einstein then deals with the electromagnetic waves as with an ensemble of particles, the photons, having discrete amounts of energy. Moreover, only an integer number of photons can be emitted and absorbed, as if they were indivisible corpuscles.

On the other side, de Broglie [9] associated a wave to the elementary particles. One can then deduce that particles too can develop an interference pattern, as was widely confirmed experimentally.

Quantum mechanics thus originated from the elimination of the asymmetry between particles and waves: every object at a microscopic level is a wave-particle or a particle-wave, or whatever we would like to call them [10]. As Heisenberg demonstrated [11], this is possible only if one gives up with determinism: it is not possible to associate to a particle at a given instant of time a position and a velocity, with arbitrary precision.

Upset by such a statement, Einstein himself reportedly said that it was as if *we were feeling lost*. According to Galileo's lesson, one can make physics only neglecting the irrelevant aspects of a phenomenon. What we have said above shows, however, that some aspects that are irrelevant at a *scale* of the order of a centimetre, say, are not irrelevant at a scale of the order of 10^{-8} cm.

Usually, different symmetries, and thus different physical laws, are associated with different scales. That's why it is important to specify the limits of validity of our theories.

The Fermi Gas

Besides the symmetries of spacetime and of particles-waves, also the symmetries of matter play a fundamental role in physics. In the following, I shall introduce this topic by speaking of the *Fermi gas,* which is one of the most important "models" of condensed matter physics.

Physical systems are usually quite complex, so that it is indispensable to introduce approximations, or models. Some models can work very well, therefore one tries to exploit them up to their limits of applicability.

In the 1920s, Thomas [12] and, independently, Fermi [13] introduced a model whereby the electrons in an atom were treated as an homogeneous gas of electrons, i.e. a gas characterized by a constant density (the density is the number of electrons per unit volume). As a matter of fact, the electron density in an atom is *not* constant at all. Yet Fermi and Thomas applied the model "locally", by assuming that in every point there was a neighbourhood where many electrons had a constant density. Such a model, corroborated by Fermi's statistics [14] and by Heinseberg's indetermination principle [11], affords a different method, with respect to Schrödinger's method, to solve quantum mechanical problems. Energy is not expressed in terms of a wavefunction, but rather in terms of the electron density, which is an observable quantity of the system. One can then avoid some of the paradoxes of quantum mechanics. One can moreover derive [15] useful relationships among the properties of the atoms as the atomic number Z is varied. Such relations cannot be obtained with more sophisticated methods.

The Fermi gas model has been generalized to molecules [16], solids [17], and recently also to fullerenes [18]. Nowadays, most refined calculations of the energy levels of the solids are performed by means of density functional theory, which is a development of Fermi–Thomas theory. It is interesting to note how the Fermi–Thomas equation enables to derive the properties of all atoms, provided lengths are properly scaled for each atomic species.

Symmetry Breaking

Many a scientist agrees upon the reductionist hypothesis, consisting in reducing everything to its simplest and most fundamental roots. In a certain sense, this implies that everything can be traced back to the four fundamental forces of nature: the strong, weak, electromagnetic and gravitational force. Indeed, physicists hope that

everything can be reduced to one single force, and some important achievements have been obtained in this direction.

On the other hand, what cannot be agreed upon is the "constructionist" hypothesis, i.e. that every law of nature can be derived only from these four forces.

The presumption of some physicists in this respect is well synthesized by P.A.M. Dirac's contribution to a conference, where he said that physicists discovered the electromagnetic force, *"all the rest is chemistry"*.

A specific answer to Dirac's statement has been provided by another distinguished physicist, P.W. Anderson [19]. The physical laws of a system cannot be extrapolated to those of another system, if the two systems have different symmetries. For example, we cannot hold that the laws governing the atoms, having spherical symmetry, are the same as those of diatomic molecules, that are symmetric only along the axis between the two nuclei. Every time that a "symmetry is broken", the laws of nature do change!

> « At each stage entirely new laws, concepts, and generalizations are necessary, requiring inspiration and creativity to just as great a degree as in the previous one.» [19]

Anderson's work emphasizes the importance of the study of symmetry breaking.

Another kind of symmetry breaking can take place in systems endowed with a very large number of atoms or molecules, such as the crystals. In a crystal, e.g. all the molecules may be "oriented" along a given direction, thus giving rise to a privileged axis. Actually, all the molecules may be oriented in exactly the opposite direction, but the system is so large that no quantum or thermal fluctuation can make it oscillate from a configuration to the other: the molecules of the crystal always stay oriented in one direction only, and the overall symmetry is thus effectively broken.

In many cases, we may conclude that a large system, built up by many small elements, is not only a larger system, but is also qualitatively different by each of its parts: to quote Anderson again, *"Many is different"* [19]. Moreover, it is important to recognize that, if one considers a very large number of particles, as a function of parameters such as pressure, temperature, or volume, one may have "phase transitions", such as that between a liquid and a gas, in which both the macroscopic properties and the microscopic symmetries of the system change completely.

One studies the fundamental physical properties of a system by studying the different phases of a system, and the transitions between one another. A characteristic of phase transitions is that there are particular values of pressure and temperature, defining a *critical point,* where it is *not* possible to distinguish two or more coexisting phases. In this case, the system looks like the *same* at every *scale* whatsoever. In other words, whatever unit we adopt to measure lengths (decimetre, millimetre, or thousandths of millimetre), the system will appear always identical. This particular kind of symmetry is known as *scale invariance* [20] (Fig. 25.2).

Another characteristic of phase transitions is the existence of an *order parameter,* i.e. of some quantity that takes on different values in the different phases. Phase transitions having an order parameter with the same symmetries are absolutely identical to one another (universality). That is why it is of the utmost importance to study

Fig. 25.2 "Big fleas have little fleas/Upon their backs to bite'em;/Little fleas have smaller fleas,/And so ad infinitum". (Jonathan Swift, *On scaling laws in Biology*, 1733)
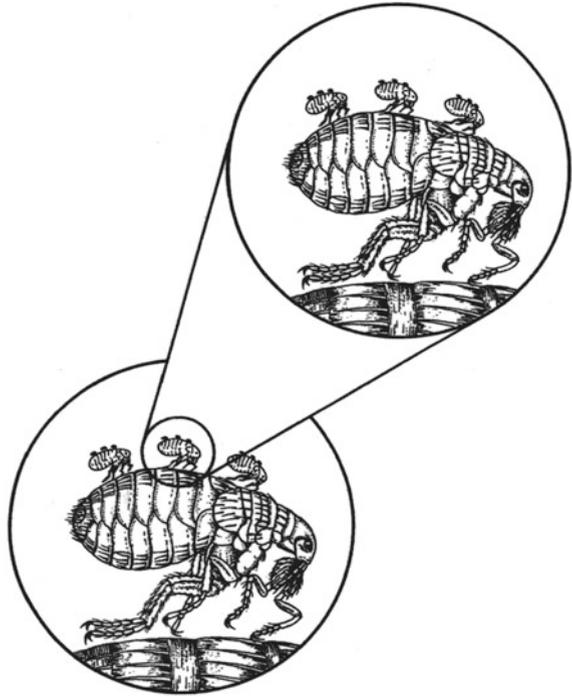
also the symmetries of the order parameter. Once more, symmetry determines the dynamics and the nature of the transition.

Solids Under High Pressure

Concerning materials, crystals provide the clearest example of the relationship between symmetries and the properties of a system. On the basis of the fact that a crystal has a given structure, which repeats itself rigidly and periodically in space, thus forming a lattice, one is able to derive the properties of the electrons moving within the crystal. From the particular symmetries of the crystalline lattice and of the atoms composing it, one may establish whether a given material is an insulator, or a metal, or a semiconductor.

The most "natural" phase of a solid should be the most symmetric one, where all the ions are found at equal distances from their nearest neighbours, immersed in a Fermi gas. Such a behaviour should take place, at least, at very high pressures (of the order of those encountered e.g. in the white dwarfs), where all the chemical bonds are broken.

"Fortunately", there are elements that behave according to the model described above also at atmospheric pressure. These elements are the alkaline metals (Li, Na, K, ...), which possess only one electron in the outermost orbital. These electrons delocalize within the crystal, thus forming the Fermi gas, having metallic properties.

In 1935, Wigner and Huntington [21] asked themselves *"why hydrogen, having a single electron, forms an insulating molecular solid at low pressure?"* These workers proposed that solid hydrogen at high pressure should undergo a molecular to monoatomic transition and, therefore, an insulator to metal transition. Since then, quite a number of papers, both theoretical and experimental, have attempted at detecting such a new phase of hydrogen, to no avail.

The quest for broken symmetries prompted us [22, 23] quite the opposite question, namely *why, at low pressure, the alkalis do not exhibit a broken symmetry, like hydrogen?*

In 1997, my collaborators and I published a paper with the provocative title *"Are the light alkali metals still metals at high pressure?"* [22]. In that paper, we suggested that under high pressure (not so high as those found in the white dwarfs), the alkalis could be unstable towards less symmetric crystalline structures. This means that the ions can pair themselves, thus forming a less symmetric phase. One then has a broken symmetry phase, of the kind just described (Fig. 25.3).

Recently, some experimental confirmations of our hypothesis have been published [24].

Fig. 25.3 Lithium under high pressure ($P \simeq 44$ GPa) in the cI16 structure. The background shows the electronic charge density distribution evaluated at 48.8 GPa. (Redrawn after Ref. [24]) (color figure online)
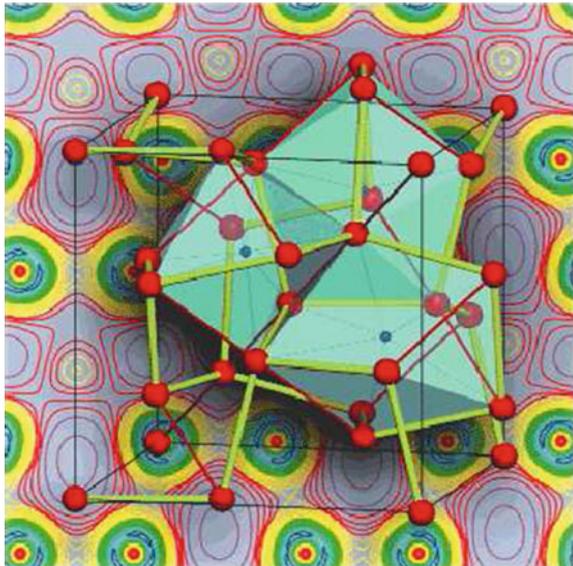
High Critical Temperature Superconductors

Superconductivity [25–27], namely the uninterrupted flow of electrical current below some critical temperature T_c, is a phase transition [28, 29]. Indeed, it is probably the transition in which a spontaneous symmetry breaking manifests itself in the most uncontroversial way.

While in the normal state one may describe the electrons by means of the Fermi gas model, in the superconducting state electron pairs are formed, which *condense* in a collective state [30]. If a material is in such a state, photons cannot enter its bulk (the Meissner–Ochsenfeld effect [31]). Therefore, they acquire a *mass* (it is as if photons would slow down and become heavy [32]).

The condensation of electron pairs in a collective state implies the formation of a *gap* between the energy of such a state and the first available state. The energy gap is the order parameter of the superconducting transition. It vanishes in the normal state, and is different from zero in the superconducting state.

Up to 1987, the maximum value reached for T_c was about 30 K. In that date, however, Bednorz and Müller [33] discovered a new family of compounds with much larger critical temperatures. Nowadays, critical temperatures of the order of 140 K have been reached.

The high-T_c superconductors are strongly anisotropic materials. They are characterized by different properties in the direction parallel to the layers containing the atoms, than in the direction perpendicular to them. One has a large conductance in the direction parallel to the layers, e.g. than in the direction perpendicular to them. That's why the high-T_c superconductors are considered quasi-bidimensional systems (Fig. 25.4).
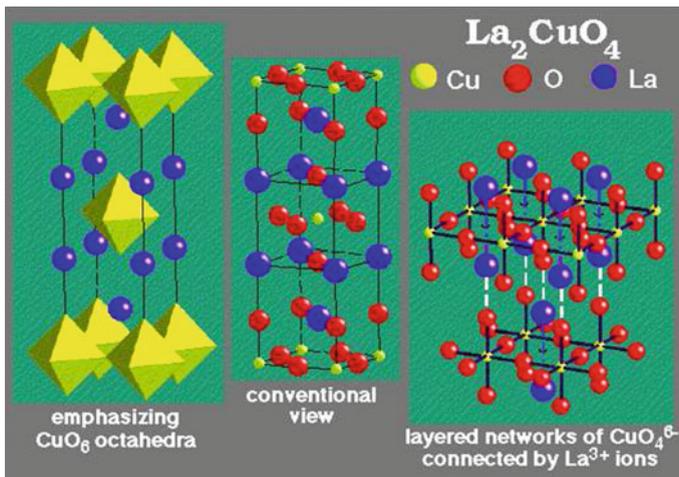
Fig. 25.4 Alternative representations of the structure of a typical cuprate high-T_c superconductor: La_2CuO_4 (color figure online)
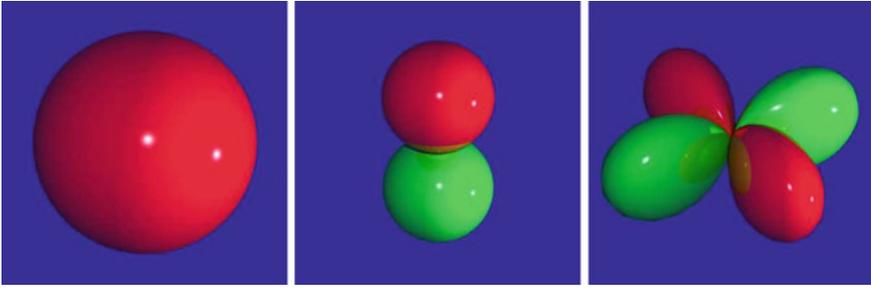
Fig. 25.5 Possible symmetries of the superconducting order parameter: *s*, *p*, *d* (*left* to *right*) (color figure online)

The order parameter, *viz.* the energy gap, is not characterized by the spherical symmetry (*s*-wave), which is typical of the "conventional" superconductors, but is rather characterized by a *d*-wave symmetry [34, 35] (Fig. 25.5).

A satisfactory theory of high-T_c superconductivity is not yet available, to date, but many a scientist is convinced that the symmetry of the lattice (i.e. its quasi-bidimensional character) and the symmetry of the order parameter must play a prominent role in the quest for the solution of such a problem.

Conclusions and Perspectives

In this opening lecture, I tried to show that it is not a paradox to state that symmetries *are* the physics, and that the quest for hidden symmetries has been, and probably will be, the main guideline to uncover the fundamental laws of nature.

Einstein, e.g. demonstrated how the symmetries of space and time, found by Galileo and Newton, could be unified in more general symmetries of the spacetime.

As far as the symmetries of matter are concerned, Wilson [36] demonstrated that there exists an amazing parallel between the physics of superconductors and that of the elementary particles. Probably, it is fair to say that elementary particles have a mass for the same reason why photons cannot enter a superconductor.

At a certain scale one may have broken symmetries, which can manifest itself at a different scale. Clear instances of such an aspect are found both in condensed matter theory, and in the electroweak theory (a theory which unifies electromagnetism and the physics of weak interactions).

Symmetries determine the physical quantities required to describe the phenomena, and the dynamical equations governing them. To reveal the symmetries of a system thus amounts to discover its fundamental properties.

One of the most important symmetries found in nature is scale invariance. This is found at the critical point of a phase transition, and consists in the system being perfectly identical to itself, when we change the length units.

The study of phase transitions, the idea of scale invariance and the renormalization method have been important precursors of the theory of chaos [37]. Also in the latter case, by means of computer "experiments", unexpected symmetries were found, such as bifurcations, strange attractors and fractals [38]. In phenomena where nonlinear effects are relevant [39], these ideas have often found important applications. In this case, I am not only referring to physics (including turbulence, nonlinear optics, etc.), but also to several other fields of science, such as biology, meteorology and geology.

Some researchers are convinced that the origin of the Universe may be described by a critical point, and that therefore there was a "scale invariance" at the beginning. It is perhaps this symmetry, at present hidden to our eyes, that we should unveil to fully understand the Universe, in its simplicity and beauty.

References

1. L.M. Krauss, *Fear of Physics: A Guide for the Perplexed* (Basic Books, New York, 1993)
2. K.R. Popper, *Logik der Forschung: zur Erkenntnistheorie der modernen Naturwissenschaft* (Springer, Wien, 1935)
3. For a deeper description and more references, see R. Pucci, in *Inizio e Futuro del Cosmo: Linguaggi a confronto,* edited by R. Pucci and G. Ruggieri (S. Paolo, Alba (TO), 1999).
4. T.S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago, 1962)
5. G. Galilei, *Discorsi e Dimostrazioni Matematiche intorno a due Nuove Scienze* (Edizione Nazionale diretta da A. Favaro e G. Barbera, Firenze, 1890–1909)
6. G. Holton, in *Int. School of Physics "E. Fermi" — LVII Course*, ed. by C. Weiner (North Holland, Amsterdam, 1972)
7. A. Einstein, Ann. Physik **17**, 891 (1905)
8. A. Einstein, Ann. Physik **17**, 132 (1905)
9. L. de Broglie, *The Revolution in Physics* (Grossett and Dunlop, New York, 1955)
10. R. Feynman, R. Leighton, M. Sands, *La Fisica di Feynman* (Masson, Milano, 1992)
11. W. Heisenberg, Z. Physik **43**, 172 (1927)
12. L.H. Thomas, Math. Proc. Cambridge Phil. Soc. **23**, 542 (1926). DOI 10.1017/S0305004100011683
13. E. Fermi, Rendiconti dell'Accademia Nazionale dei Lincei **6**, 602 (1927)
14. E. Fermi, Z. Physik **36**, 902 (1926)
15. R. Pucci, N.H. March, Phys. Rev. A **33**, 3511 (1986). DOI 10.1103/PhysRevA.33.3511
16. N.H. March, Proc. Cambridge Philos. Soc. **48**, 665 (1952). DOI 10.1017/S0305004100076441. Reprinted in Ref. [40]
17. J. Callaway, N.H. March, in *Solid State Physics*, *Solid State Physics*, vol. 38, ed. by H. Ehrenreich, D. Turnbull, F. Seitz (Academic Press, 1984), pp. 135–221. DOI 10.1016/S0081-1947(08)60313-6
18. F. Siringo, G. Piccitto, R. Pucci, Phys. Rev. A **46**, 4048 (1992)
19. P.W. Anderson, Science **177**(4047), 393 (1972). DOI 10.1126/science.177.4047.393
20. K. Huang, *Statistical mechanics* (Wiley, New York, 1963)
21. E. Wigner, H.B. Huntington, J. Chem. Phys. **3**(12), 764 (1935). DOI 10.1063/1.1749590
22. F. Siringo, R. Pucci, G.G.N. Angilella, High Press. Research **15**(4), 255 (1997). DOI 10.1080/08957959708244246. arXiv:cond-mat/9512011

23. G.G.N. Angilella, F. Siringo, R. Pucci, Eur. Phys. J. B **32**, 323 (2003). DOI 10.1140/epjb/e2003-00105-8
24. M. Hanfland, K. Syassen, N.E. Christensen, D.L. Novikov, Nature (London) **408**, 174 (2000). DOI 10.1038/35041515
25. H. Kamerlingh Onnes, Comm. Phys. Lab. Univ. Leiden **120b** (1911)
26. H. Kamerlingh Onnes, Comm. Phys. Lab. Univ. Leiden **122b** (1911)
27. H. Kamerlingh Onnes, Comm. Phys. Lab. Univ. Leiden **124c** (1911)
28. F. London, *Superfluids* (Wiley, New York, 1950)
29. V.L. Ginzburg, L.D. Landau, Zh. Eksp. Teor. Fiz. (Sov. Phys. JETP) **20**, 1064 (1950)
30. J. Bardeen, L.N. Cooper, J.R. Schrieffer, Phys. Rev. **108**, 1175 (1957)
31. W. Meissner, R. Ochsenfeld, Naturwiss. **21**(44), 787 (1933). DOI 10.1007/BF01504252
32. M. Baldo, G. Giansiracusa, U. Lombardo, R. Pucci, Phys. Lett. A **62**(7), 509 (1977). DOI 10.1016/0375-9601(77)90084-6
33. J.G. Bednorz, K.A. Müller, Z. Phys. B **64**, 189 (1986). DOI 10.1007/BF01303701
34. G.G.N. Angilella, R. Pucci, F. Siringo, A. Sudbø, Phys. Rev. B **59**, 1339 (1999). DOI 10.1103/PhysRevB.59.1339
35. G.G.N. Angilella, N.H. March, R. Pucci, Phys. Rev. B **65**(9), 092509 (2002). DOI 10.1103/PhysRevB.65.092509
36. K.G. Wilson, in *Nobel Lectures, Physics 1981–1990*, ed. by T. Frängsmyr, G. Ekspong (World Scientific, Singapore, 1993)
37. J. Gleick, *Chaos. Making a new science.* (Penguin, New York, 1988)
38. B.B. Mandelbrot, *Les objets fractals: forme, hasard et dimension* (Flammarion, Paris, 1975)
39. A. La Magna, R. Pucci, Phys. Rev. B **55**, 14886 (1997). DOI 10.1103/PhysRevB.55.14886
40. N.H. March, G.G.N. Angilella (eds.), *Many-body Theory of Molecules, Clusters, and Condensed Phases* (World Scientific, Singapore, 2009)

Chapter 26
Science and Religion: A Difficult Relationship

R. Pucci and G.G.N. Angilella

Introduction

The relationship between science and religion is not a well-posed question, at least over the last decades. Indeed, it was long ago when scientists such as Boyle or Newton endeavoured to justify in which way their faith could be reconciled with their own scientific discoveries. Nowadays, most scientists seem to share the conviction that region has been and still is a hindrance towards scientific progress. On the other hand, most theologians seem to consider scientists as bold and incapable to understand that, beyond their experiments and calculations, there exists a deeper word on the human being.

Recently, the main attitude of scientists has been reinvigorated by Rovelli [1]. Rovelli emphasizes that science could only be born when men (he often makes reference to the ancient Greek philosopher Anaximander) could get free of their own

R. Pucci (✉) · G.G.N. Angilella
Dipartimento di Fisica e Astronomia, Università di Catania, Via S. Sofia, 64, 95123 Catania, Italy
e-mail: renato.pucci@ct.infn.it

G.G.N. Angilella
e-mail: giuseppe.angilella@ct.infn.it

R. Pucci · G.G.N. Angilella
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci · G.G.N. Angilella
IMM-CNR, UdR Catania, Via S. Sofia 64, 95123 Catania, Italy

G.G.N. Angilella
INFN, Sez. Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
Scuola Superiore di Catania, Università di Catania, Via Valdisavoia, 9, 95123 Catania, Italy

mythical-religious conceptions. Indeed, it was with Thales and Anaximander when natural phenomena found an explanation without making recourse to any deity. It is possible to explain the universe through a rational investigation: the supernatural, as introduced by mythology, disappears. The idea of a mutation cycle makes its appearance: a cycle which, from air, earth, and water, through plants and animals, gets back again to air, earth, and water [2].

Rovelli [1] underlines several elements that made this revolution possible, among which: (1) Anaximander learnt a lot from his mentor Thales, but was able to radically change some aspects of the latter's teachings; (2) within a Greek *polis*, not subject to any emperor, citizens were more likely to devise their own laws, rather than have them imposed from above; (3) an increase in the exchange of goods and the mobility of people induced an encounter among different civilizations, such as the Greek and Egyptian ones; this naturally favoured the birth of the critical sense.

Even though these ideas are fully reasonable, in what follows we shall try to emphasize some tracts of the religious thinking that, in our opinion, did contribute to the development of science. Finally, we will surmise that theology evolved as a consequence of scientific thinking. Due to our scarce knowledge of other religions, we will principally make reference to the Judeo-Christian religious tradition.

Popper and Kuhn

One of the most relevant contribution to modern epistemology was provided by Popper through his theory of falsificationism [3]. This approach allowed to draw a boundary between what is scientific and what is not. What is scientific, is what can be falsified, not what can be verified. For instance, the statement 'God exists' is not a scientific proposition as it cannot be falsified. At variance with the principle of verifiability, this approach endows with dignity and autonomy all human activities other than science, such as art, economy, psychology, politics, and religion. In addition, Popper insists that the true scientist must seek all that can confute his own theory, and not proofs in favour of the latter.

Kuhn's reply [4] could not be more devastating: never in the history of science did the revolutions evolve as Popper would have wished. On the contrary, in scientific revolutions conservatives and innovators struggled between each other making recourse to all the weapons at their disposal, even the less dignified. Kuhn, moreover, connects scientific revolutions to profound changes in economy, society and politics. Thus, science does not appear as an isolated castle, but rather as a palace embedded in a great town. Even though each of these two important authors tried to polish his own thought so as to take into account the other's criticism, the discrepancy between the two appears insanable.

Nonetheless, we believe that one should recognize Popper's merit of having defined the peculiarity of scientific thinking without denigrating other forms of thinking, and Kuhn's merit of having underlined that different forms of thinking are often not isolated, but rather influence one another. We will not be surprised, therefore,

if in the following, even admitting that science and religion are different languages [5], we shall look for possible reciprocal influences, which have been recognized historically.

Job, and the Scientist as a Rebel

In his book *The scientist as a rebel,* Dyson writes [6]

> There is no such thing as a unique scientific vision, any more than there is a unique poetic vision. Science is a mosaic of partial and conflicting visions. But there is one common element in these visions. The common element is rebellion against the restrictions imposed by the locally prevailing culture, Western or Eastern as the case may be. The vision of science is not specifically Western. …And what is true of science is also true of poetry.

Rovelli [1] is even more precise, by saying that a scientist introduces relevant advances if he is capable to contradict his masters' teachings in any of their fundamental aspects. Rovelli illustrates the above statement by describing the aforementioned case with Anaximander as an example. At variance with his master Thales, Anaximander does not believe that the world is made of water, neither he believes that the Earth floats on water, nor that earthquakes are due to the fluctuations of the Earth's disk in the medium where it floats. To Anaximander, [1]

> the Earth is not lying on something else (columns, turtles, an ocean, earth down forever), but rather it floats free in space. The sky is not just above our heads: it is all around us, including under our feet.

There is not even other water underneath [5]. Rovelli [1] recognizes in Anaximander the promotor of a 'third way'. He leans on some of Thales' achievements, but criticizes some essential traits. Similarly is how Copernicus behaves with respect to Ptolemy. In his *De revolutionibus*, Copernicus keeps the epicycles and deferents from Ptolemy's Almagest, but sets the Sun at the centre of the Universe, rather than the Earth. Such a 'third way' is halfway between unrestricted criticism, as is e.g., the case in the Bible for the Babylonian religion, and the acritical acceptation of his master, "as it was for Saul of Tarsus with respect to Jesus" [1]. In our opinion, the example by Rovelli [1] is however not quite correct. Saul (Paul the Apostle) did not confront himself with Jesus. He states that he does not know him by the flesh, and that such a way of knowledge is completely useless (2 COR 5, 16). Saul only confronted himself with the apostolic announcement, that is with the Crucifix, dead and resurrected, as a way of salvation. We do not know anything else.

It is appropriate to note, at this point, that such a 'third way' was also that of Job's with respect to God, who was much more than a teacher for him. Job does not abjure God, nor does he maledicts him, but asks him to account why He, the Just, lets an innocent suffer. Job legitimately claims that God participates in a direct confrontation with him. According to the Jewish legislation, this is actually a *ryb*, which, as is explained by Zagrebelsky [7], used to be a debate between two subjects, where each aimed not to destroy his own adversary, but rather to make him repent,

thus recuperating the relationship between the two. Another *ryb* is, for example, that of Christ with respect to his own people, who unjustly condemnates him. In his silence, he emphasizes the enormous injustice that is being perpetrated by his people, whereto he has offered himself as a 'silent sheep'. This is probably one of the largest problems of modern democracies: is it possible to establish a right that avoids the sufferings of the innocent?

The poetic corpus of Job's book is likely to date back to the seventh–sixth century BC, but one can certainly finds precursors to Job already in 2000 BC [8, 9]. But the genius of the author of the book of the Old Testament consists "in changing the ancient legend of patient and submissive Job into the tragedy of Job the rebel, who dismantles the tradition and struggles against God" [8]. The parallelism between Anaximander's and Job's attitudes has been here mentioned to remind that what Rovelli calls the 'third way' [1] was not alien to Jewish culture and probably also to other cultures [8, 9]. One may recognize the spirit of rebellion of Anaximander of Miletus, now in Turkey, with respect to his master also, and more vigorously, in the relationship between Job and his God.

Hints of Influence of the Religious Thinking in the Development of Scientific Thought

It is more difficult to find examples of *ryb* in the debates between two scientists, or two currents of scientific thought. That's because these debates are more similar to usual trials in a tribunal: the two opponents are judged by a third subject, which in this case is the rest of the scientific community, or even better the experimental data.

One of the most notable, and probably unique, example of *ryb* in the scientific context can be found in the epistolary between Bohr and Einstein [10]: that is a *ryb* which did not result in a positive outcome. But if the influence of the Jewish literature on authors such as Anaximander, as described above, may seem uncertain, we can assert with relative certainty that Christian theology did have a relevant role on the birth of modern Western science.

If by theology we mean any proposition on God [11], including those based on stories, myths, believes and rhytes, every religion obviously possesses its own theology. But we are here referring to the critical and rational theology, that one which had one of its greatest exponents e.g., in St Thomas Aquinas. Such a theology is a characteristic only of Christianism. Again Dyson writes [6]

> It is probably not an accident that modern science grew explosively in Christian Europe and left the rest of the world behind. A thousand years of theological disputes nurtured the habit of analytical thinking that could also be applied to the analysis of natural phenomena. On the other hand, the close historical relations between theology and science have caused conflicts between science and Christianity that does not exist between science and other religions.

Evidently, this close relationship between science and Christian theology was also source of conflicts between themselves, with no holds barred.

The last aspect that we would like to underline within the relationship outlined above is nothing but a mere suggestion. In support to Kuhn's thesis that scientific revolutions more easily take place where also economical, political or social revolutions are in progress, several authors [12, 13] asked themselves why quantum mechanics was mainly born in Germany. Ciccotti [12] and Baracca [13] emphasized that the Industrial revolution, albeit landed to Germany after the English counterpart, was in Germany shorter and deeper, thus influencing more decidedly teaching and scientific research in that country.

Such an analysis looks sufficiently convincing, but if one wants to underline that scientific revolutions arise in those places where criticism and rationalism is more diffused and better rooted, so as to react against the consolidated believes of a given society, then, in order to study the development of science, one should study that cultural climate of rebellion, which established itself in Germany at the beginning of the sixteenth century, because of Protestantism.

It is appropriate at this point to emphasize that Reformation did not at the beginning stimulate any scientific revolution. On the other hand, Protestants, and in particular Lutherans, condemned Copernicanism quite rapidly and severely. They considered it a false and dangerous system, as it was in contrast with the Bible. Marin Luther reportedly [14] said, in one of his *Tischreden* (dinner-table conversations), in open contrast against Copernicus,

> There is talk of a new astrologer who wants to prove that the earth moves and goes around instead of the sky, the Sun, the Moon, just as if somebody were moving in a carriage or ship might hold that he was sitting still and at rest while the Earth and the trees walked and moved. But that is how things are nowadays: when a man wishes to be clever he must needs invent something special, and the way he does it must needs be the best! The fool [Copernicus] wants to turn the whole art of astronomy upside-down. However, as Holy Scripture tells us, so did Joshua bid the sun to stand still and not the earth.

The considerations above suggest that the interaction between religious and scientific culture sometimes takes place directly, as was the case with Galileo, Darwin, etc., but often takes place in a fashion mediated by the dominant culture in a given society. Any scientific revolution should first finds its roots in culture, then in society, and it will be eventually accepted by the Church's establishment. That is why there have been delays even of centuries between the occurrence of a scientific revolution, and its acceptance by religion. The converse is probably true as well.

God and Symmetries

If God created the world, then the laws governing the latter must be simple and symmetrical. This conviction has been deeply shared by many men of science, who also were believers in God. Therefore, they sought manifestations of divine perfection in the study of Nature. As already said, Copernicus maintained almost all the ptolemaic model, with the exception of equants. However, he claimed that the simplifications

introduced in his theory, with the Sun, rather than the Earth, at the centre of the Universe, reflected of that harmony that the Universe 'must have'. It was not possible that God had created such a complicated Universe, such as Ptolemy's. Copernicus was convinced of the simplicity and harmony of the Universe, and these two attributes are valid elements in support of his model [15].

Even Descartes wrote [16]

> Besides, I have pointed out what are the laws of nature; and, with no other principle upon which to found my reasonings except the infinite perfection of God, I endeavored to demonstrate all those about which there could be any room for doubt, and to prove that they are such, that even if God had created more worlds, there could have been none in which these laws were not observed.

The fundamental fact is that symmetries are closely related with the laws of Nature. Space and time invariance imply the conservation laws of momentum and energy, respectively, and from these it is possible to derive the laws of motion [17]. It is remarkable that Einstein's paper on special relativity [18] begins with the words

> It is known that Maxwell's electrodynamics —as usually understood at the present time— when applied to moving bodies, leads to asymmetries which do not appear to be inherent in the phenomena.

Similarly, it is possible to assume that Quantum mechanics was originated by the attempt of removing the asymmetries which are present in classical physics between waves and particles [17]. Thus, the quest for symmetries has become even more importantly a method to study the laws of Nature, also for scientists who are non-believers. On the other hand, it has been observed that both the fundamental forces and some properties of matter appear as they do, because the breaking of a given symmetry has taken place. To reveal the symmetries of a system is equivalent to discover its fundamental properties, also for the physics of matter. The quest for hidden symmetries has become one of the most powerful means of investigation of Nature. One has succeeded to unify three out of four fundamental forces of Nature, and present efforts are aimed at adding to these three also the gravitational force. One endeavours to overcome the apparent contradictions between relativity and quantum mechanics within a more symmetrical scheme. The quest for hidden symmetries is also fundamental in the study of phase transitions and in several other fields. Coming back to the original point, one might say that Anaximander may have concluded that the Earth floats in the void and does not fall, from the observation that all the directions along which it might have fallen are equivalent.

Hints of Influence of Scientific Thinking in the Development of Religion

For nearly four centuries, the 'Galileo affair' has been a paradigm of the clash between science and the Church hierarchy, and is still nowadays the object of several controversies [19, 20]. Without getting involved with the numerous different issues of

the problem, we would just like wot emphasize, with Melloni [21], that the 'Galileo affair' has become for the fathers of the Second Vatican Council a 'spyglass', wherewith to scan

> (i) the abuse, as it gets manifested concretely in well ascertained facts; (ii) the theological basis over which that abuse had been founded, and the resistance to disclaim it.

One has to acknowledge, however, that this message, rather than from the official proceedings of the Council, emerges from the interpretation thereof by Pope John Paul II, during his address to the Pontifical Academy of Sciences, on November 30, 1979, on the occasion of the centennial of Albert Einstein's birth. On that occasion, he reminded [22] that Galileo Galilei

> had to suffer much [...] owing to men and bodies of the Church.

According to him, Galileo also suffered from one of those 'undue interventions' already condemned by the Second Vatican Council.

This fact, in our opinion, is suggestive not only that the Catholic Church has recognized most achievements of scientific research, but also of the profound changes occurred in theology. It would be absurd for theologians to claim that the Scripture should be interpreted literally. On the other hand, those who wish to interpret the Bible should find out what the author is likely to tell us, through the metaphors and the myths he adopts, well conscious that he is embedded in the culture of his times [11].

The exegetical research of the twentieth century operated a process of demythologizing of the Old Testament, to the aim of distinguishing between "the mythical vision of the world, which conditions the language of the Bible, and the message thereby conveyed" [11]. One may therefore surmise that the typical conviction of scientific research, that our knowledge is precarious and not definitive, has been transferred to the religious thinking, at least in part. Indeed, it is believed that the Scripture will become clearer through deeper study and religious attitude. In his *Homiliæ in Hiezechielem prophetam* (I, 7, 8), Pope Gregory the Great (twelfth century AD) used to say

> Scriptura crescit cum legente
> [Scriptures grow with him, who reads it].

In so saying, he meant that our knowledge of the Bible gets enriched by the thoughts of the men of faith that pursue it with heart sincere and free from prejudices.

Conclusions

In this brief paper, we addressed the question whether the issue of God be compatible with the new scientific knowledge, even though this question has been dealt with quite recently [23]. We have shared Popper's epistemologic attitude, according to which religious and scientific thinkings are sharply distinct from each other. As Kuhn

observes, however, it is possible to remark that such a distinction is not historically found. Dyson [6] notes that an interference between the two forms of thinking occurs especially when Christian religion comes into play. Scientists belonging to other religious confessions do not usually feel the urge to justify their being scientists and believers.

On the contrary, Christian scientists feel rather obliged to justify their being scientists and Christians at the same time. This is probably due to the fact that science and Christianism have almost always suffered from conflicts, at least in their public manifestations. Among the several examples that could be quoted, suffice here to mention the anecdote recently reported by Boncinelli [24]. In 1860, at the end of Huxley's conference on Darwin's evolutionism in Oxford, Anglican bishop Wilderforce asked him whether he descended from the apes from his mother's, or his father's side.

> Quietly Huxley replied that he did not consider a shame to have an ape as an ancestor, but rather to be kindred with a man who employed his own intelligence to hide the truth. [24]

It is appropriate here to mention that Lennox [25] remarks that this story, quite known in the account reported above, is not quite correct from the historical point of view, for several reasons, among which: (a) the clash was not between science and religion, but between two different scientific models; (b) the struggle did not end in a striking victory on behalf of Huxley, but rather quits. Nonetheless, the fact that scientists often report the contrast between Huxley and Wilderforce along the terms reported by Boncinelli [24] above, thereby ignoring the historical facts, denotes the fact that there is still some diffuse and sometimes justified fear among them, that the ecclesiastical hierarchy might hinder the development of scientific research. On the other hand, it is not difficult to find numerous theologians quite convinced that scientists, with their scientific method, are refractory to any other form of knowledge.

We believe that this situation is due to easily change in the near future, but we hope that the present contribution by two Christian scientists might be useful to calm down the controversy. After all, we have here surmised, even though some of our statements could be questioned, that there exist historical evidence that scientific and religious thinking, sometimes unconsciously, have reciprocally served to the development of each other.

References

1. C. Rovelli, *The first scientist. Anaximander and his legacy* (Westholme, Yardley, 2011)
2. W.C. Dampier, *A History of Science* (Cambridge University Press, Cambridge, 1929)
3. K.R. Popper, *Conjectures and refutations* (Routledge, London, 1963)
4. T.S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago, 1962)
5. R. Pucci, in *Inizio e futuro del cosmo: linguaggi a confronto*, ed. by R. Pucci, G. Ruggieri (San Paolo, Torino, 1999), p. 47
6. F.J. Dyson, *The Scientist as Rebel* (New York Review Books, New York, 2006)
7. G. Zagrebelsky, *Il* Crucifige! *e la democrazia* (Einaudi, Torino, 1995)
8. L. Alonso Schökel, J.L. Sicre Diaz, *Giobbe* (Borla, Roma, 1985)
9. G. Ravasi, *Giobbe* (Borla, Roma, 2005)

10. N. Bohr, *A. Einstein: Philosopher Scientist* (Cambridge University Press, Cambridge, 1949)
11. G. Ruggieri, *Prima lezione di teologia* (Laterza, Bari, 2011)
12. G. Ciccotti, E. Domini, in *L'Ape e l'Architetto*, ed. by G. Ciccotti, M. Cini, M. de Maria, G. Jona-Lasinio (Feltrinelli, Milano, 1976)
13. A. Baracca, S. Ruffo, A. Russo, *Scienza e industria: 1848–1915* (Laterza, Bari, 1979)
14. A. Koyré, *La rivoluzione astronomica* (Feltrinelli, Milano, 1966)
15. E.J. Dijksterhuis, *The mechanization of the world picture* (Clarendon Press, Oxford, 1961)
16. R. Descartes, *Discours de la méthode pour bien conduire sa raison, et chercher la vérité dans les sciences (Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth in the Sciences)* (1637)
17. R. Pucci, p. 347. (This book)
18. A. Einstein, Ann. Physik **17**, 891 (1905)
19. R. Chiaberge, Saturno, 10 febbraio 2012 (2012)
20. E. Brovedani, Il Sole 24 Ore, 12 febbraio 2012 (2012)
21. A. Melloni, Cristianesimo nella storia **31**, 131 (2010)
22. P. Rossi, *La nascita della scienza moderna in Europa* (Laterza, Bari, 1997)
23. P.C.W. Davies, *God and the new physics* (Simon and Schuster, New York, 1983)
24. E. Boncinelli, *Charles Darwin. L'uomo: evoluzione di un progetto?* (L'Espresso, Roma, 2012)
25. J.C. Lennox, *God's undertaker: Has science buried God?* (Lion Hudson, Oxford, 2007)

Chapter 27
The Bold and the Humble: Physics and Epistemology

R. Pucci and G.G.N. Angilella

For many years, the epistemological debate in Italy has been polarized between 'internists' and 'externists'. That is, between those who think that the scientific 'progress' should be explained by means of criteria internal to science itself (and these are related to Popper), and those who think that scientific revolutions could be understood only in connection with economical, political, and social revolutions. To Popper, the true scientist is he, who tries to demonstrate, with all means, that his own theory is false. When he succeeds in doing so, he abandons it, and tries and introduces a new theory, which he will endeavour to falsify again, but with a progressive shift of the problems. To the latter proposal, Kuhn objects that scientists

R. Pucci (✉) · G.G.N. Angilella
Dipartimento di Fisica e Astronomia, Università di Catania, Via S. Sofia 64, 95123 Catania, Italy
e-mail: renato.pucci@ct.infn.it

G.G.N. Angilella
e-mail: giuseppe.angilella@ct.infn.it

R. Pucci · G.G.N. Angilella
CNISM, UdR Catania, Via S. Sofia 64, 95123 Catania, Italy

R. Pucci · G.G.N. Angilella
IMM-CNR, UdR Catania, Via S. Sofia 64, 95123 Catania, Italy

G.G.N. Angilella
INFN, Sez. Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
Scuola Superiore di Catania, Università di Catania, Via Valdisavoia, 9, 95123 Catania, Italy

never behave this way, but rather make recourse to all means, including rhetorics, to try and convince their opponents that their own theory is correct. Scientific revolutions do not take place thanks to geniuses, but rather in connection with other revolutions of society. Young scientists adhere to a new paradigm more as an act of faith, rather than because a new theory is well established. Kuhn brilliantly demonstrates his thesis in several papers, including his most beautiful book, entitled *The structure of scientific revolutions* [1].

Our unconditioned admiration for Kuhn's work is apparent. There is, however, one aspect in Popper's falsificationism, *viz.* that on the distinction between scientific and non-scientific propositions, which is fundamental for modern epistemology, which has been not yet overcome, in our opinion. Certainly, Popper's falsifictionist criterion is more convincing that the criterion of verifiability introduced by the neopositivists. Kuhn does not face such a problem. In our opinion, the *humbler* position by Popper in contrast to the *arrogance* of the neopositivists leads him to support that non-scientific disciplines have all the same dignity. Not everybody agree with this connection, and among these is e.g., Gillies [2] (recently edited in Italian by Giorello [3]).

It is beyond any doubt, however, that Popper provides convincing arguments to prove that metaphysics is endowed with its own meaning, and that in certain cases it can even somehow positively influence science, as it was evidenced in the aforementioned work. It is true, therefore, that one cannot detect a clear borderline between science and other human activities, but it is also true that each of these activities possesses its own methodologies, that differentiate it from all the others. Therefore, it is important to have a criterion to distinguish them. Popper's criterion is not absolute, but is up to now the best, in order not to erect insurmountable barriers between one discipline and the other, but to understand that each of these is looking for its own way towards *truth.*

An *arrogant* attitude, similar to that of the positivists', can be recognized in some elementary particle physicists. According to these, the only fundamental aspects of Nature are elementary particles and the fundamental interaction forces among them. All the rest can be deduced by these, and if not, then it is meaningless. We will illustrate this through an anecdote. At the beginning of quantum mechanics, when there used to be congresses where both physicists and chemists participated, Dirac reportedly said *"Physicists discovered the electromagnetic force. All the rest is …chemistry."* Coulson then replied *"Physicists are those who know what to do, but do not know how."*

A much more profound answer was provided, several years later, by another great physicist, P.W. Anderson, in his seminal paper entitled *More is different* [4]. In his work, Anderson distinguishes between a *reductionist* approach to science, i.e., in which one tries to reduce all phenomena to some fundamental 'bricks' (*viz.* the elementary particles) and the fundamental interactions among them, and a *constructionist* approach, which makes use of such bricks to 'reconstruct' the entire building of Nature, as it is observed at the various scales. Anderson says that the *reductionist* approach intrinsically excludes the possibility that new phenomena may *emerge* in systems of many particles, as a consequence of the complexity of these systems, due

both to the large number of constituent particles, and to the hierarchy and intensity of the interactions among them.

Almost all physicists agree on the fact that it is necessary to go ever *deeper* in looking for the simplest constituents of the Universe, its fundamental laws. One hopes to find the most elementary particles and to unify all forces in a single force. Up to now, it has not been possible to unify quantum mechanics and general relativity, two sometimes contradicting theories, but there is still hope to find a new theory which will eventually able to eliminate the discrepancies between the two.

It is not correct, Anderson says, to think that reductionism implies constructivism, i.e., the possibility to construct, starting from a few fundamental laws, the whole Universe. Such an impossibility is due to the fact that when a system changes symmetry, i.e., there is a *spontaneous symmetry breaking,* the observed phenomenca can change, even drastically. For instance, it is not possible to extend the properties of a single atom to a diatomic molecule, since an atom is spherically symmetric, while a diatomic molecule is characterized by an axial symmetry, which is a reduced symmetry with respect to spherical symmetry, which thus appears *broken.* Another fundamental aspect concerns the comparison between a molecule and a solid, which is typically made of a large number of molecules of the same kind, of the order of 10^{23}. If to each molecule we associate an arrow, we have the same probability of finding the molecule with the arrow pointing upwards or downwards, but if we consider 10^{23} molecules of the same kind all with the arrows pointing upwards, at a sufficiently low temperature there are no enough thermal fluctuations which can flip the molecules' arrows downwards: *more is different!*

The laws governing solids are different, but equally fundamental, than those governing molecules or elementary particles.

Naturally, the same reasoning may be extended to chemistry, biology, economy, and possibly also to human 'sciences', such as psychology. It is appropriate at this point to emphasize the difference between science and technology. Up to now, we have made reference to science, but it is important to clarify that science is different from technology. On the other hand, many people think that science and technology are the same thing, and the modern society of consumerism and information unfortunately contributes (sometimes deliberately) to level out the intrinsic differences between these two concepts. One then tends to attribute to science all the positive or negative aspects typical of technological applications, which, according to the various points of view, have changed human life for better or for worse.

We will again illustrate this point with an anecdote. In 1983, one of us attended the Varenna summer school in physics. In that occasion, on one evening, he happened to share the same table for dinner with John Bardeen,[1] who told the following story

---

[1]John Bardeen (1908–1991) was the only scientist to receive two Nobel prizes, both for physics: one (in 1956, with William Shockley and Walter Brattain), for the invention of the transistor, and the other (in 1972, with Leon N. Cooper and John Robert Schrieffer) for the theory of superconductivity.

> On a day, I invited for lunch a friend of mine, a lawyer, and his wife. He arrived with a transistor radio, and told me: 'John, I apologize if I'm going to take advantage of our friendship, but since you invented the transistor, could you please repair my radio?' At which I replied: 'I can explain *why* the radio works, not *how* it works.'

From this statement one can understand most clearly the difference between science and technology.

Going back now to the extrapolation pointed forward previously from physics to psychology, we must conclude that such an extension is less trivial, when applied to religion. Indeed, in this case man must make a further act of *humility:* he cannot find the laws by himself, but must accept that another will provide him with them. The good news is that such laws, which for Christians reduce to a single one, *viz.* that of love, correspond to the most profound sense of the human being.

Such laws convey a message of hope, of joy, and of peace, but it is not always easy to accept them. On the other hand, we usually find ourselves to struggle with a mysterious God who would like to crash us, as Pino Ruggieri states in his last book [5]. We share what Bonhöffer terms an anticlerical attitude: *"that sniffing the track-of-the-human-sins to frame humanity."* On the other hand, we must recognize, as ever more people do in the world, that the meeting with Pope Francis with the Popular Movements of October 28, 2014[2] establishes a deep turning point in the active participation of Catholic Church in the contrast against the injustices and social inequalities. As in the Gospel, it becomes even now evident, in what the believers think and do, the centrality of paupers, of excluded people, of those who are at the margins of the world. The struggle of those should also be our struggle, the life of the Church.

In this context, three dimensions have been singled out: the country, the house, and the job, which must be an asset of everybody, and not just a privilege of a few. These problems should be felt with a profound spirit of 'solidarity' and in connection with the problems of peace achievement and environment protection. Scientists who share this life programme cannot but be of service of their own brethren.

References

1. T.S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, Chicago, 1962)
2. D. Gillies, *Philosophy of science in the twentieth century: four central themes* (Blackwell, Oxford, 1993)
3. D. Gillies, G. Giorello, *La filosofia della scienza nel* XX *secolo* (Laterza, Roma–Bari, 2010)
4. P.W. Anderson, Science **177**(4047), 393 (1972). doi:10.1126/science.177.4047.393
5. G. Ruggieri, *Della fede: la certezza, il dubbio, la lotta* (Carocci, Roma, 2014)

---

[2]Cf. e.g., the *Roman Observer*.

Chapter 28
Ettore Majorana's Early Scientific Production

R. Pucci and G.G.N. Angilella

Abstract Among Ettore Majorana's (1906–1938) earliest works, there is a brief memory on the Thomas–Fermi model, and a paper on the chemical bond in the helium molecular ion. Here, we emphasize Majorana's hypercritical attitude even towards his own work. Besides describing Majorana's contributions to the Thomas–Fermi model and to the understanding of the molecular bond, we argue that some of Majorana's seminal contributions in molecular physics already prelude to the idea of exchange interactions (or Heisenberg–Majorana forces) in his later works on theoretical nuclear physics. In all his papers, he tended to emphasize the symmetries at the basis of a physical problem, as well as the limitations, rather than the advantages, of the approximations of the method employed.

R. Pucci (✉) · G.G.N. Angilella
Dipartimento di Fisica e Astronomia, Università di Catania,
Via S. Sofia, 64, 95123 Catania, Italy
e-mail: renato.pucci@ct.infn.it

G.G.N. Angilella
e-mail: giuseppe.angilella@ct.infn.it

R. Pucci · G.G.N. Angilella
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci · G.G.N. Angilella
IMM-CNR, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
INFN, Sez. Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
Scuola Superiore di Catania, Università di Catania,
Via Valdisavoia, 9, 95123 Catania, Italy

Introduction

Ettore Majorana's most famous, seminal contributions are certainly those on the relativistic theory of a particle with an arbitrary instrinsic angular momentum [2], on nuclear theory [3] and on the symmetric theory of the electron and the positron [4]. In particular, the latter paper already contains the idea of the so-called Majorana neutrino [4], as has been correctly emphasized [5]. The quest for Majorana neutrinos is still the object of current fundamental research (see, e.g. Refs. [6] and [7] for a general overview).

In this note, we would like to reconsider two more papers by Majorana [8, 9], both on atomic and molecular physics, and show how they are precursor to his theoretical work on the exchange nuclear forces, the so-called Heisenberg–Majorana forces [3]. We will also try and emphasize his critical sense and great ability to catch the relevant physical aspects of a given problem, beyond his celebrated mathematical skills, as witnessed by contemporaries and colleagues who met him personally [10–12] (see especially Ref. [13] for more references).

Both Amaldi and Segrè have provided us with a vivid account of Majorana's first meeting with Enrico Fermi. Majorana and Fermi first met in 1928 at the Physical Institute in Via Panisperna, Rome. At that time, Fermi was working on his statistical model of the atom, known nowadays as the Thomas–Fermi model, after the names of the two authors who derived it independently [14–16]. Such a model provides an approximate alternative to solving Schrödinger equation [17], and paved the way to density functional theory [18].

Thomas–Fermi Model

Within Thomas–Fermi approximation, the electronic cloud surrounding an atom is described in terms of a completely degenerate Fermi gas. Following Ref. [17], one arrives at a *local* relation between the electron density $\rho(\mathbf{r})$ at position $\mathbf{r}$ with respect to the nucleus, and the momentum $p_F(\mathbf{r})$ of the fastest electron (Fermi momentum), as

$$\rho(\mathbf{r}) = 2 \cdot \frac{4\pi}{3} p_F^3(\mathbf{r}), \tag{28.1}$$

where the factor of two takes into account for Pauli exclusion. In Eq. (28.1), the Fermi momentum $p_F(\mathbf{r})$ depends on position $\mathbf{r}$ through the self-consistent potential $V(\mathbf{r})$ as

$$p_F^2(\mathbf{r}) = 2m[E_F - V(\mathbf{r})], \tag{28.2}$$

where E_F is the Fermi energy, and m is the electron mass. Fermi energy E_F is then determined via the normalization condition

$$\int \rho(\mathbf{r}) \, d^3\mathbf{r} = N, \tag{28.3}$$

where N is the total electron number, equalling the atomic number Z for a neutral atom. Inserting Eq. (28.2) into Eq. (28.1), making use of Poisson equation, and introducing Thomas–Fermi screening factor ϕ through

$$V(\mathbf{r}) - E_F = -\frac{Ze^2}{r}\phi(\mathbf{r}), \tag{28.4}$$

one derives the adimensional Thomas–Fermi equation for a spherically symmetric electron distribution,

$$\frac{d^2\phi}{dx^2} = \frac{\phi^{3/2}}{x^{1/2}}, \tag{28.5}$$

where

$$r = bx, \tag{28.6}$$

and b sets the length scale as

$$b = \frac{1}{4}\left(\frac{9\pi^2}{2Z}\right)^{1/3} a_0 = \frac{0.8853}{Z^{1/3}} a_0, \tag{28.7}$$

with a_0 the Bohr radius.

Equation (28.5) is 'universal', in the sense that the sole dependence on the atomic number Z comes through Eq. (28.7) for b. Once Eq. (28.5) is solved, the self-consistent potential for the particular atom under consideration is simply obtained by scaling all distances with b.

Asymptotic Behaviour of the Solution to Thomas–Fermi Equation

Fermi endeavoured to solve Eq. (28.5) analytically without success. On the occasion of his first meeting with Majorana, Enrico succinctly exposed his model to Ettore, and Majorana got a glimpse of the numerical results he had obtained over a week time, with the help of a primitive calculator. The day after Majorana reappeared and handled a short note to Fermi, where he had jotted down *his* results. Majorana was amazed that Fermi's results coincided with his own.

How could Majorana solve Eq. (28.5) numerically in such a short time without the help of any calculator? Various hypotheses have been proposed. Did he find an analytical solution? At any rate, there are no physically acceptable analytical solutions to Eq. (28.5) in the whole range $0 \leq x < +\infty$. The only analytical solution,

$$\phi(x) = \frac{144}{x^3}, \tag{28.8}$$

would have been found later by Sommerfeld in 1932 [19], and is physically meaningful only asymptotically, for $x \gg 1$.

The most likely hypothesis is probably that of Esposito [20], who, together with other authors [21], has found an extremely original solution to Eq. (28.5) in Majorana's own notes (see also Ref. [22]). The method devised by Majorana leads to a semi-analytical series expansion, obeying both boundary conditions for a neutral atom

$$\phi(0) = 1, \tag{28.9a}$$
$$\phi(\infty) = 0. \tag{28.9b}$$

In a recent work, Guerra and Robotti [23] have rediscovered a not well-known short communication by Majorana, entitled *Ricerca di un'espressione generale delle correzioni di Rydberg, valevole per atomi neutri o ionizzati positivamente* (*Quest for a general expression of Rydberg corrections, valid for either neutral or positively ionized atoms*) [8]. In that work, perhaps in the attempt of improving the asymptotic behaviour of the solution to Thomas–Fermi equation, Ettore requires that the potential vanishes for a certain finite value of x, say x_0, *both* for neutral atoms and for positive ions. He writes the self-consistent potential as

$$V(r) = \frac{Ze}{r}\phi + C, \tag{28.10}$$

where, for an atom positively ionized n times ($n = Z - N$), the constant C equals

$$C = \frac{n+1}{bx_0}e, \tag{28.11}$$

where

$$b = 0.47 \frac{1}{Z^{1/3}} \left(\frac{Z-n}{Z-n-1}\right)^{2/3} \text{Å}, \tag{28.12}$$

and the boundary conditions to Eq. (28.5) now read

$$\phi(0) = 1, \tag{28.13a}$$
$$\phi(x_0) = 0, \tag{28.13b}$$
$$-x_0\phi'(x_0) = \frac{n+1}{Z}. \tag{28.13c}$$

One immediately notices that, due to the new boundary conditions, Eq. (28.10) does not reduce to Eq. (28.4) even for $n = 0$, i.e. for a neutral atom. In other words, Majorana does not consider the potential $V(\mathbf{r})$ in a generic location of the electron

Fig. 28.1 Thomas–Fermi screening factor ϕ for the self-consistent potential of a neutral Ne atom ($Z = N = 10$). *Solid line* is Fermi's solution, *dashed line* is Majorana's solution, while the *light dashed line* has been obtained within Hartree–Fock approximation
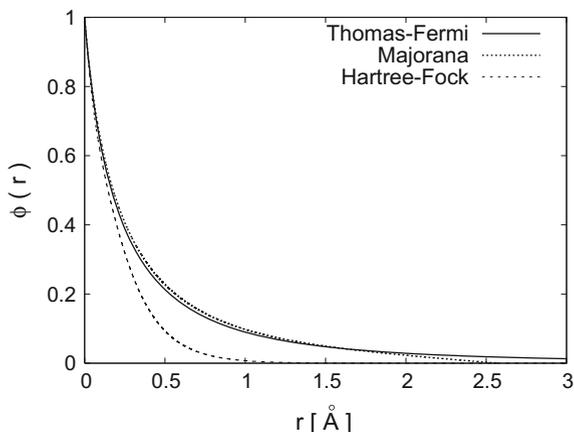
cloud, but the effective potential acting on a single electron, thus excluding the interaction of an electron with itself.

Probably, owing to his profound critical sense (let us remind that his colleagues in the Panisperna group nicknamed him the 'Great Inquisitor'), Majorana must have not excessively relied on his own solution [20], which however reproduced the numerical solution of Thomas–Fermi equation quite accurately. Probably, Majorana was looking for a solution which should not decrease so slowly as $x \to \infty$, as Eq. (28.8) does.

In Fig. 28.1 we report Thomas–Fermi screening factor ϕ as a function of r for a neutral Ne atom ($Z = N = 10$). The solid line refers to Fermi's numerical solution, with boundary conditions given by Eqs. (28.9), the dashed line refers to Majorana's solution, with boundary conditions given by Eqs. (28.13) with $n = 0$, while the light dashed line has been obtained within the Hartree–Fock approximation (see App. "Appendix: Derivation of Thomas–Fermi screening factor within Hartree–Fock self-consistent approximation" for a derivation). As it can be seen, Majorana's solution introduces only a minor correction to Fermi's solution at finite x values, but is strictly zero for $x \geq x_0$.

In his work on positive ions [24], Fermi considers a potential vanishing at a finite value $x = x_0$. However, instead of Eqs. (28.13), he employs the boundary conditions

$$\phi(0) = 1, \tag{28.14a}$$

$$-x_0\phi'(x_0) = \frac{n}{Z}, \tag{28.14b}$$

which in particular imply Eq. (28.9b) in the case $n = 0$, corresponding to a neutral atom.

In Fig. 28.2, we again report Thomas–Fermi screening factor ϕ as a function of r according to Fermi, Majorana and Hartree–Fock, respectively, but now for a

Fig. 28.2 Thomas–Fermi screening factor ϕ for the self-consistent potential of the Ne$^+$ ion ($Z = 10$, $N = 9$). *Solid line* is Fermi's solution, *dashed line* is Majorana's solution, while the *light dashed line* has been obtained within Hartree–Fock approximation
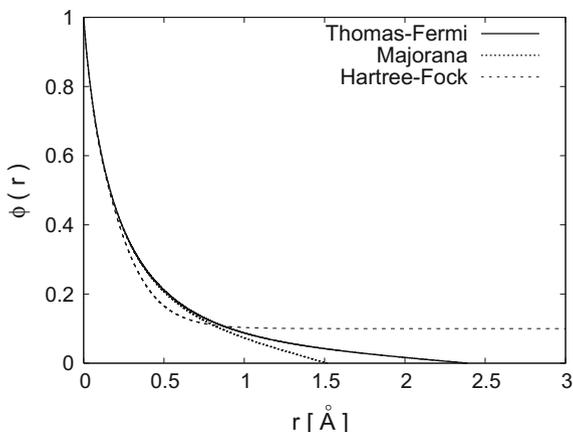
positively ionized Ne atom, Ne$^+$ ($Z = 10$, $N = 9$, $n = 1$). Majorana's solution again differs but marginally from Fermi's solution, but while for a neutral Ne atom Fermi's solution decreases too slowly, it decreases too rapidly for Ne$^+$.

Here, we are not disputing whether Majorana's note, Ref. [8], should be considered as a 'full' paper [25], nor do we want to undervalue the importance of the contribution analyzed in Ref. [20]. We would rather like to emphasize that Majorana was conscious that his correction[1] did not lead to substantial modifications to Fermi's solution of Eq. (28.5), including in the asymptotic limit ($x \gg 1$) [28].

Ettore never published anything else on this subject.

Although the present work is mainly based on Majorana's actually published works, rather than on speculations, it is remarkable that Majorana's work (Ref. [8]), which was presented by Majorana himself at the 22nd Adunanza Generale of the Italian Physical Society, and which was therefore certainly known to the group of the Via Panisperna boys, would never be referred to by Majorana's colleagues. Actually, Majorana's correction [8] is reported in full in the work by Fermi and Amaldi [27], but without any reference to the work (Ref. [8]) itself.

One may draw several hypotheses as to the reason of that, but we agree with Esposito's opinion [25], that this was due to Majorana's own 'hypercritical' attitude, which led him not to divulge the results of his own research, before these had reached a degree of completeness which was satisfactory to him.[2]

At the end of this section, we would like to mention that Ettore Majorana's most important contribution to atomic physics was Ref. [29], which deals with the behaviour of oriented atoms in a varying magnetic field. In this work [29], Majorana not only sets out the problem, but also provides it with a theoretical solution, employing for the first time the concept of 'Bloch sphere' and 'generalized Bloch sphere'.

---

[1]Flügge [26] erroneously attributes this correction to Amaldi. Probably, he was only aware of Fermi and Amaldi's final work, Ref. [27].

[2]See also below on the 'discovery' of the neutron.

Actually, Felix Bloch himself, in his 1945 work with Isidor Rabi [30], makes extensive reference to Majorana's paper, emphasizing its fundamental importance for the solution of the problem. Remarkably enough, the title of Bloch and Rabi's paper [30] is almost literally the English translation of the title of Majorana's paper [29]. (See also Refs. [31–33] for further comments.)

Helium Molecular Ion

In his successive work [9], Majorana deals with the formation of the molecular ion He_2^+. There again, Majorana demonstrates his exceptional ability to focus on the main physical aspects of the problem, while showing the limitations of his own theoretical approximations. He immediately observes that the problem is more similar to the formation of the molecular ion H_2^+, than to the reaction He + H. The most relevant forces, especially close to the equilibrium distance, are therefore the *resonance* forces, rather than the polarization ones. By *exchanging* the two nuclei, the system remains unchanged. Majorana makes then use of the method of Heitler and London [34], and emphasizes the importance of inversion symmetry with respect to the middle point between the nuclei, set at a distance R apart.

Heitler and London [34] introduced a relatively simple expression for the wave-function ψ of the two electrons in a hydrogen molecule H_2 in terms of the wave-function a and b of one electron in the atomic orbital corresponding to atom A and B, respectively:

$$\psi(1, 2) = a(1)b(2) \pm b(1)a(2), \tag{28.15}$$

where 1 and 2 denote the coordinates of the two electrons, respectively. The wave-function ψ_S, corresponding to the choice of the plus sign in Eq. (28.15), is symmetric with respect to the exchange of the coordinates of both electrons and nuclei, while ψ_A (minus sign in Eq. (28.15)) is antisymmetric. The full wave-function is globally anti-symmetric, but here we are neglecting its spin part, since the Hamiltonian is spin independent.

Despite its simplicity, the success of Heitler–London approximation relies on the fact that it explained the stability of the H_2 molecule, and could reproduce with remarkable accuracy the dependence of the total electronic energy E_I on the internuclear distance R. One obtains the attractive solution in correspondence with the eigenfunction ψ_S. It is relevant to stress at this point that, if one had considered only $a(1)b(2)$, or $b(1)a(2)$, in Eq. (28.15), the agreement with experimental data would have been rather poor. Therefore, the resonance or exchange term is quite decisive for establishing the chemical bond.

Heitler–London theory is even more accurate than the method of molecular orbitals [35–38] (see, e.g. Ref. [39] for a more detailed discussion), which in addition to Eq. (28.15) takes into account also for the ion-like configurations

$$a(1)a(2) \quad \text{and} \quad b(1)b(2), \tag{28.16}$$

corresponding to having both electrons on atom A, or B, respectively, on the same footing and with equal weights as the terms in Eq. (28.15). However, the theory can be improved by adding to Eq. (28.15) the two contributions in Eq. (28.16) with appropriate weights, to be determined variationally. In the large R range, also the polarization contributions become relevant, and the importance of ionic terms is even larger for the doubly ionized He molecule, He_2^{++} [40].

Owing to his usual critical sense, Majorana does employ Heitler–London approximation[3] for He_2^+, but he notes [9] that their approximation becomes

«unreliable for distant atoms.»

Therefore, instead of performing any numerical calculation, he introduces an effective nuclear charge, Z_{eff}, which affords an analytical solution. The fundamental effect he considers is nowadays dubbed as screening: a single electron does not see a bare nuclear charge Ze, but rather a reduced one, $Z_{eff}e$, because of the screening cloud of the other electrons.

The concept of an effective nuclear charge, already introduced for the helium atom, had been extended by Wang [41] to the hydrogen molecule. Probably Majorana was not aware of Wang's work, since he does not refer to it in his 1931 paper. In any case, Majorana is the first one to make use of such a method for He_2^+. In making reference to his own work [42], where Z_{eff} is used as a variational parameter for He_2^+, Pauling reports in a footnote[4]

«The same calculation with Z_{eff} given the fixed value 1.8 was made by E. Majorana [9].»

The above statement by Pauling is not quite correct. It seems to suggest the conclusion that Majorana did not perform a variational calculation to estimate Z_{eff} in the case of He_2^+, but this is not possible. The only values known at that time for Z_{eff} were those for the helium atom ($Z_{eff} = 1.69$) [17], and for the hydrogen molecule ($Z_{eff} = 1.166$) [41]. Majorana could determine the value $Z_{eff} = 1.8$ for He_2^+ only performing a variational calculation, even though this calculation was not explicitly reported. The variational value successively obtained by Pauling for He_2^+ was, in fact, $Z_{eff} = 1.833$.

By making use of his results, Majorana evaluates the equilibrium internuclear distance as $d = 1.16$ Å, in good agreement with the experimental value 1.087 Å. He can then estimate the vibrational frequency as $n = 1610$ cm^{-1}, which he compares with the experimental value 1628 cm^{-1}. Majorana concludes his paper by stating [9] that his own result is

«*casually* in perfect agreement with the experimentally determined value»

(our italics). Any other author would have emphasized such a striking agreement as a success of his own method, whereas Majorana rather underlines the drawbacks of his own approximations.

---

[3]In analogy with Heitler–London, Majorana writes the wave functions for He_2^+ as linear combinations of products of three atomic-like wave functions, of which at least one is relative to one of the two nuclei.

[4]See footnote on p. 359 of Ref. [43].

Fig. 28.3 Variational energies of the molecular ion He$_2^+$, as a function of the internuclear distance R. *Solid line* refers to the symmetric wave-function in Eq. (28.15), *dashed-dotted line* to the antisymmetric one, while *dashed line* refers to the 'non-bonding' case, where position exchange is neglected. Redrawn after Ref. [42] (see also Ref. [40])
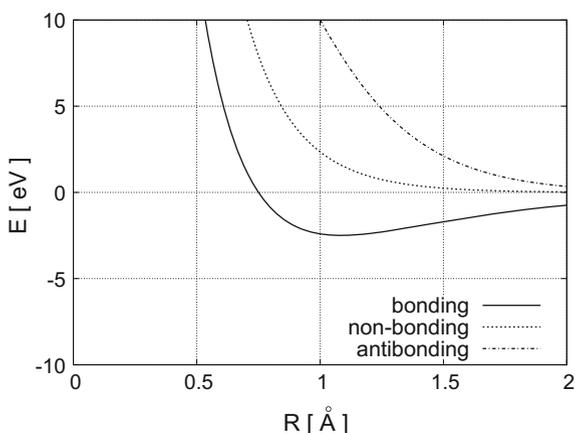
We would like to remind that he also estimates the minimum energy, i.e. the dissociation energy, finding the value $E_{min} = -1.41$ eV, but he had no available experimental data to compare with, at that time. However, he is not satisfied with such a result and collects [9]

«all the errors of the method under the words '*polarization* forces',»

which he estimates for very distant nuclei using the polarizability of the neutral He atom. He then finds $E_{min} = -2.4$ eV. More recent theoretical calculations, using the method of configuration interactions [44] or ab initio variational methods [45], have estimated the value $E_{min} = -2.47$ eV. The experimental value has been accurately determined quite recently [46] as $E_{min} = -2.4457 \pm 0.0002$ eV. We are not claiming that Ettore's result is more accurate than the theoretical results mentioned above. However, he certainly understood the essential physical effects for that system, and made use of appropriate approximations to estimate them. In particular, it is interesting how he emphasizes the quest for the *symmetries* of the system (see the translation of a paper by Majorana in Ref. [21]). As in the case of H$_2$, also for He$_2^+$ it is essential to include the position exchange term between He and He$^+$, in order to have chemical bonding, as it can be seen in Fig. 28.3, redrawn after Ref. [40]. If one had neglected the resonance He : He$^+$ $\rightleftharpoons$ He$^+$: He (see dashed line in Fig. 28.3), chemical bonding would have been impossible.

Recently, further considerations on Majorana and Pauling's works, in relation to the variational method and the genesis of the quantum theory of the chemical bond, have been made by Esposito and Naddeo [47–49].

The Discovery of the Neutron

Rutherford's pioneering work [50] paved the way not only to Bohr's atomic model, but also to nuclear physics.

In 1930 Bothe and Becker [51], like Rutherford, employed α particles against a berillium target in a scattering experiment. They observed the emission of a very penetrating radiation, which they interpreted as γ rays. In successive experiments, Irène Curie and Frederic Joliot [52, 53], her husband, developed further these experiments, but they arrived at similar conclusions. According to Emilio Segrè's account [54], Majorana thus commented the Joliots' results:

> «They haven't realized they have discovered the neutral proton.»

At this point we should remind that at that time it was believed that the nucleus was composed by protons and electrons. It was Chadwick [55] who soon after demonstrated that the radiation emitted in the Joliots' experiments was made up by neutral particles, whose mass is very close to the proton's mass. It was probably Fermi [54] who first distinguished between the *neutrinos* conjectured by Pauli, and the *neutrons* discovered by Chadwick.

Meanwhile, Majorana developed a theory of the nucleus containing protons and neutrons and then, according to Segrè [54],

> «he analyzed, as far as it was possible, the nuclear forces on the basis of the available experimental results, and he estimated the binding energies of the lightest nuclei. When he presented his work to Fermi and ourselves, we immediately recognized its importance. Fermi encouraged Majorana to publish his own results, but Majorana refused to do so, saying they were yet too incomplete.»

More than that, when Fermi asked Majorana whether he could make reference to his results during a forthcoming conference in Paris, Ettore mockingly replied he would agree, provided the reference was attributed to an old professor of electrochemistry, who was also going to attend the same conference. Obviously, Fermi could not accept Majorana's condition, and no reference was then made to his results during the conference.

Meanwhile, people started feeling the lack of a theory of nuclear forces, conveniently taking into account for the presence of both protons and nucleons in the nucleus. But where to begin with?

Heisenberg–Majorana Forces

To this aim, in three fundamental contributions [56–58], Heisenberg assumed hydrogen molecular ion H_2^+ as a model. He recognizes that the most important nuclear forces are not the polarization forces among the neutrons, or Coulombic repulsion among protons, but the exchange forces between protons and neutrons.

Fig. 28.4 Exchange interactions. Resonant forms in the hydrogen molecular ion, H: $H^+ \rightleftharpoons H^+$: H (*upper row*), and in the proton–nucleon pair inside a nucleus, $p: n \rightleftharpoons n: p$ (*lower row*)
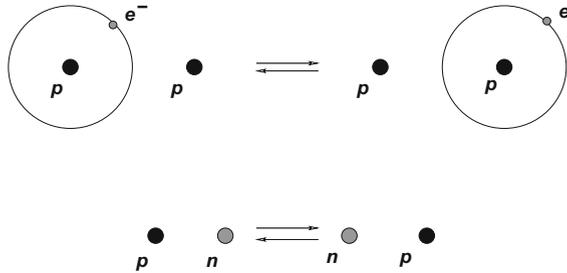
Heisenberg emphasizes that neutrons obey to Fermi statistics. Moreover, since a neutron possesses spin $\frac{1}{2}\hbar$, it cannot be simply thought of as composed of a proton plus an electron, unless the latter has zero spin, when inside a neutron.[5] A neutron is an elementary particle per se. The interactions postulated by Heisenberg are characterized by the exchange of both position coordinates and spins of the two nucleons.

Similarly, Majorana assumed that the fundamental nuclear forces are of exchange nature between protons and neutrons. However, he fully exploits the analogy with H_2^+ (see Fig. 28.4), regardless of spin.[6]

Let $\mathbf{r}_1, \sigma_1$ and $\mathbf{r}_2, \sigma_2$ stand for the position and spin coordinates of the first and the second nucleon, respectively, and let $\psi(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2)$ be the wave-function for a given nucleon pair [60]. Then Heisenberg exchange P^H implies

$$P^H \psi(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2) = \psi(\mathbf{r}_2, \sigma_2; \mathbf{r}_1, \sigma_1), \tag{28.17}$$

whereas Majorana exchange P^M implies

$$P^M \psi(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2) = \psi(\mathbf{r}_2, \sigma_1; \mathbf{r}_1, \sigma_2). \tag{28.18}$$

In Majorana's own notation (a part from a minus sign here included in the definition of $J(r)$), the exchange interaction then reads [3]

$$(Q', q'|J|Q'', q'') = J(r)\delta(q' - Q'')\delta(q'' - Q'), \tag{28.19}$$

where Q and q are the position coordinates of the neutron and the proton, respectively, and $r = |q' - Q'|$ is their relative distance. Majorana then plots a qualitative sketch of $J(r)$ (cf. Fig. 2 in Ref. [3]), which closely resembles the behaviour of the potential energy in H_2^+, when the internuclear repulsion is neglected (Fig. 28.5).

In the same paper [3], in addition to his knowledge of molecular physics, Majorana fully exploits also his acquaintance with the atomic statistical model. Indeed, he

---

[5]Besides considerations concerning the spin, such a model would require an enormous amount of energy to localize the electron within the neutron [57].

[6]Current literature usually employs the formalism of isotopic spin to describe the exchange character of the nuclear forces. However, as noted by Blatt and Weisskopf [60], this is equivalent to a description which makes use of the forces of Bartlett, Heisenberg, Majorana and Wigner.

Fig. 28.5 Kinetic, potential and total energies for the ground state of H_2^+, excluding nuclear repulsion, within the linear combination of atomic orbitals (LCAO) approximation. Cf. Fig. 2.4 in Ref. [59], where the same quantities have been obtained within a variational method
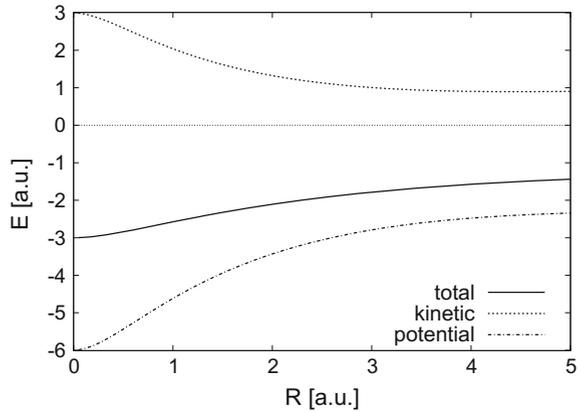
defines the nuclear density as

$$\rho = \frac{8\pi}{3h^3}(P_n^3 + P_p^3), \tag{28.20}$$

in complete analogy with Eq. (28.1), where P_n and P_p are the Fermi momenta of neutrons and protons, respectively. From this model, he derives an asymptotic expression ($\rho \to \infty$) for the exchange energy per particle,

$$a(\rho)|_{\rho\to\infty} = -\frac{n_2}{n_1 + n_2}J(0), \tag{28.21}$$

where n_1 and n_2 are the numbers of neutrons and protons, respectively. As in Thomas–Fermi model, the kinetic energy per particle, t say, is given by

$$t \propto \rho^{2/3}. \tag{28.22}$$

From the competition between kinetic and potential energy, the total energy attains a minimum as a function of r (cf. Fig. 1 in Ref. [3]).

Majorana's model explains two fundamental properties of nuclear physics [60]: (a) the density of nucleons is about the same for all nuclei (*density saturation*); (b) the binding energy per nucleon is about the same for all nuclei (*binding energy saturation*).

Concluding Remarks

In some of his fundamental papers, Majorana mainly focussed on the asymptotic properties of the potential and of the wave-function of an atomic or molecular system.

This is clearly demonstrated in his work on helium molecular ion, He$_2^+$ [9]. On the basis of his hypercritical spirit, Majorana was probably unsatisfied with the asymptotic behaviour of the screening factor ϕ within Thomas–Fermi model, but his note [8] is too short to confirm that. What we can certainly emphasize is his taste for the quest of symmetries, and their relevance to determine the main properties of a physical system [25]. This led him to demonstrate that the exchange symmetry is essential to the formation of the chemical bond. Exchange symmetry is also central in his model of the nuclear forces.

The quest for symmetries is evident in his famous work on the symmetrical theory of the electron and the positron [4]. There, he notes that

«all devices suggested to endow the theory [61] with a symmetric formulation, without violating its contents, are not completely satisfactory. […] It can be demonstrated that a quantum theory of the electron and the positron can be formally symmetrized completely by means of a new quantization procedure. […] Such a procedure not only endows the theory with a fully symmetric formulation, but also allows one to construct a substantially new theory for chargeless [elementary] particles (neutrons and hypothetical neutrinos).»

Several important experiments [6, 7] are currently under way to observe the 'Majorana neutrino'.

Appendix: Derivation of Thomas–Fermi Screening Factor within Hartree–Fock Self-Consistent Approximation

The solution of Hartree–Fock equations enables one to determine the (spherically symmetric) radial electron density

$$D(r) = 4\pi r^2 \rho(r), \tag{28.23}$$

normalized to the total electron number as

$$\int_0^\infty D(r)\, dr = N \tag{28.24}$$

(see, e.g. Ref. [62]). Figure 28.6 shows $D(r)$ for a neutral Ne (solid line; cf. Fig. 8.6 in Ref. [62]) and for Ne$^+$ (dashed line). The two peaks refer to the $1s$ and $2s\,2p$ shells, respectively. Although $D(r)$ is always strictly different from zero over the whole r range, it is an exponentially decreasing function of r, with $D(r) \approx 0$ roughly defining the atomic (respectively, ionic) radius.

By relating the electric field

Fig. 28.6 Radial electron
density $D(r)$ within
Hartree–Fock self-consistent
approximation for neutral Ne
(*solid line*) and singly
ionized Ne$^+$ (*dashed line*)
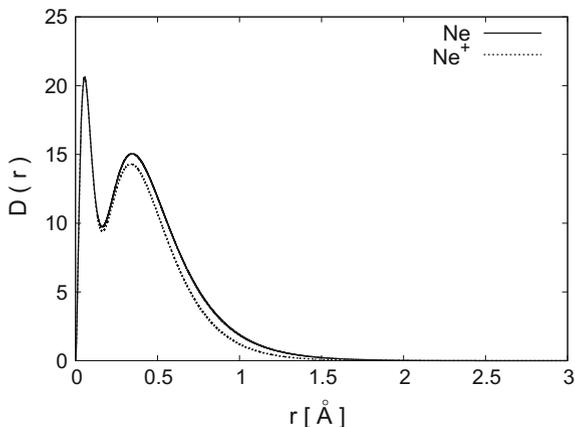
$$|\mathbf{E}| = \frac{1}{e}\frac{\partial V}{\partial r} = -\frac{Ze}{r}\phi' + \frac{Ze}{r^2}, \tag{28.25}$$

corresponding to the self-consistent potential, Eq. (28.4), to that generated by the
nucleus and the electron cloud within a distance r from the nucleus, by Gauss law,

$$|\mathbf{E}| = \frac{\sum_{r' \le r} Q}{r^2} = \frac{1}{r^2}\left[Ze - e\int_0^r D(r')\,dr'\right], \tag{28.26}$$

one finds

$$\phi' = \frac{1}{r}\left[\phi - 1 + \frac{1}{Z}\int_0^r D(r')\,dr'\right], \tag{28.27a}$$

$$\phi(0) = 1, \tag{28.27b}$$

where a prime here refers to derivation with respect to r.

Within Thomas–Fermi approximation, $\phi(r_0) = 0$, where r_0 is the ionic radius,
and the integration in the normalization condition, Eq. (28.24), should actually be
performed up to $r = r_0$. Then, Eq. (28.27a) yields

$$-r_0\phi'(r_0) = 1 - \frac{N}{Z}, \tag{28.28}$$

in agreement with Fermi's boundary condition, Eq. (28.14b).

Within Hartree–Fock approximation, $D(r)$ is in general nonzero at any finite r.
However, as $r \to \infty$, the potential experienced by a test charge is that of a charge
$(Z - N)e$, i.e.

$$V(r) - E_F \sim -\frac{(Z-N)e^2}{r}. \tag{28.29}$$

Comparing the above asymptotic behaviour of the potential with the definition of the screening function ϕ in Eq. (28.4), one has

$$\lim_{r \to \infty} \phi(r) = 1 - \frac{N}{Z}. \tag{28.30}$$

On the other hand, making use of the latter result, from Eq. (28.27a) it follows that

$$\lim_{r \to \infty} r\phi'(r) = 0, \tag{28.31}$$

thus implying that $\phi'(r)$ vanishes as $r \to \infty$ more rapidly that $1/r$ (in fact, it vanishes exponentially).

Finally, from Poisson equation

$$-\nabla^2 V = 4\pi e^2 [-Z\delta(r) + \rho(r)], \tag{28.32}$$

for a given electron charge distribution, Eq. (28.23), one obtains

$$\phi'' = \frac{1}{Z} \frac{D(r)}{r} \tag{28.33}$$

at any $r > 0$. Integrating once between r and ∞, and making use of the limiting value Eq. (28.31), one obtains

$$\phi'(r) = -\frac{1}{Z} \int_r^\infty \frac{D(r')}{r'} \, dr', \tag{28.34}$$

which combined with Eq. (28.27a) yields the desired screening factor

$$\phi(r) = 1 - \frac{1}{Z} \int_0^r D(r') \, dr' - \frac{r}{Z} \int_r^\infty \frac{D(r')}{r'} \, dr' \tag{28.35}$$

in terms of the Hartree–Fock self-consistent radial density $D(r)$. In particular, Eq. (28.35) manifestly fulfills the boundary conditions

$$\phi(0) = 1 \tag{28.36a}$$

$$\phi(\infty) = 1 - \frac{N}{Z}. \tag{28.36b}$$

In Figs. 28.1 and 28.2, light dashed lines represent Eq. (28.35), with $D(r)$ numerically obtained within Hartree–Fock self-consistent approximation for Ne and Ne$^+$, respectively, as shown in Fig. 28.6.

References

1. R. Pucci, G.G.N. Angilella, Found. Phys. **36**, 1554 (2006). DOI 10.1007/s10701-006-9067-7
2. E. Majorana, Nuovo Cimento **9**, 335 (1932)
3. E. Majorana, Z. Physik **82**, 137 (1933)
4. E. Majorana, Nuovo Cimento **14**, 171 (1937)
5. E. Amaldi, in *Strong and Weak Interactions. Present problems*, ed. by A. Zichichi (Academic Press, New York, 1966), p. 10
6. P. Sapienza for the NEMO collaboration, Nucl. Phys. B: Proc. Suppl. **145**, 331 (2005)
7. A. Bettini, *Fisica subnucleare* (Università degli Studi di Padova, Padova, 2004). Available at http://www.pd.infn.it/~bettini
8. E. Majorana, Nuovo Cimento **6**, xiv (1929)
9. E. Majorana, Nuovo Cimento **8**, 22 (1931)
10. E. Amaldi, Giornale di Fisica **9**, 300 (1968)
11. B. Preziosi (ed.), *Ettore Majorana: Lezioni all'Università di Napoli* (Bibliopolis, Napoli, 1987)
12. L. Bonolis, *Majorana, il genio scomparso* (Le Scienze, Milano, 2002)
13. E. Recami, Quaderni di Storia della Fisica del Giornale di Fisica **5**, 19 (1999). Also available as preprint arXiv:physics/9810023
14. L.H. Thomas, Math. Proc. Cambridge Phil. Soc. **23**, 542 (1926). DOI 10.1017/S0305004100011683
15. E. Fermi, Rendiconti dell'Accademia Nazionale dei Lincei **6**, 602 (1927)
16. E. Fermi, Z. Physik **48**, 73 (1928). DOI 10.1007/BF01351576
17. N.H. March, *Self-Consistent Fields in Atoms* (Pergamon Press, Oxford, 1975)
18. R. Pucci, Giornale di Fisica **27**, 256 (1986)
19. A. Sommerfeld, Rend. R. Accademia dei Lincei **15**, 293 (1932)
20. S. Esposito, Am. J. Phys. **70**, 852 (2002)
21. S. Esposito, E. Majorana Jr., A. van der Merwe, E. Recami, *Ettore Majorana: Notes on Theoretical Physics* (Kluwer, New York, 2003)
22. E. Di Grezia, S. Esposito, Foundations of Physics **34**, 1431 (2004)
23. F. Guerra, N. Robotti, (2005). Preprint arXiv:physics/0511222
24. E. Fermi, Mem. Accad. Italia (Fis.) **I**, 139 (1930)
25. S. Esposito, (2005). Preprint arXiv:physics/0512259
26. S. Flügge, *Practical Quantum Mechanics* (Springer, New York, 1974)
27. E. Fermi, E. Amaldi, Mem. Accad. Italia (Fis.) **6**, 119 (1934)
28. R. Pucci, N.H. March, J. Chem. Phys. **76**(8), 4089 (1982). DOI 10.1063/1.443483
29. E. Majorana, Nuovo Cimento **9**, 43 (1932)
30. F. Bloch, I.I. Rabi, Rev. Mod. Phys. **17**, 237 (1945)
31. G.G.N. Angilella, F.S. Cataliotti, R. Pucci, PoS Proc. Sci. **EMC2006**, 015 (2006)
32. G.G.N. Angilella, F.S. Cataliotti, R. Pucci, in *Idee, cultura e storia per la Città della Scienza*, ed. by P. Finocchiaro, M. Alberghina (G. Maimone, Catania, 2007), p. 38
33. G.G.N. Angilella, S. Bartalini, F.S. Cataliotti, I. Herrera, N.H. March, R. Pucci, in *Trends in Boson Research*, vol. 1, ed. by A.V. Ling (Nova Science, New York, 2006), p. 35. Preprint arXiv:cond-mat/0410287
34. W. Heitler, F. London, Z. Physik **44**, 455 (1927)
35. F. Hund, Z. Physik **51**, 759 (1928)
36. R.S. Mulliken, Phys. Rev. **32**, 186 (1928)
37. R.S. Mulliken, Phys. Rev. **32**, 761 (1928)
38. E. Hückel, Z. Physik **60**, 423 (1930)
39. C.A. Coulson, *Valence* (Oxford University Press, Oxford, 1952)
40. L. Pauling, L.O. Brockway, J.Y. Beach, J. Am. Chem. Soc. **57**, 2705 (1935)
41. S.C. Wang, Phys. Rev. **31**, 579 (1928)
42. L. Pauling, J. Chem. Phys. **1**, 56 (1933)
43. L. Pauling, E. Bright Wilson, *Introduction to Quantum Mechanics with Applications to Chemistry* (McGraw-Hill, New York, 1935)

44. J. Ackermann, H. Hogreve, Chem. Phys. **157**, 75 (1991)
45. P.N. Reagan, J.C. Browne, F.A. Matsen, Phys. Rev. **132**, 304 (1963)
46. L. Coman, M. Guna, L. Simons, K.A. Hardy, Phys. Rev. Lett. **83**, 2715 (1999)
47. S. Esposito, A. Naddeo, Found. Phys. **42**(12), 1586 (2012). DOI 10.1007/s10701-012-9685-1
48. S. Esposito, A. Naddeo, Majorana, Pauling and the quantum theory of the chemical bond (2013). Preprint arXiv:1306.6153 [physics.hist-ph]
49. S. Esposito, A. Naddeo, The genesis of the quantum theory of the chemical bond (2013). Preprint arXiv:1309.4647 [physics.hist-ph]
50. E. Rutherford, *Collected papers* (Interscience: J. Wiley & Sons, New York, 1963)
51. W. Bothe, H. Becker, Z. Physik **66**, 289 (1930)
52. I. Joliot-Curie, F. Joliot, Compt. Rend. **194**, 273 (1932)
53. I. Joliot-Curie, F. Joliot, Compt. Rend. **194**, 876 (1932)
54. E. Segrè, *Enrico Fermi, Physicist* (The University of Chicago Press, Chicago, 1970)
55. J. Chadwick, Nature **129**, 312 (1932)
56. W. Heisenberg, Z. Physik **77**, 1 (1932)
57. W. Heisenberg, Z. Physik **78**, 156 (1933)
58. W. Heisenberg, Z. Physik **80**, 587 (1933)
59. J.C. Slater, *Quantum theory of molecules and solids*, vol. 1 (McGraw-Hill, New York, 1963)
60. J.M. Blatt, V.F. Weisskopf, *Theoretical Nuclear Physics* (J. Wiley & Sons, New York, 1952)
61. P.A.M. Dirac, Proceedings of the Cambridge Philosophical Society, Mathematical and Physical Sciences **30**, 150 (1924)
62. B.H. Bransden, C.J. Joachain, *Physics of Atoms and Molecules*, 2nd edn. (Prentice Hall, London, 2003)

Chapter 29
Einstein and His Struggle for Peace

R. Pucci and G.G.N. Angilella

This year [2015] is the centennial of the publication of Einstein's paper (1915) on general relativity [1–5]. Among the numerous celebrations of Einstein's life and works, we will here remind a probably less emphasized aspect of his activity: the quest for peace in the world. Let us move from a pamphlet by Albert Einstein and Sigmund Freud [6]. This contains several short essays published by Freud in 1915, and the epistolary between Einstein and Freud in 1932. In the recent Italian edition, published by Boringhieri, the only contribution by Einstein is a five-pages letter sent to Freud. Its importance, evidently, resides in the author, rather than in the number of pages. The occasion of this epistolary stems from a proposal made to Einstein by the League of Nations and by International Institute of Intellectual Co-operation at Paris, led by Freud, *"to invite a person, to be chosen by* Einstein, *to a frank exchange*

R. Pucci (✉) · G.G.N. Angilella
Dipartimento di Fisica e Astronomia, Università di Catania,
Via S. Sofia, 64, 95123 Catania, Italy
e-mail: renato.pucci@ct.infn.it

G.G.N. Angilella
e-mail: giuseppe.angilella@ct.infn.it

R. Pucci · G.G.N. Angilella
CNISM, UdR Catania, Via S. Sofia, 64, 95123 Catania, Italy

R. Pucci · G.G.N. Angilella
IMM-CNR, UdR Catania, Via S. Sofia 64, 95123 Catania, Italy

G.G.N. Angilella
INFN, Sez. Catania, Via S. Sofia, 64, 95123 Catania, Italy

G.G.N. Angilella
Scuola Superiore di Catania, Università di Catania, Via Valdisavoia, 9, 95123 Catania, Italy

of views on any problem that Einstein himself *might select."* Most people, including Freud himself, expected that Einstein would choose a *"problem at the frontiers of today's knowledge"* (Freud's words). On the other hand, Einstein invites Freud to debate on the theme *"Is there any way of delivering mankind from the menace of war?"*

Einstein is fully conscious of his own enormous popularity and therefore of the great influence that his statements and ideas might have. However, he chooses to make use of such an influence not to exalt further his own name or his discoveries, but to rather spread a message of peace. In other words, he feels as a moral duty to employ his authoritativeness to induce the world to get organized in order to avoid war. To this mission he constantly devoted himself during his public life, which was dedicated and concerned, even in the remoteness implied by the severity of his studies.

This civil engagement begins with the countermanifest, written in 1914 by Einstein, Nicolai and Förster, in protest against the manifest in support of the war signed by 93 German intellectuals. World War I is about to break out, and Freud does in fact join with enthusiasm the idea of this possibility. This should not surprise, as many German, Austrian and Hungarian Jewish people participated with enthusiasm to the Great War, since they hoped therefore to finally become full citizens of the Fatherland they had chosen for themselves.

However, Freud soon realizes his own mistake, and is overwhelmed by a tremendous bewilderment. Not only has war upset the lives of men, but also the traditional way in which death has been conceived.

But let us get back to 1932. Einstein is not primarily concerned to assess the correctness of relativity (he knows it *is* correct),[1] rather than try to avert World War II, which he feels is initiating. The only solution he can imagine is the establishment of a supernational political organization, whereto the individual countries would entrust the authority to compose the inevitable reciprocal quarrels. Every state must waive some of its own jurisdiction.

Einstein does not behave naïvely: he can easily see the difficulties in such a proposal, and addresses several questions to Freud, which should provide him with advice on how to achieve his goals. Einstein's questions were: (a) How is it possible that a minority can enslave to its own greed the people's mass, which would only suffer from a war? (b) How is it possible that such a mass allows to get excited by the communication media, the school, and sometimes even by the religious organizations, up to the fury and its own holocaust? (c) Is there any possibility to guide

---

[1] In 1919, Sir Arthur Eddington led one of the expeditions organized in order to obtain an experimental verification of one of the predictions of Einstein's general relativity: the measurement of the gravitational deflection of light. Such a measurement can be performed only during a solar eclipse [7]. Therefore, Eddington and collaborators had to go to the island of Príncipe, near Africa. Despite some clouds, the observation was effected on May 29, 1919, and it fully confirmed the predictions of the new theory. Some time later, a journalist asked Einstein how he would have reacted if Eddington's observations did not confirm his theory. Einstein mildly replied: "I would have been very sorry for that nice gentleman [Eddington]. The theory is correct".

men's psychic evolution, so that they become able to resist to the psychosis of hate and destruction?

Freud provides Einstein with precise answers to each of his questions, and notes that some answers are already contained in his letter (such as that the supernational structure should be endowed with some force, otherwise its deliberations would not be respected, or that of a school which should be able to kindle the fire of peace in its pupils' hearts), but also makes recourse to his psychoanalytic instruments to illustrate his opinion.

As days pass by, the wave of anti-Zionism increases in Germany, along with the hate towards Einstein and his 'Jewish' science. Einstein moves to Belgium, but even that country is too close to Germany, and he could thus be easily reached by killers. At the beginning of 1933, when the Nazi took the power, the path of his emigration to the United States of America had been already laid out, and in winter 1933 Einstein started his new appointment at the Institute of Advanced Studies, which had been recently established by Flexner in Princeton. (In 1930, Louis Bamberger and Caroline Frank Fuld had donated 5 millions dollars to the aim of founding this Institute.) It should be noted that Einstein entered the USA as a visitor: at the beginning he had no residence permit, let alone the US citizenship.

Some feminist associations opposed to Einstein's entrance into the US, since in their opinion he favoured subversive and pacifist doctrines. In this respect, Einstein once declared to a representative of the Associated Press: *"Why should one admit such a man, as vulgar as to oppose to any war, but the inevitable one with his own wife?"* [8] Similarly, with the same seriousness, he stated to the New York Times (June 20, 1932): *"Only a life lived for others is the life worth while."* With time, however, he won the liking of a large part of the American people with his anticonformism, not only through his way of dressing himself, but because of his ideas: his constant rejection to take sides with an ever more inhuman society.

Meanwhile, World War II is approaching, and many scientists are aware that it will be probably fought with new and terrific weapons, the atomic bombs. Two physicists from Columbia University, Enrico Fermi and Leó Szilárd, a Hungarian national who fled away from Berlin University, and Eugene Wigner, a physicist from Princeton, a Hungarian himself and a Nobel prize awardee, tried to convince President Roosevelt on the opportunity to build the atomic bomb. Their efforts were not successful, and so they induced Einstein to make use of his great authoritativeness. Einstein, a convinced pacifist, realized that the Nazi could indeed build the atomic bomb themselves, and that, in the hands of that raving insane of Hitler, could only mean the end of human civilization and of individual rights.

On August 2, 1939, Einstein sent a letter to President Roosevelt, its first sentences being as follows:

> Some recent work by E. Fermi and L. Szilard … leads me to expect that the element uranium may be turned into a new and important source of energy in the immediate future. … A single bomb of this type, carried by boat and exploded in a port, might very well destroy the whole port together with some of the surrounding territory.

Thus the Manhattan Project got started.

When the first atomic bomb, significantly called 'Trinity', exploded in Alamogordo, on July 16, 1945, everybody knows that Oppenheimer said *"The physicists have known sin; and this is a knowledge which they cannot lose."* In an interview after the event, Oppenheimer then said

> We knew the world would not be the same. A few people laughed, a few people cried, most people were silent. I remembered the line from the Hindu scripture, the Bhagavad-Gita; Vishnu is trying to persuade the Prince that he should do his duty and, to impress him, takes on his multi-armed form and says, 'Now I am become Death, the destroyer of worlds.' I suppose we all thought that, one way or another.

Several scientists, including Einstein, tried, to no avail, to prevent any further use of the atomic bomb, which was launched on Hiroshima and Nagasaki. Some physicists had a crisis of conscience. Oppenheimer himself was indicted with communism, for having delayed the construction of the hydrogen bomb, and he was forbidden to access to military classified information. Rasetti, one of the Via Panisperna boys, not only refused to take part in the Manhattan Project, but even changed job, and successfully dedicated himself to biology.

When the war was over, Einstein undertook again his struggle for peace. He writes many papers, among which the most important are 'Atomic war or peace' (1947), and 'The war is won but peace is not' (1945). In this latter essay [9], he writes

> Alfred Nobel invented the most powerful explosive ever known up to his time, a means of destruction par excellence. In order to atone for this, in order to relieve his human conscience, he instituted his awards for the promotion of peace and for achievements of peace.
>
> Today, the physicists who participated in forging the most formidable and dangerous weapon of all times are harassed by an equal feeling of responsibility, not to say guilt. We cannot desist from warning, and warning again, we cannot and should not slacken in our efforts to make the nations of the world, and especially their governments, aware of the unspeakable disaster they are certain to provoke unless they change their attitude toward each other and toward the task of shaping the future.
>
> We helped in creating this new weapon in order to prevent the enemies of mankind from achieving it ahead of us, which, given the mentality of the Nazis, would have meant inconceivable destruction and the enslavement of the rest of the world.
>
> We delivered this weapon into the hands of the Americans and the British people as trustees of the whole [of] mankind, as fighters for peace and liberty. But so far, we fail to see any guarantee of peace. We do not see any guarantee of the freedoms that were promised to the nations in the Atlantic Charter. The war is won, but the peace is not.
>
> …Otherwise, human civilization will be doomed.

Einstein proposes again, with numerous more details, the idea of a supernational government to guarantee peace in the world His position is now shared by many, among whom Emery Reves, a collaborator of Churchill's, who writes in his book 'The anatomy of peace'

> Today, sovereignty has far too narrow a basis; it no longer has the power it should and was meant to have. The word is the same. The conception it expresses is the same. But the surroundings have changed. The conditions of the world have changed. And this changed situation calls for corresponding changes in the interpretation of this basic principle, if we desire to preserve this, the only foundation of democratic society yet discovered.

However, many are still accusing him to be an idealist, without any practical sense. Or worse! In the United States, there are some who accuse him to be a communist. In 1947, he sends an open letter to the General Assembly of the United Nations. There, Einstein denounces that

> The progress of technological development has not increased the stability and the welfare of humanity. Because of our inability to solve the problem of international organization, it has actually contributed to the dangers which threaten peace and the very existence of mankind.

He then makes his own proposal.

From the Soviet Union, several scientists (S.I. Vavilov, A.N. Frumkin, A.F. Ioffe, and N.N. Semyonov) reply with an open letter entitled 'The wrong ideas of Dr Einstein' (1947). Basically, they surmise that Einstein is proposing a supernational government in order to have the whole world ruled by the United States [9].

To no avail there came Einstein's reply. Einstein continued to struggle for a different school, a school which would teach the ideals of freedom and peace, for conscientious objection up to civil disobedience, a point of view which he shared with don Milani, for disarmament and the end of armies, for a Jewish state in Palestine.

On April 11, 1955, for the last time, he undersigned a pacifist manifesto (this time written by Bertrand Russell), in which all Nations were invited to give up with nuclear weapons. On the morning of April 13, the ambassador of Israel paid a visit to Einstein to record a message that he wanted to convey to the radio and the television, on the occasion of the forthcoming Yom ha-Atzma'ut, Israel's Independence Day. The incomplete text ended thus

> Not one statesman in a position of responsibility has dared to pursue the only course that holds out any promise of peace, the course of supranational security, since for a statesman to follow such a course would be tantamount to political suicide. Political passions, once they have been fanned into flame, exact their victims.

These were probably Einstein's last words.

References

1. A. Einstein, Sitzungsber. K. Preuss. Akad. Wiss.: Phys. Math. Kl. **1915 (part 1)**, 315 (1915)
2. A. Einstein, Sitzungsber. K. Preuss. Akad. Wiss.: Phys. Math. Kl. **1915 (part 2)**, 778 (1915)
3. A. Einstein, Sitzungsber. K. Preuss. Akad. Wiss.: Phys. Math. Kl. **1915 (part 2)**, 831 (1915)
4. A. Einstein, Sitzungsber. K. Preuss. Akad. Wiss.: Phys. Math. Kl. **1915 (part 2)**, 844 (1915)
5. A. Einstein, Sitzungsber. K. Preuss. Akad. Wiss.: Phys. Math. Kl. **1916 (part 1)**, 688 (1916)
6. A. Einstein, S. Freud, *Warum Krieg? Ein Briefwechsel* (Institut International de Coopération Intellectuelle, Société des Nations, Paris, 1933). Perché la guerra? (Torino, Boringhieri, 2012)
7. A. Pais, *Subtle is the Lord . . . The science and the life of Albert Einstein* (Oxford University Press, Oxford, 1982)
8. P. Frank, *Einstein: his life and times* (Knopf, New York, 1947)
9. A. Einstein, *Out of my later years* (Estate of Albert Einstein, New York, 1956)

Author Index

Subject Index