David L. Block
Kenneth C. Freeman
Ivânio Puerari *Editors*

Galaxies and their Masks

A Conference in Honour of
K.C. Freeman, FRS

Springer

Galaxies and their Masks

David L. Block · Kenneth C. Freeman ·
Ivânio Puerari
Editors

Galaxies and their Masks

A Conference in Honour of K.C. Freeman, FRS

Springer

Editors
David L. Block
School of Computational
and Applied Mathematics
University of the Witwatersrand
Johannesburg
WITS 2050, South Africa
igalaxy@iafrica.com;
david.block@wits.ac.za

Kenneth C. Freeman
Australian National University
Weston Creek, ACT 2611, Australia
kcf@mso.anu.edu.au

Ivânio Puerari
Instituto Nacional de Astrofísica
Óptica y Electrónica (INAOE)
Santa Maria Tonantzintla, Puebla 72840, México
puerari@inaoep.mx

ISBN 978-1-4419-7316-0 e-ISBN 978-1-4419-7317-7
DOI 10.1007/978-1-4419-7317-7
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2010936375

Preface

The year: 1660.
The date: November 28.
Present: The Lord Brouncker, Mr Boyle, Mr Bruce, Sir Robert Moray, Sir Paule Neile, Dr Wilkins, Dr Goddard, Dr Petty, Mr Ball, Mr Hooke, Mr Wren, and Mr Hill.
Occasion: A lecture by Mr Wren at Gresham College, United Kingdom.

After Christopher Wren had delivered his lecture at Gresham College on that historic occasion in November 1660, "they did according to the usual manner, withdraw for mutual converse."

It was in 1660 that the Royal Society was founded, with 12 persons present. This year, 2010, is thus a special year for scientists worldwide: it celebrates the 350th anniversary of the founding of the Royal Society, whose current President is Martin Rees.

One of the enormous challenges facing scientists in the 1600s was the great need for the classification of objects they were studying, particularly in the field of botany.

The seeds for classification lie in the works of the British naturalist John Ray (1628–1705), who commencing in 1660 with his *Catalogus plantarum circa Cantabrigiam nascentium* (Catalogue of Cambridge Plants) – published in the year in which the Royal Society was founded – and ending with the posthumous publication of *Synopsis Methodica Avium et Piscium* in 1713, pioneered systematic studies on plants, birds, mammals, fish, and insects.

In these works, Ray brought order to the chaotic mass of names in use by the naturalists of his time. Like Linnaeus, Ray searched for the "natural system," a classification of organisms that would reflect the Divine Order of Creation. Unlike Linnaeus, whose plant classification was based entirely on floral reproductive organs, Ray classified plants by overall morphology: the classification in his 1682 book *Methodus Plantarum Nova* draws on flowers, seeds, fruits, and roots. Ray's plant classification system was the first to divide flowering plants into monocots and dicots. This method produced more "natural" results than "artificial" systems based on one feature alone; it expressed the similarities between species more fully. Ray's system greatly influenced later botanists such as Jussieu and de Candolle and schemes based on total morphology came to replace those which focused on only

one feature or organ. Ray cautioned against blind acceptance of authorities: in *The Wisdom of God*, he wrote:

> Let it not suffice to be book-learned, to read what others have written and to take upon trust more falsehood than truth, but let us ourselves examine things as we have opportunity, and converse with Nature as well as with books.

And so it is with galaxies. We need to observe them; we need to classify them as comprehensively as possible.

The early seeds for galaxy classification were sown by an amateur astronomer, John Reynolds, who rose to the position of President of the Royal Astronomical Society of London. Hubble had urged Reynolds to throw his ideas into a galaxy classification scheme, which Reynolds duly did. The name of John Reynolds today is almost unknown to the modern astronomical community; for a full discussion of Hubble's failure to acknowledge his source Reynolds, the reader is referred to the book *Shrouds of the Night* authored by D.L.B. and K.C.F.

What has emerged is that galaxies are extremely complex systems, each presenting a different set of "masks." The use of the word "masks" in galaxy morphology can be traced back approximately two decades, as described in a paper "Cosmic Masks Still Dance" by one of us (DLB). The important point here is that without a detailed understanding of galaxies and their masks, astronomers walk in a land of shadows.

There are many masks: masks of cosmic dust, dark matter masks, dynamical masks – the list continues! These masks "dance" – in other words, they are dynamically active.

Masks of dust have today taken central stage, whether it be in the study of supernovae (Chapter "Supernovae, Dust, and Cosmology" by Schmidt, this volume) or of galaxies. In earlier years, dust was considered as somewhat of a *nuisance*; something not to be a grand focal point of research. The late Mayo Greenberg often reflected on the challenges he encountered in his early laboratory days.

Enter astronomer Ben Gascoigne, who was a close colleague and friend of one of ours (KCF), who passed away in March of this year. Gascoigne did so much of his pioneering observations of variable stars in the Large Magellanic Cloud by having the Director at Mt. Stromlo, Sir Richard Woolley, allocate 9 months of time to him and to his colleague, Gerry Kron.

The instrument: the 30-in. Reynolds reflector, donated by the same John Reynolds we met earlier. In those formative years, this telescope ranked as one of the largest telescopes in the southern hemisphere.

While the German astronomer Walter Baade had access to the giant 200-in. reflector telescope in California, Ben recalls how thrilled he was to be given unlimited access to the Reynolds instrument in the early 1950s. At this time, Gascoigne, Kron, and Baade were intensely interested in determining distances to nearby galaxies. From the northern hemisphere, Baade could observe Cepheid variable stars in the Andromeda Spiral, but not in the Magellanic Clouds – which are solely the domain of observers in the southern hemisphere.

Those Magellanic Clouds held many secrets: it is no wonder that they have been referred to as *Rosetta Stones*. The results of 9 months of observing time with a modest telescope follow in Gascoigne's own words:

> . . . the dust content of the galaxy [the Milky Way] came out three times greater . . . the Magellanic Clouds twice as far away [as hitherto assumed], the cosmic distance scale twice as large, and the universe twice as old . . . And the piece I liked, the key observations were made with a 30-inch [the Reynolds telescope], as opposed to Baade's 200-inch.

Gascoigne could clearly see the effects of masks of cosmic dust in his observations of our Milky Way galaxy. Prior to the photographic era, such masks were captured visually. In this context, one is reminded of an exquisite set of drawings of the Milky Way, produced at Birr Castle in Ireland, by the German astronomer Otto Boeddicker (1853–1937). Boeddicker became the astronomical assistant to Lawrence Parsons, the 4th Earl of Rosse. Boeddicker' drawings of the Milky Way (one of these is reproduced in Fig. 1) are breathtaking and the set, made over a period of 6 years, was published in 1892.

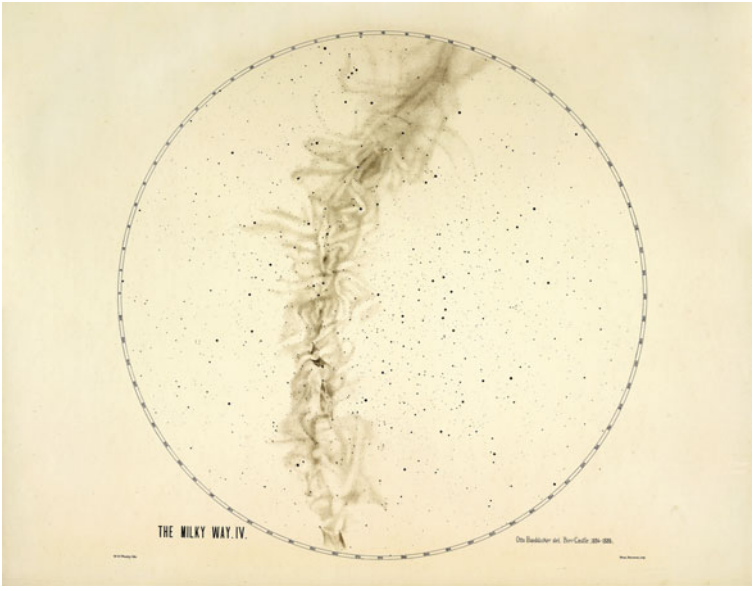


Fig. 1 A remarkable drawing of the Milky Way by Otto Boeddicker, astronomical assistant to Lawrence Parsons, 4th Earl of Rosse

It is fitting to hold a conference "Galaxies and their Masks" now, to assess the situation in 2010. (Note from David: The principal reason for selecting the year 2010 was that Ken turns 70 in 2010, and astronomers around the globe were eager to celebrate Ken's 70th birthday).

The venue chosen was in the Namib Desert: Sossusvlei, with its spectacular set of orange-reddish dunes, to be precise. Sossusvlei lies within the Namib-Naukluft National Park and is fed by the Tsauchab river. The area boasts a veritable sea of

sand dunes, some rising to more than 300 m above the desert floor. Life in the desert sands is diverse, having adapted to arid conditions. One is reminded, for example, of the "fog basking" beetle in the Namib Desert known as *Onymacris unguicularis*. The beetle derives its supply of water from aperiodic fog collection. To quote ecologists Hamilton and Seeley, "The Namib Desert along the south-western coast of Africa supports a sand dune fauna without counterpart elsewhere in the world."

Namib is the *Nama* word for open space, and the Namib Desert (considered to be the oldest desert on the planet) provides the name for the land of open spaces – Namibia. Open spaces below, and open spaces above: each night, the *Via Lactea* or Milky Way delineated a gargantuan arch above our heads; the Large Magellanic Cloud was particularly magnificent.

The conference itself took 2 years of planning. It was a continual interchange of emails between Namibia and South Africa on the one hand, and another interchange between SOC co-chairs K.C.F. and D.L.B., on the other. Our conference venue was remote, but spectacular (Fig. 2). The nearest large towns or cities are 5 h away, on gravel roads. Everything, down to the last bottle of wine, had to be planned well in advance. One of us (DLB) visited Namibia twice to carefully plan all logistics together with Andrew Bassingthwaighte and Franzpeter Ackermann, including the décor of the conference venue. Approximately 50 of Ken's closest friends and collaborators joined us at the conference, which, due to limited accommodation facilities in the desert, was by invitation only. There were two sets of presentations: review talks (of 30 min duration) and shorter (20 min) contributed talks.

Fig. 2 The venue of our conference "Galaxies and their Masks" – the Sossusvlei Lodge in Namibia. The rustic colors of the chalets blended in beautifully with the sands of the surrounding Namib Desert. Photo: Trevor Gould

The scientific program was exceptional. K.C.F. remains very grateful to those who, despite the most demanding schedules, still registered for the conference 2 years ago.

The only guidance that we offered delegates for the theme of the meeting "Galaxies and their Masks" was that masks should be discussed, in the broadest possible sense. For example,

— masks of cosmic dust, particularly relevant not only for galaxies, but also for supernovae too!
— masks of dark matter
— magnetic fields
— very faint galactic exteriors
— phase-mixed substructures
— difficulties in estimating chemical abundances, understanding nucleogenesis, and chemical evolution
— lack of telescope aperture and appropriate instrumentation
— difficulties in simulating disk galaxies
— hot halos of galaxies, and
— how globular clusters fit into galaxy formation scenarios.

A special "Bush Dinner" (Fig. 3) had been organized in honour of the 70th birthday of K.C.F. The dinner took place approximately 45 min drive from the Lodge, in the actual desert. Candles blazed brilliantly (Fig. 4), fires roared, a marimba band and musicians played, while the starry vault of the Milky Way beamed from above.

A special moment during the Bush Dinner occurred when some of Ken's former students stood up to sing a special song, entitled "Observing with Ken".

The singers were Jayanne English, Gayandhi de Silva, Isabel Perez, Carl Grillmair, Mary Putman, and Mary Williams.

Some inside information from Gayandhi:

> "It [the composition-planning] was done at the bar together with wine, martini and champagne." They considered adapting their song to those "from the Sound of Music, or the Macarena. After trying out all these options we decided that the Waltzing Matilda was the best."

Here is their song, which meant so much:

Observing with Ken

"Once a jolly supervisor camped on Mt. Stromlo
Under the shade of a Chinese Elm tree
And he sang as he made red marks upon our thesis drafts
Who'll come observing at Siding Spring with me?

Observing with Ken, observing with Ken
We'll go observing together with Ken
And he sang as he made red marks upon our thesis drafts
Who'll come observing at Siding Spring with me?

Down came a postgrad to drink at the Whig and Pen
Up jumped the supervisor, grabbed them with glee
And he sang as he showed that student tricks with DBS
You'll come observing at Siding Spring with me!

Observing with Ken, observing with Ken
We'll go observing together with Ken
And he sang as he showed that student tricks with DBS
You'll come observing at Siding Spring with me!

Then came the thesis, printed and submitted
Down came the referees; one, two, three
Where has the postgrad gone to work overseas?
Who'll come observing at Siding Spring with me?

Observing with Ken, observing with Ken
We'll go observing together with Ken
And he sang as the student went to work overseas
Who'll come observing at Siding Spring with me?

Up jumped the supervisor checking all the applications
Who'll be the 55th student asked he?
And his inspiration's felt as you pass by MSSSO
Who'll come observing at Siding Spring with me?

Observing with Ken, observing with Ken
We'll go observing together with Ken
And his inspiration's felt as you pass by MSSSO"

An unusual photograph of Ken and Margaret, photographed in silhouette during the song, appears in Fig. 5.

At night, astronomer Trevor Gould set up an 8-in. telescope; delegates could marvel at ω Centauri, 47 Tucanae, the Large Magellanic Cloud, the rings of Saturn, and much more. Trevor Gould deserves a very special note of thanks, for his remarkable commitment to "Galaxies and their Masks". The telescope itself had to be sent by courier to Namibia and was re-assembled in Sossusvlei. As each day dawned, the sunrises were spectacular (Fig. 6).

The proceedings of "Galaxies and their Masks" are contained in this volume, which we trust will capture the intellectual enrichment of our incredible meeting. Not every delegate was able to present a written version of their talk, but we are very grateful to all of the delegates for finding the time in their busy schedules to participate in the Sossusvlei conference. Our rule of thumb in drawing up the index is that, as closely as possible, printed papers should appear in the same order as that in which they were delivered at Sossusvlei. The groupings of talks at the conference was according to themes, ending with the high-redshift universe. The editors have included several photographs taken by delegates; these appear, space permitting, between papers in the volume. There is no specific ordering of these

Fig. 3 Fires roared and a marimba played as delegates entered the site for the "Bush Dinner" in the heart of the Namib Desert. Delegates were driven in open vehicles to the site – a drive of approximately 45 min. The Milky Way blazed brilliantly from above. Photo: Robert Groess

Fig. 4 A glass of wine seen in silhouette against a candle. Each candle was placed in a brown paper bag filled with sand; dozens of these candles contoured the environs of the "Bush Dinner." Photo: Robert Groess

Fig. 5 Ken and Margaret Freeman, photographed in silhouette, as a special song was sung by some of Ken's former Ph.D. students. Photo: Lourdes Verdes-Montenegro

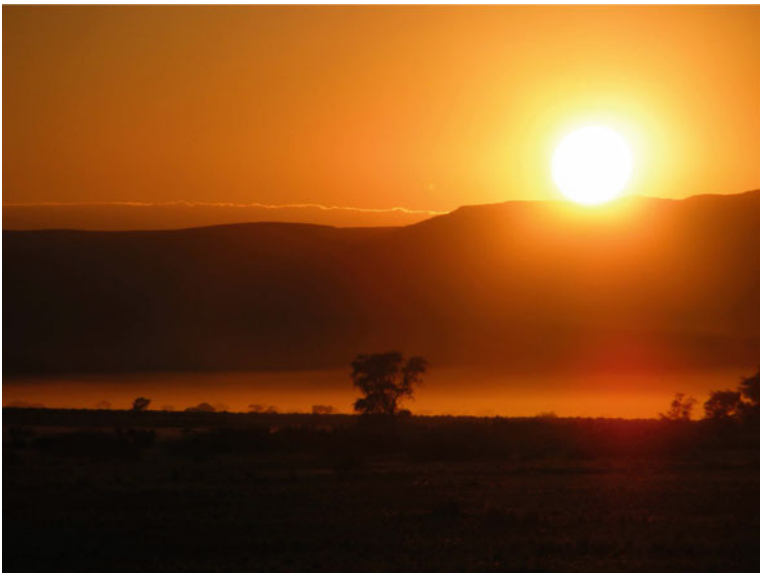


Fig. 6 Sunrise at Sossusvlei: each sunrise is unique. The entire countryside suddenly "comes alive" as the first rays of the sun impinge upon this most exquisite section of our planet. Brilliant hues of yellow and of gold are frequently seen, as captured in this image. Photo: Trevor Gould

photographs interspersed between papers. Only images referenced in the text of papers bear captions. We warmly thank those delegates who sent us their images for editorial consideration and possible inclusion. Our cover photograph was secured by Jacqueline Riffault-Silk; the description thereof appears below. We thank every Chairperson for their dedication and careful attention to time.

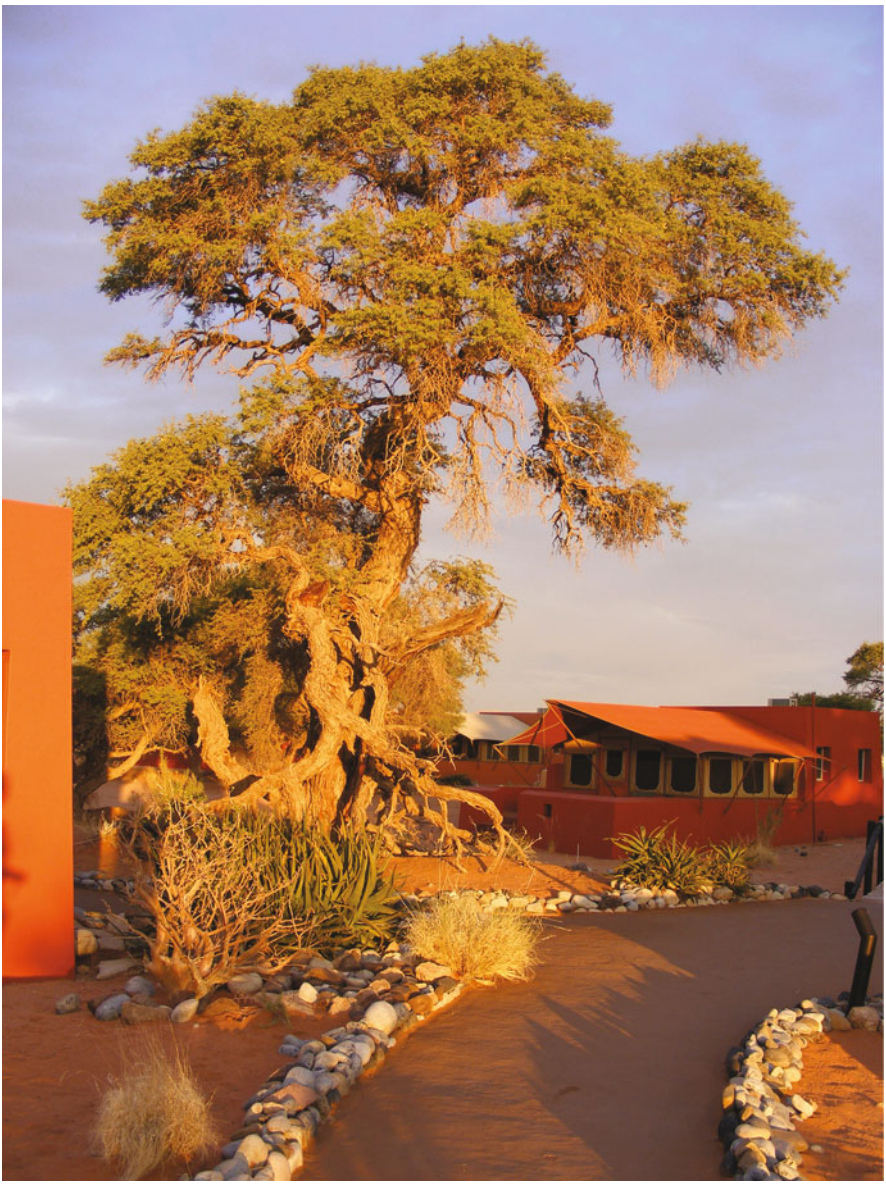
We are indebted to AVENG and to F. Titi for their partnership in the Sossusvlei conference. We acknowledge the visionary insights of Roger Jardine, Fani Titi, and Kim Heller.

Much has changed since Sir John Herschel, observing from the Cape in 1834–1838, demarcated the first known bar – in the Large Magellanic Cloud – which he described as an "axis of light." From 12 persons in 1660, we present the latest observations and theoretical perspectives in the year 2010 in the pages to follow.

Editorial note: Sentences with direct reference to K.C.F. in the preface were written by DLB.

Johannesburg, South Africa — David L. Block
Canberra, Australia — Kenneth C. Freeman
Puebla, México — Ivânio Puerari

Front cover – caption: Are we viewing a planet? Are we possibly imaging Mars? The answer is no – Jacqueline Riffault-Silk has captured an extraordinary photograph of a dune from the air. A section of the dune is bathed in sunlight, while the other sector lies in the shade. The unusual arc of the windswept ridge of the dune gives the false impression of a neighbouring reddish planet. A veritable terrestrial mask in the Namib Desert. Photo: Jacqueline Riffault-Silk.



Contents

Preface v
David L. Block, Kenneth C. Freeman, and Ivânio Puerari

A Tribute to Ken Freeman xix
David L. Block

Stars are Small Dark-Coloured Things That Live in Holes in the Ground 1
Pippa Skotnes

Shrouds of the Night – Galaxies and René Magritte 23
David L. Block, Kenneth C. Freeman, and Ivânio Puerari

Twin Masks of Spiral Structure? A Local Perspective 45
Thomas Y. Steiman-Cameron

The Mask of Complexity in Disk Galaxies 59
Daniel Pfenniger

Cosmic Magnetic Fields – An Overview 67
Richard Wielebinski and Rainer Beck

The Gaseous Halo Mask 87
Mary E. Putman, M. Ryan Joung, Jana Grcevich, and Fabian Heitsch

Molecular Gas Properties of Galaxies: The SMA CO(2-1) B0DEGA Legacy Project 97
Daniel Espada, S. Martin, P.-Y. Hsieh, P.T.P. Ho, S. Matsushita, Lourdes Verdes-Montenegro, J. Sabater, Simon Verley, M. Krips, and V. Espigares

The DiVA's Mask: Iconifying Galaxies and Revealing HI Anomalies 105
Jayanne English, Jason Fiege, Theresa Wiegert, Baerbel Koribalski, Wolfgang Kerzendorf, and Kenneth C. Freeman

Enigmatic Masks of Cosmic Dust: Lessons from Nearby Galaxies Through the Eyes of the Spitzer Space Telescope 113
Robert Groess, David L. Block, and Giovanni G. Fazio

The Large Magellanic Cloud: A Power Spectral Analysis of Spitzer Images 121
Ivânio Puerari, David L. Block, Bruce G. Elmegreen, and Frédéric Bournaud

Light Cores Behind Dark Masks 129
Ruben J. Díaz, Damián Mast, Germán Gimeno, Horacio Dottori, Irapuan Rodrigues, María Paz Agüero, and Peter Pessev

Globalization, Open Access Publishing, and the Disappearance of Print: Threat or Opportunity? 139
J.J. Blom

Super Star Clusters and Supernovae in Interacting LIRGs Unmasked by NIR Adaptive Optics 147
Petri Väisänen, Zara Randriamanakoto, Erkki Kankare, Seppo Mattila, and Stuart Ryder

Structure, Mass, and Stability of Galactic Disks 153
Pieter van der Kruit

What Can the Radial Surface Brightness Profiles of Galaxy Discs Tell Us About Their Evolution? 169
John E. Beckman, Leonel Gutiérrez, Peter Erwin, Ruyman Azzollini, and Inma Martínez-Valpuesta

The Complex Interplay of Dust and Star Light in Spiral Galaxy Discs ... 187
Maarten Baes, Dimitri Gadotti, Joris Verstappen, Ilse De Looze, Jacopo Fritz, Edgardo Vidal Pérez, and Marko Stalevski

Galaxy Morphology Revealed By SDSS: Blue Elliptical Galaxies 195
Hong Bae Ann

Rings and Bars: Unmasking Secular Evolution of Galaxies 201
Johan H. Knapen

Bars and Bulges Through Masks of Time 221
Isabel Pérez, Patricia Sánchez-Blázquez, A. Zurita, G. Popping, Bard K. Gibson, and Pierre Ocvirk

Tidal Trails and Mass-Segregated Isothermal Clusters 233
Donald Lynden-Bell

Stellar Debris Streams: New Probes of Galactic Structure and Formation 247
Carl J. Grillmair

Chemical Enrichment in Galaxies: Constraints on Nucleogenesis and Galaxy Evolution 261
Francesca Matteucci

Chemodynamical Simulations of Galaxies 277
Chiaki Kobayashi

Elemental Abundance Patterns of Disk Substructure 293
Gayandhi De Silva

Searching for Structures and Streams in the Extended Solar Neighbourhood with RAVE 305
Mary Williams and The RAVE Collaboration

On the Age–Metallicity–Velocity Relation in the Nearby Disk Using the RAVE Survey 313
Borja Anguiano, Kenneth C. Freeman, Matthias Steinmetz, and Elizabeth Wylie de Boer

The HERMES Project: Reconstructing Galaxy Formation 319
Kenneth C. Freeman

Stellar Halos: Unmasking a Galaxy's History 327
Amina Helmi

The Outer Halos of Elliptical Galaxies 339
Ortwin Gerhard

Galaxies: Lighthouses in the Shoals of Dark Halos 347
R. Brent Tully

Dark Haloes as Seen with Gravitational Lensing 361
Konrad Kuijken

Behind the Mask: Resolving the Core–Cusp Problem in Spiral Galaxies . . 373
George Rhee

A GALAXY BASELINE: Multiwavelength Study of a Sample of the Most Isolated Galaxies in the Local Universe 379
Lourdes Verdes-Montenegro, J. Sulentic, G. Bergond, Daniel Espada, S. Leon, U. Lisenfeld, V. Martinez-Badenes, J.E. Ruiz, J. Sabater, and Simon Verley

Diffuse Light and Galaxy Interactions in the Core of Nearby Clusters . . . 385
Magda Arnaboldi

Feedback in Star and Galaxy Formation 399
Joseph Silk

When Bad Masks Turn Good 409
Roberto G. Abraham

Spitzer's View of Galaxies in the High-Redshift Universe 425
Giovanni G. Fazio

Bandshifting and Other Masks of the Clumpy Populations in High-Redshift Galaxies 437
Bruce G. Elmegreen and Debra Meloy Elmegreen

Supernovae, Dust, and Cosmology 451
Brian P. Schmidt and Joerg Fischera

Endpiece – On Location in Dead Vlei 463

List of Participants 471

CODA 475
David L. Block

A Tribute to Ken Freeman

David L. Block

Quietness at the Walter Sisulu Botanical Gardens near Johannesburg, South Africa, where Ken Freeman and myself were steeped in thought.

As Ken cast his eyes to the future, he highlighted some of the major questions and problems to hopefully be solved in the future.

Meanwhile, black eagles, *Aquila verreauxii*, soared loftily above us in the African skies.

Allow me to pay a brief tribute to Ken, in the context of masks; specifically, galaxies and their masks.

Over the years, Ken (Fig. 1) has focused his research on several masks and he continues to play a pioneering role in each of these areas. There is the enigmatic mass mask. The actual mass of stars in the disk of a spiral galaxy such as our Milky Way or the Andromeda Spiral is only a small fraction of its total mass. It may not be widely known, but the first recognition of dark matter in spiral galaxies (as detected from the manner in which galaxies rotate) was made by Ken in 1970.

There is a fascinating historical interlude here, from the pen of Princeton astrophysicist, the late John Bahcall. The letter is dated December 23, 1982 and addressed to Bart Bok, then in Arizona. In it, Bahcall cites a paper published in 1970 (ApJ, 160, 811) in which we find these words, written by Ken:

> . . . there must be undetected matter beyond the optical extent of NGC 300 . . . For NGC 300 and M33, the 21-cm data give turnover points near the photometric outer edges of these systems . . . there must be in these galaxies additional matter which is undetected . . . Its mass must be as least as large as the mass of the detected galaxy . . .

In another letter, Bahcall comments that this Freeman 1970 paper *"is the earliest explicit recognition of the problem that I know about from rotation curves."*

A thought for the future, which Ken has often shared with me. Imagine if astronomers could construct a special telescope, which could directly image how matter in a spiral galaxy is distributed. Such a telescope would image dark matter

D.L. Block (✉)
School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa
e-mail: igalaxy@iafrica.com

Fig. 1 Ken Freeman outside the building which formerly housed the 30-in. "Reynolds telescope" in Canberra. Photo: David Block

as well as luminous matter. Galaxies would look supremely different through such mass-detecting telescopes, for their dark matter halos would be imaged too!

Dark matter itself is a mask: Although it does not actually obscure any stars in galaxies, it is a mask in the sense that its presence means that astronomers are afforded a misleading picture as to how the total mass in spiral galaxies is distributed.

In the words of Johann Wolfgang von Goethe,

"Wo viel Licht is, ist starker Schatten" or translated,

"Where there is much light, the shadows are deepest."

Another favourite mask of Ken Freeman is the dynamical mask of a galaxy. Pertaining to the formation of galaxies: What happens during the settling process of the gas itself? What does the birthing process of galaxies actually reveal? After all, a galaxy becomes quite different from what it was before. Has it not changed from a rather chaotic system into a well-ordered disk?

A crucial point is that a lot of information is lost during the settling (or birthing) process of a galaxy. By examining stars in the disk of a galaxy now, it is almost impossible to learn much about the properties of the galaxy before the disk formed. One can think of this loss of information as the dynamical mask.

A spiral galaxy may present a rather well-ordered dynamical face to us now, with the stars moving mostly in near-circular orbits. Yet behind this mask lies a history untold – the chaos which invariably initially reigned supreme before the disk settled. Most of the fossil history of spiral galaxies has forever been lost in the processes which led to the formation of galactic disks. If one wishes to recover information about a galaxy in its early history, one needs to focus upon chemical fossils or signatures rather than dynamical signatures. The way in which stars move today obscure most of the dynamical fossil information.

To retrieve that information, chemical tagging, much like human beings are genetically tagged by their DNA, is required. To quote Ken:

> We would like to reconstruct the ancient star-forming aggregates of the thick disk, which have largely dispersed by phase-mixing... Chemical tagging relies on similarities in the detailed chemical abundances of thick disk stars ([Fe/H], [α/Fe], r– and s– process elements) to tag them to common ancient star-forming aggregates ... The detailed abundance pattern reflects the chemical evolution of the gas from which the aggregates formed...

Among Ken Freeman's future goals is the chemical identification of stellar fossils in our Milky Way Galaxy. Chemical archaeology – of which Ken is a Pioneer with a capital P – will allow astronomers to penetrate the dynamical mask and facilitate the unveiling of a primordial cosmic landscape of our River of the Night – the *Via Lactea* or Milky Way. Galactic archaeology in its finest sense. Such a landscape will not have been created by paint, but by the distinguishing footprints or signatures of the chemical elements of the stars. The painting will unfold in the years to come, as Ken, who serves as Principal Investigator of the HERMES project, leads the way forward.

The above are but two examples in which the sands of astronomical research bear the footprints of Ken Freeman. At this gathering of astronomers from around our planet, to celebrate the 70th birthday of this most remarkable scientist, we address some of the key challenges and enigmas in understanding galaxies and their masks.

Birthday greetings to Ken flowed in from everywhere. Ken had two supervisors for his Ph.D.: Leon Mestel and Donald Lynden-Bell. It was a great joy to have Donald present at the conference; unfortunately Leon could not attend.

Ken's Ph.D. thesis was entitled:

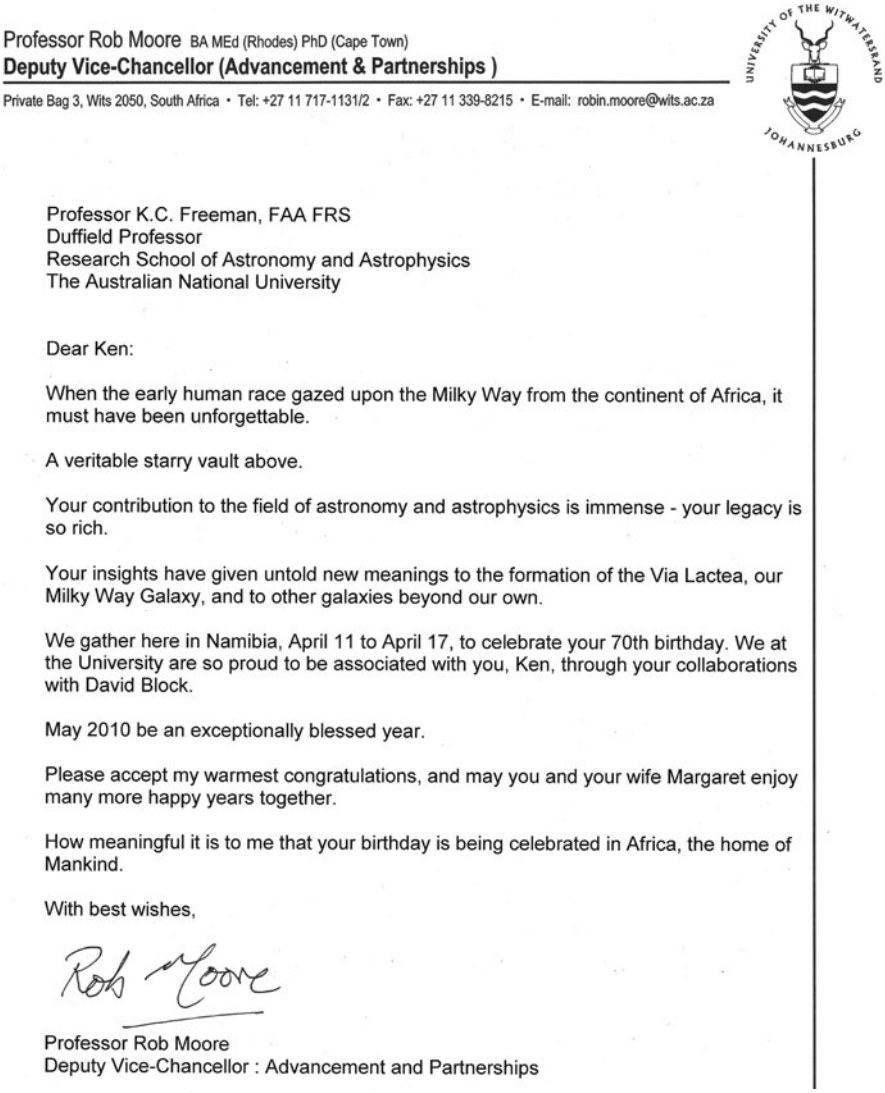
"On the Structure and Evolution of Barred Galaxies" and was submitted to the University of Cambridge (Trinity College) in 1965.

Leon Mestel sent Ken the following note of congratulations on his 70th birthday:

> The decades rush by – but life would be unbearable, if one were not fully occupied with pushing back the frontiers. I have watched with gratification and pride your steep, monotonically upward trajectory, taking in Trinity College High Table and the Royal Society en route. Having you as a research student has been one of my main contributions to the advance of Astronomy, both world-wide, but also in the land of my birth. Well done, Ken!

> And I am sure that, like others known to you – who would willingly let you have a decade or so – you are already demonstrating that 'retirement' is a highly notional concept.

Our Deputy Vice-Chancellor Rob Moore deeply regretted that he could not attend Ken's birthday fest, but wrote a special letter for Ken, reproduced in Fig. 2.



Professor Rob Moore BA MEd (Rhodes) PhD (Cape Town)
Deputy Vice-Chancellor (Advancement & Partnerships)
Private Bag 3, Wits 2050, South Africa • Tel: +27 11 717-1131/2 • Fax: +27 11 339-8215 • E-mail: robin.moore@wits.ac.za

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Professor K.C. Freeman, FAA FRS
Duffield Professor
Research School of Astronomy and Astrophysics
The Australian National University

Dear Ken:

When the early human race gazed upon the Milky Way from the continent of Africa, it must have been unforgettable.

A veritable starry vault above.

Your contribution to the field of astronomy and astrophysics is immense - your legacy is so rich.

Your insights have given untold new meanings to the formation of the Via Lactea, our Milky Way Galaxy, and to other galaxies beyond our own.

We gather here in Namibia, April 11 to April 17, to celebrate your 70th birthday. We at the University are so proud to be associated with you, Ken, through your collaborations with David Block.

May 2010 be an exceptionally blessed year.

Please accept my warmest congratulations, and may you and your wife Margaret enjoy many more happy years together.

How meaningful it is to me that your birthday is being celebrated in Africa, the home of Mankind.

With best wishes,

Rob Moore

Professor Rob Moore
Deputy Vice-Chancellor : Advancement and Partnerships

Fig. 2 Letter to Ken Freeman from Professor R. Moore, one of the Deputy Vice-Chancellors at the University of the Witwatersrand, Johannesburg, South Africa

Martin Rees had so much wanted to attend the conference, but his schedule as President of the Royal Society would not permit this. We had both met with Martin at the launch of our book in London in 2009 (Fig. 3). Martin's words to Ken are as follows:

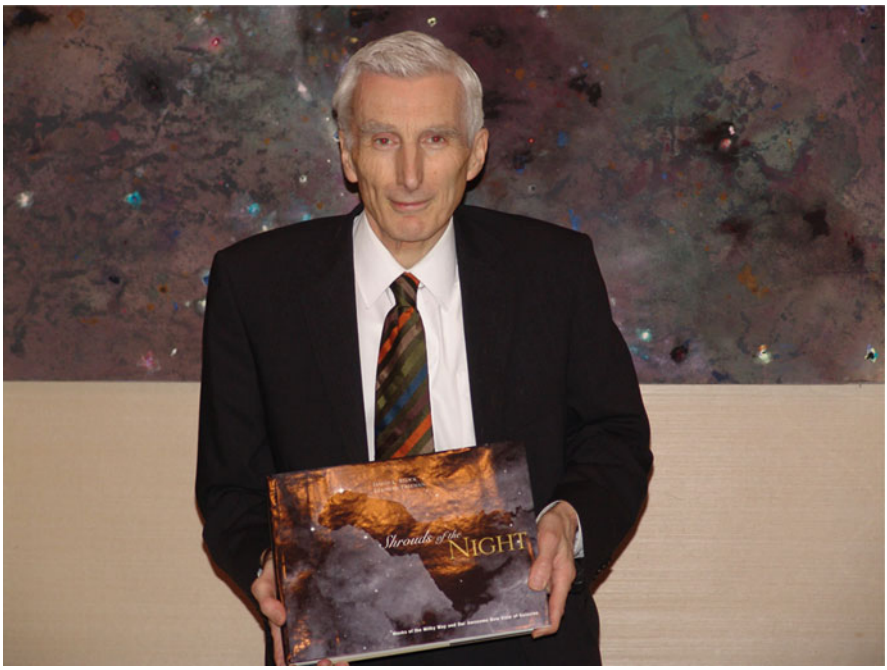


Fig. 3 Martin Rees at the launch of "'Shrouds of the Night" in London. Birthday wishes from Martin to Ken appear in the accompanying tribute. Photo: Robert Groess

> I first encountered Ken when I attended his lectures in Cambridge on galactic dynamics in the 1960s. Since that time, galaxies have been revealed as even more complex and fascinating than they seemed then . . . and Ken has been a leader in this great enterprise. As one of his students, as a colleague, and as a fellow alumnus of Trinity College, I send good wishes to Ken as he enters his years of maturity and discretion!

Ken has supervised 54 Ph.D. students, including 5 still working on their Ph.D. There have been 46 Ph.D. completions to date (3 gave up), and 5 of the 46 later became Hubble fellows. They are Hui Xiaohui, Heather Morrison, Yong-Ik Byun, Mary Putman, and Martin Bureau. The absolute and unflinching dedication of Ken to his Ph.D. students is but one of the distinguishing characteristics of this gentleman – allow me to repeat, gentleman – in the truest and fullest meaning of the word.

I asked Garth Illingworth, Ken's first Ph.D. student – *"numero uno"* – to share some thoughts about his then supervisor (1969–1973):

> It was with much regret that I was unable to attend the conference (teaching and project commitments). But I appreciate that David asked me to add a few comments to his tribute to Ken. At another celebration for Ken a decade ago on a little island off the coast of Queensland I realized that I was actually Ken's first graduate student. When that realization struck home I came to appreciate even more his willingness to gamble, and even to "leap in off the deep end". While those days as a student under Ken's guidance were certainly a key point in my career and development as a scientist, I wonder every now and then if Ken ever had second thoughts about taking on more students.. . . Obviously he overcame

any reluctance to do so! It is a great credit to Ken that his long and productive career has encompassed a remarkable series of important scientific contributions, and that his enthusiasm for astronomy and for a robust application of the scientific process has infected his many students as well.

Ken was a patient, very helpful supervisor. Ken taught me the value and the importance of thinking about the "big picture" aspects of one's scientific research. His theoretical background provided an excellent framework not only for interpreting the observations but also for helping him identify the important issues and questions in a particular scientific problem. His willingness to push the limits of the observational capabilities available to us was also a great lesson for me. Deriving velocity dispersions (and hence masses) in globular clusters in the 1970s was right at the limit of what could be done. It was Ken's enthusiasm and ability to take measured risks to push the image intensifier-fed photographic plates to their limits while observing at the Mt Stromlo 74-inch telescope Coude spectrograph that set me on the path to success on those challenging globular cluster observations. And that gamble finally got me a thesis! Ken's approach to science was a combination of rigorous methodology combined with a willingness to gamble on the implications of a result. It is not easy to balance such an approach, but it is key to making progress and to attracting attention to one's work and ideas. Ken's successes attest to his ability to find the right balance. I am looking forward to Ken's 80th celebration and I will make sure I attend that one!

I have studied a histogram giving the number of Nobel prize recipients in physics in the last 20 years versus their h-index. Ken's h-factor is much higher than that of many Nobel laureates in physics over the past 20 years! He remains one of the most highly cited astronomers on our planet, and his seminal contributions to the field have been recognized worldwide, not least in 1998, when he was elected FRS.

One of Ken's favourite hobbies lies in the world of ornithology (Fig. 4). Ken also appreciates music. He appreciates art. He introduced me to Aboriginal X-ray paintings. As a little birthday present to Ken, I thought of the masks of the San, as portrayed in the work of George Stow. On my desk lay a copy of *'Unconquerable Spirit – George Stows's History Paintings of the San'* authored by Pippa Skotnes, Professor of Fine Art and Director of the Centre for Curating the Archive at the Michaelis School of Fine Art. The world of the San people in Namibia and in South Africa – and our world of galaxies – are inextricably linked: through the theme of masks.

The paintings of the San as seen through the eyes and hands of George Stow had fascinated me for a number of years. One of my favourite plates is his "Dance of women in masquerading dresses of Buck-heads and Porcupine quills" (see Fig. 5, which appears as plate 13 in the book containing his masterful works: *Rock-Paintings in South Africa*). According to Stow's San informant Dia!kwain, the "things which the people here have put on are caps which they have made for themselves of young gemsbok' heads."

The late Dorothea F. Bleek elaborates "The clothes worn by Bushmen consisted in skin aprons and 'karosses' or skin cloaks made from the hide of bucks . . . Caps of leather with the fur on, head-dresses of hyenas' tails, or the plumage of some birds, were also worn. The whole of a bird's or buck's head as assumed for purposes of the chase, for dances and ceremonies.. . ."

In the context of a person under the skin of a large antelope, an extraordinary photograph was shown to me by my colleague and friend, Professor Francis Thackeray;

Fig. 4 Ken Freeman and the world of ornithology, as captured by caricature artist C. Brown

it is dated circa 1934, secured by W.H.C. Taylor, and is reproduced here (Fig. 6). The person is described as a *buckjumper*.

San folklore contains fascinating stories of the stars; I am reminded of the moving story of a young San girl who gently threw up wood ashes into the sky, to become the Milky Way; I am reminded of other stories, too, such as a porcupine on the hunting ground, watching those celestial bodies signifying the approach of day: "He [the porcupine] returns home; for, he is used to look at these Stars; they are those which he watches; while he feels that the dawn's Stars they are." An ever changing world of dance: the dance of earth and sky. The classic work *Bushman Folklore* by the late W.H. Bleek and L.C. Lloyd is a veritable vault of folklore coming from the San: the oldest known lineage of modern human.

Fig. 5 "Rocks in the mountain to the N.W. of the Komani near Queenstown. 24th April, 1867. G.W.S. " Stow wrote of this work, "Dance of women in masquerading dresses of Buck-heads and Porcupine quills" and which, according to his San informant, shows caps "which they have made for themselves of young gemsboks' heads." Photo: Iziko South African Museum collection

Why not approach Professor Skotnes (Fig. 7) to fly to Namibia to address our delegates and their spouses/partners at a pre-dinner lecture? I did so, and she immediately agreed. Not only did she focus her attention on George Stow, but she also introduced us to two young !kun boys who came from Namibia to South Africa (Fig. 8). Her contribution *Stars are small dark-coloured things that live in holes in the ground* was received by Ken and Margaret with the highest of praise. At Ken's request, the contribution by Professor Pippa Skotnes appears first in this volume.

Concluding thoughts. . .

Ken, it has been an enormous sense of inspiration collaborating with you over a grand number of years. We have written papers together. We have shared many meals together. We have visited each others homes. Our wives Margaret and Liz have shared many special moments together (Fig. 9). You have visited the University at which I work, on several occasions. We have lectured at Harvard together, and the memories of both of us being hosted by Lord and Lady Rosse at Birr Castle are forever etched in my mind. What a singular joy to co-author the book *Shrouds of the Night* with you. May Margaret and yourself (Fig. 10) share many more happy years together.

How well do I remember your enthusiasm at a theatrical play in Boston featuring the life of Galileo Galilei. The play was performed during the International Year of Astronomy. Galileo beheld two books: the *Book of Nature* and the *Book of Scripture* (Fig. 11). He believed both had the same author (Fig. 12).

Fig. 6 An extraordinary photograph (circa 1934) of a person under the skin of a large antelope. Masks serve as a common theme in both the terrestrial and the astronomical worlds. An ingenious way of hunting would be for the hunter to masquerade as the animal being hunted. Photographed by W.H.C. Taylor at Logageng (also known as Logagani), north of the Wonderwerk Cave on the southern border of the Kalahari desert. A detailed exploration of these ideas follows in the contribution by Skotnes (chapter "Stars Are Small Dark-Coloured Things that Live in Holes in the Ground," this volume). Photo: McGregor Museum, Kimberley, South Africa

It was Galileo who penned these words:

> Philosophy is written in this grand book, the universe, which stands continually open to our gaze. But the book cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wonders around in a dark labyrinth.

In this, the 350th anniversary of the founding of the Royal Society, of which you are a Fellow, I have the following words:

> Ken, you soar like an eagle, but you are as gentle as a dove. Your roots are etched in the soil of a rare breed – a very rare breed – of family. The Grand Family. You typify 'La Grande Famille' painted in 1947 by the Belgian surrealist artist René Magritte.

Fig. 7 Margaret and Ken Freeman, together with world-renowned expert on the world of the San, Professor Pippa Skotnes. As a birthday gift, Skotnes presented Ken and Margaret with a copy of a watercolor entitled "The Star Country." The watercolor (drawn in 1881) was produced by a young !kun boy, formerly from the land of our conference: Namibia. Photo: Bruce Elmegreen

Fig. 8 A portrait of the !kun boy Tamme, who was captured and traveled along the trade routes of Namibia by foot, wagon, and then by ship to the Cape of Good Hope. Tamme features prominently in the birthday lecture for Ken, entitled "Stars are small dark-coloured things that live in holes in the ground" by Skotnes (Chapter "Stars Are Small Dark-Coloured Things that Live in Holes in the Ground," this volume). Tamme and his friend !nanni produced magnifient watercolors of the microcosm and the macrocosm. Their world of art was a dance of earth and sky. Photo: Centre for Curating the Archive and LLAREC

Fig. 9 En route to some of the highest dunes on our planet: from left to right, Liz Block, Joe Silk, and Margaret Freeman. Photo: David Block

Fig. 10 Ken and Margaret Freeman at the Bush Dinner near Sossusvlei. Photo: Lourdes Verdes-Montenegro

Fig. 11 Galileo beheld two books: the *Book of Nature* and the *Book of Scripture*. He believed that both had the same author. Reproduced here is a magnificent leaf (dated 1526) of William Tyndale's translation of the New Testament into English, from the original Greek. The first sentence on this leaf reads "In the begynnynge was that worde . . ."

Fig. 12 Penetrating "The Shroud of God" could be the most challenging mask of all. "Gloria in Profundis" or "Glory to God in the lowest" (G.K. Chesterton). A response to Deus Absconditus: The Wager of Blaise Pascal. Image courtesy: Ismar David Collection, RIT Cary Graphic Arts Collection

Stars are Small Dark-Coloured Things That Live in Holes in the Ground

Pippa Skotnes

Abstract In this chapter, attention is drawn to some indigenous stories about disguise and transformation, taken from this magnificent country Namibia and from southern Africa as a whole, with a focus on a complex of ideas about the stars. Astonishingly, these stories suggest a meeting of minds from different worlds – Ken Freeman's and that of the San or Bushmen 130 years ago. San stories about the stars contribute to our understanding of our relationship to the universe in a similar way to that which Ken Freeman's revelations of masking and shrouding help understand the universe itself. The cosmology of the San universe construed at times through the agency of masked and shrouded beings known as therianthropes was partly illuminated by the work of George Stow (the father of San rock art research in southern Africa), two !kun San boys, a group of |xam San and a remarkable woman called Lucy Lloyd.

1 Introduction

When David Block asked me to make a presentation to this gathering, I was, all at once, delighted, perplexed, intimidated, intrigued. I read a little bit about Professor Ken Freeman and began to understand why it is that so many interesting scholars from all over the world have gathered to honour him. But I know next to nothing about dark matter, globular clusters, stellar debris streams, outer halos of elliptical galaxies or galactic archaeology. The list of paper titles reads, to me, like a poetry of an alien world – word conjuring an unknown yet resonant universe of ideas. But it is this strangeness that is my very connection to this conference, for I am deeply intrigued by what I imagine must be a kind of astronomical sense of awe in the face of the immensity of what is around us, an immersion not only in science but also in the limits of humanunderstanding.

P. Skotnes (✉)
Centre for Curating the Archive, University of Cape Town, 7701, South Africa
e-mail: pippa.skotnes@uct.ac.za

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_1,

David Block, as a hunter-gatherer of extraordinary ideas from a vast range of subjects had seen a book I had produced on the early geologist and father of San rock art research in southern Africa, George Stow. One of the intriguing things about Stow's marvellous paintings, and which I think immediately appealed to David is that amidst the beautifully painted animals are humans apparently cloaked in animal heads. I think that for David, as was the case later for me when I read the fabulous book *Shrouds of the Night*, the image of these part-human part-animal beings resonated with ideas about galactic masks and shrouds. One of the seminal contributions Ken Freeman has made is to stretch the limits of our understanding of the universe by announcing the presence of dark matter as a fundamental component of galaxies. Dark matter thus forms a veritable set of masks in a shrouded sky. Both Ken Freeman and David, in their joint publication *Shrouds of the Night*, showed how the brilliance and passion of individuals can shape knowledge in astonishing ways, and that a diversity of ideas of disguise and transformation in many places can expand our understanding not only of the universe but of ourselves and our place in the cosmos.

What I want to do today is to draw your attention to some indigenous stories about disguise and transformation, taken from this magnificent country Namibia and from southern Africa as a whole, with a focus on a complex of ideas about the stars. Astonishingly, these stories suggest a meeting of minds from different worlds – that of Freeman's and that of the San or Bushmen's 130 years ago. I would argue that the San stories about the stars contribute to our understanding of our relationship to the universe in a similar way to that which Ken Freeman's revelations of masking and shrouding help understand the universe itself. The cosmology of the San universe – construed at times through the agency of masked and shrouded beings known as therianthropes – was partly illuminated by the work of George Stow, two !kun San boys, a group of |xam San and a remarkable woman called Lucy Lloyd (Fig. 1). Bear with me, dear audience, while I take you back into their world described 130 years ago, and I hope you will see, as I do, not a totally alien interruption to your days of talk about astronomical phenomenon, but a meeting of different but sympathetic minds.

2 !nanni, Tamme and George Stow

This conference takes place in an extraordinary part of the world, characterised by vast open spaces, huge protected reserves and an enormous coast line where, in places, the footprints of desert lions can be tracked to the shore line and the carcasses of sea animals and wrecks of ships. The night sky is almost unparalleled. Namibia covers an area of just over 8,00,000 km^2 and is flanked by two major deserts – the Namib along the west and the Kalahari in the east. Less than 1% of the land is arable. The population, estimated at just over 2 million, comprises a majority of Ovambo but with smaller groups including Herero (Fig. 2), Damara and Nama as well as the Bushmen or San, once the only human inhabitants and who have had the longest unbroken occupation of the country of all.

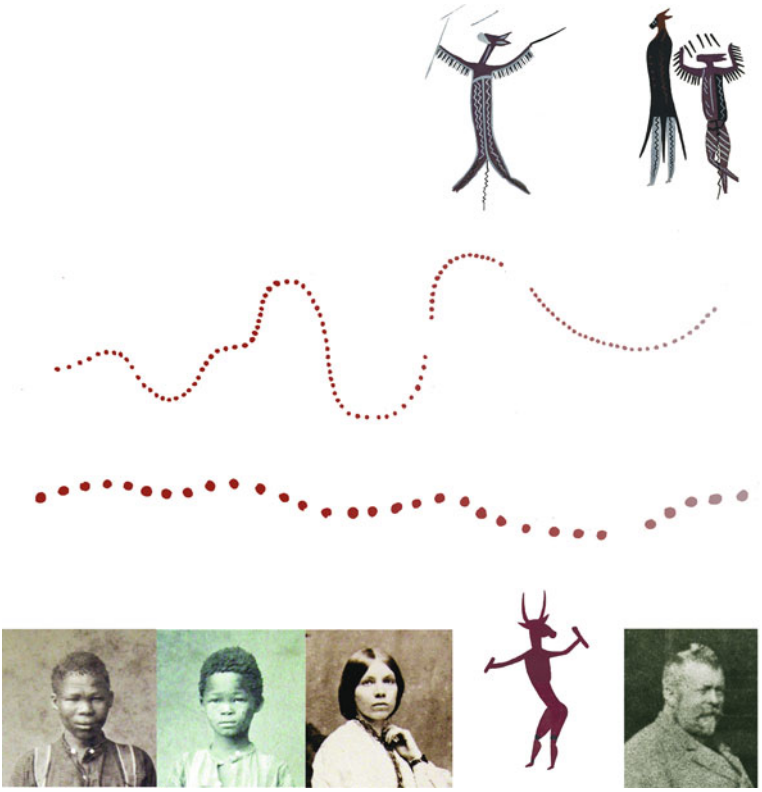


Fig. 1 The *lower panel* shows (from *left* to *right*) !nanni, Tamme, Lucy Lloyd and George Stow, who are the focus of this chapter. The *painted figures* were copied by George Stow, from rock paintings in southern Africa

The story of the title of this talk was first recorded on the 25 June 1881 in a home in colonial Mowbray, Cape Town. The story teller was !nanni, assisted by his friend Tamme, both of whom had reached that place after an epic journey of almost a year in which they had travelled along the dusty trade routes of Namibia by foot and by wagon from the northern border with Angola to Walvis Bay and then by ship to the Cape of Good Hope. !nanni and Tamme spoke !kun, one of a cluster of dialects spoken by the San or Bushmen in the region and came from a group little known or studied even today. Their country was partly wooded and relatively well watered, and it was rich with edible roots and flowering plants that produced berries and fruits of many kinds. At night, !nanni and Tamme would hear the yapping of jackals and the whooping of hyaenas. Lions, known to be afraid of the sun, would growl from the distant abandoned grass huts they occupied, and owls were heard calling in the surrounding trees. The sparks from the little campfires were mirrored in the dark sky, where the moon's mother (the upper pointer to the Southern Cross) would signal the deepening night. In the day the boys would play with homemade

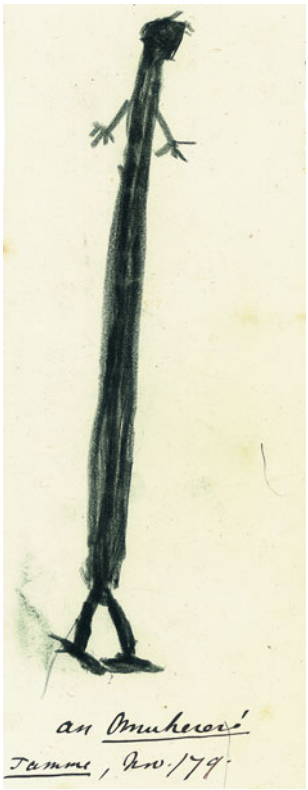


Fig. 2 Tamme's drawing of a Herero man made in November 1879

toys, while singing the song of the ≠na≠n'arro chameleon. They would pick the pods of sweet smelling shrubs to make a !ne (a musical instrument) or tap gum from the !yu tree. The skins of monkeys, genets and weasels were tanned for front and back aprons. Orchids flowered in !nanni's father's hunting grounds and the cooing of the wood pigeons at the water's edge was a sound of home.

!nanni and Tamme's neighbours were the Makoba who lived nearby in a village at the edge of a large waterway. The Makoba had boats and dogs that they would sometimes have to eat when the season was dry and food sources failed. They made pots which they would trade with !nanni's people and also kept cattle. In exchange for a cow or tobacco – perhaps cannabis too – the Makoba would demand from !nanni's parents a regular supply of veldkos and, most importantly, elephant tusks.

This was the 1870s and the height of the ivory trade. African elephants were being slaughtered almost to extinction in the east and west of the continent to supply ivory to the international market. In the year that !nanni and Tamme told their story of the stars between 50 and 80,000 elephants were killed to furnish a global industry with knife handles, billiard balls, combs and the keys of pianos. Indeed, the supply of upright pianos to Germany and the USA in a single year alone required over

1,00,000 kg of ivory. While much of the trade was continued north and east of the country of the boys, the American whalers and sealers losing profits along the coast turned their attention to ivory and helped deplete Namibian and Angolan herds of elephants as well.[1] !nanni and Tamme (Fig. 3) told many stories about elephants whose hearts alone could feed a family for 5 days. Some stories observe the family structures of the herds in the areas, others tell of the dying of an aunt, trampled to death by elephants or describe the trade negotiations with the Makoba for tusks.

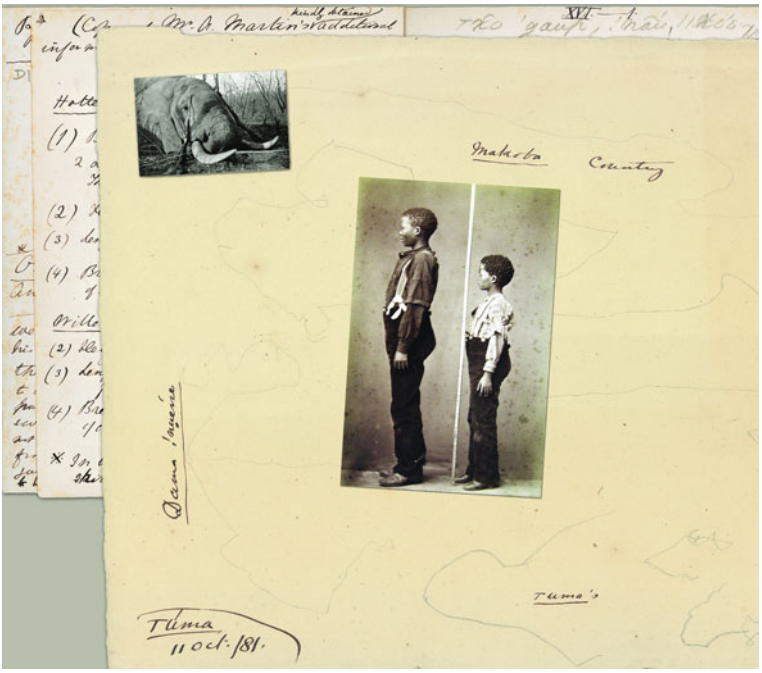


Fig. 3 A studio photograph of !nanni and Tamme, on top of a map of Makoba country drawn by another of the !kun boys, |uma, on 11 October, 1881

At the same time as !kun men were hunting elephants for the ivory trade on the northern border with Angola, the Cape Colony in the south was experiencing an acute labour shortage. The end of slavery and the expanding capitalist economy had left the farms and the colonial households under-supplied with workers. Cape Commissioner Coates Palgrave, deeply familiar with the possibilities in central Namibia and willing to exploit them, instigated a programme of labour recruitment that resulted in several hundred individuals, mainly Damara, being displaced from their homes and shipped to Cape Town. These included women and children and while the recruits mostly came from central Namibia, trader networks with the Makoba meant some came from further north too.

[1] See Robert Gordon (1992) quoted in Skotnes (2007) 324–5.

One day when Tamme was on the bank of the waterway near his home the Makoba arrived in their boats and took him away. He was afraid of the Makoba who were known for violent acts. They had captured several other boys, throwing some to the crocodiles and had killed their parents who tried to protect them.[2] Tamme was then sold to the Ovambo and, along with !nanni who had similarly been abducted, tracked his way down the sandy trade roads with his masters to Walvis Bay where both boys were shipped to Cape Town. After they arrived in the colonial city they were placed at a home in Mowbray belonging to Jemima Bleek who lived there with her sisters and her five daughters. One of her sisters, Lucy Lloyd, was well known to be a scholar and expert on the San.

!nanni's story was about the nature of the stars (Figs. 4 and 5). *"Stars,"* he told Lucy Lloyd, *"are small dark-coloured things that live in holes in the ground. They have small arms but have not horns."* In the morning, the star burrows into an earthen hole, afraid of the sun, but at dusk the star ascends into the sky. In the day, the boys would hunt for stars that had fallen in the night, and see their burrows on the horizon. *"The star falls to the earth,"* !nanni said, *"and, we take it up; we look at it, throwing it away, for it falls down, dying."* The boys would play with the star, they would kill it, but the star was already dying. *"Stars are dark-coloured, little and black,"* said !nanni, *"and, we kill it; we do not merely set it down but throw it away."*[3]

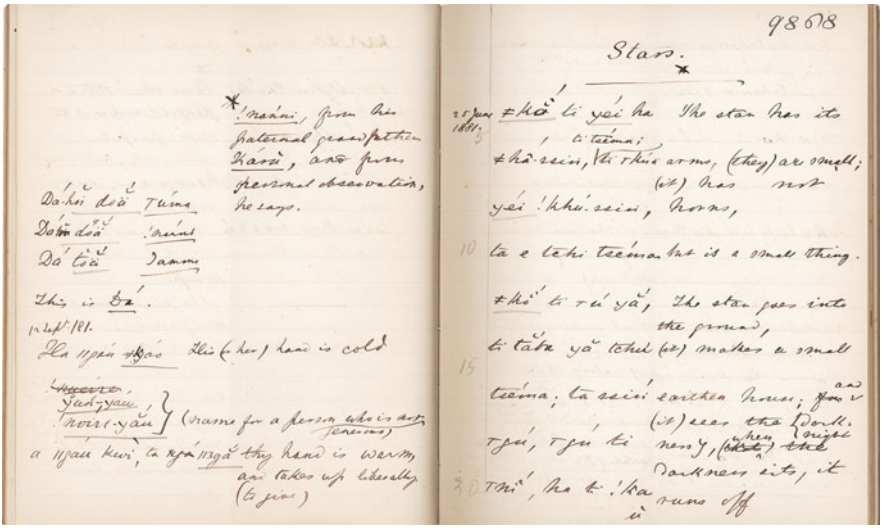


Fig. 4 Page 9868 of Lucy Lloyd's !kun notebooks, the first page of an entry about stars told to her by !nanni on 25 June 1871

[2] The story of two of these boys is described in Skotnes (2007) 236–71.

[3] The whole |xam and !kun archive of texts is included in DVD in Skotnes (2007) and can be found at http://lloydbleekcollection.cs.uct.ac.za

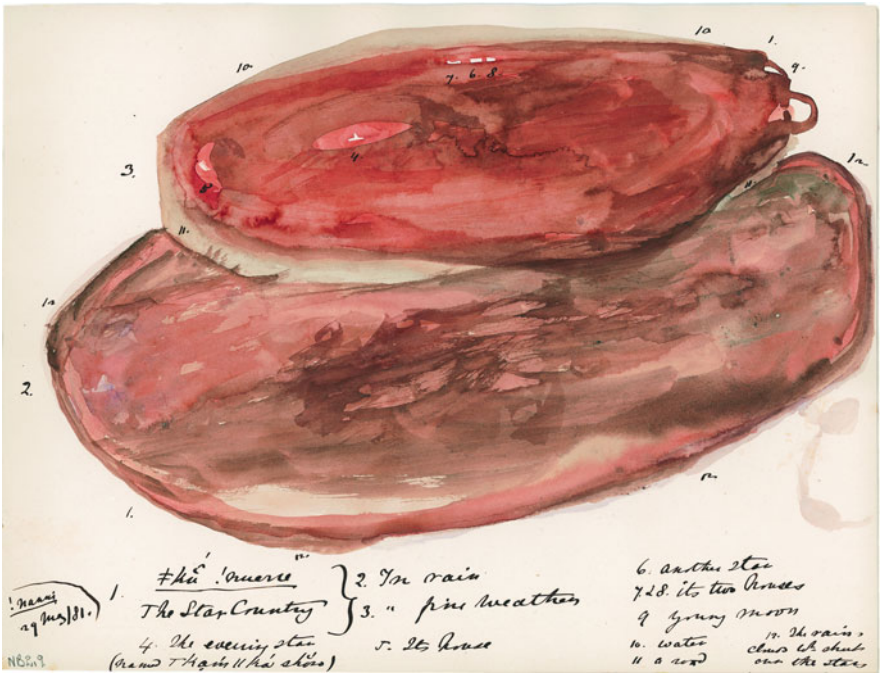


Fig. 5 The Star Country, drawn by !nanni in May 1881. This painting, made for Lucy Lloyd, includes the evening star and its house, an area of rain and of fine weather, another star and its two houses, the young moon, water and a road and the rain's clouds which shut out the stars

At once strange and unfamiliar, this counterintuitive view of the stars was part of a whole complex of San ideas about life and death, game and hunting, sorcerers and transformation. In this universe of ideas, a community of therianthropes constantly penetrated the shroud that separated the living from the dead. These penetrations are described in many stories and pictured in the San art of southern Africa.

At the time that !nanni and Tamme were living at Mowbray, Lucy Lloyd was working with a remarkable collection of rock paintings made by George Stow, a medical student drop-out turned geologist who had come to the Eastern Cape in South Africa from England in the 1840s. In the course of searching the landscape for fossils and composing poetry about the miraculous nature of creation (and later mapping the geology of a vast tract of land between the Orange and Vaal Rivers) Stow had become obsessed with the rock art of the country and the scattered bands of San who still survived colonial expansion. After serving with British forces in the Seventh Frontier War, Stow left Port Elizabeth and travelled into the interior. Against great odds and with a zealous passion Stow copied paintings at several hundred rock art sites, carrying his paintings across thousands of miles of countryside.[4]

[4] See Skotnes (2008).

His was the earliest systematic study of rock art in southern Africa. He identified what he thought to be the key themes in the art, writing:

> The paintings are not only of different degrees of excellence – but are also of different kinds. Some may have a mythological character but others are certainly historical paintings – and not only represent some particular events but also are, from the marked character and diversity of figure, intended for portraits of various individuals belonging to the several tribes. These are generally representations of battles – huntings and dances – in which the Bushmen are represented wearing the different disguises they are known to adopt on such occasions but which to European eyes when they see representations of men enveloped in skins with Buck's or Bird's heads or with tails of iguanas and other animals fastened round their waists, make them believe they must have a mythological, or fabulous meaning – instead of representing the manners and customs of the Bushmen as they really were.

Stow also recorded interviews with remaining San individuals who had escaped the carnage of the frontier wars and the commandos that had been established in the 18th and 19th centuries to exterminate them. He visited massacre sites and wrote with a painful passion about the tragic fate of the San and their way of life:

> Many caves are to be found in this mountainous region, several of them of immense size . . . There is every evidence that at one time densely populated centres were sprinkled throughout the whole of these mountain glens . . . But if the extent and number of the caves and paintings contained in them make known the numerous clans which once occupied these picturesque glens, and the surprising degree of excellence at which some of their leading artists arrived, so also do these spots proclaim in an equally unmistakable manner the tragic fate which befell their former inhabitants; they tell us but too plainly of the infernal storm of lead which was poured in upon them by their vindictive and remorseless pursuers.
>
> The sides of the great cave of the White Rhinoceros and Serpent, in a rocky ravine on the right bank of the Washbank Spruit, are so thickly bespattered with hundreds of the bullet marks of the assailants, that one could almost write an account of its siege and point out where in their desperate struggle the intrepid defenders were forced back from point to point, where they from time to time turned at bay in their attempts to keep back their enemies, and where, behind a great heap of piled rocks at the end of the cave, they turned for the last time, overpowered but unsubdued, and resolutely continued the conflict until the shout and the turmoil closed with the final discharge of musketry, in the silence of death.[5]

For George Stow (Fig. 6) the paintings of the San were inextricably tied up with their fate. He was convinced that the San were the only true indigenous people in southern Africa and that their paintings were not only the evidence of their long occupation but also, as he phrased it, their title deeds. He saw their long battle against the invading British, Dutch and various African herders and farmers as a patriotic struggle to defend their home and country, and their actions in defence of their hunting grounds as acts of great heroism. He understood their paintings not only to reflect this valiant resistance but also to reveal a way of life depicted with brilliance and imaginative genius.

George Stow had been corresponding with Lucy Lloyd and after his premature death in 1882 she bought almost 200 of his copies of rock paintings.[6] Several had been at the Mowbray home from the mid-1870s but the bulk arrived after 1882

[5] See Skotnes (2008) and Stow's own publication (1905).

[6] See Schoeman (1997).

Fig. 6 One of the few extant photographs of George Stow, the father of San rock art research in southern Africa

when the !kun boys were in residence. The house in Mowbray (Fig. 7) was an extraordinary place. It had been bought by Wilhelm Bleek (Fig. 8), a scholar well connected with the European intelligentsia and a trained philologist who, having first worked on a Zulu dictionary in the Natal Colony, had relocated to Cape Town (and Mowbray) to work in the library of the governor Sir George Grey.[7] Here he had met Jemima Lloyd and after a courtship by correspondence had proposed to her and

[7] For the story of the Bleek and Lloyd archive, see Skotnes (2007), also Deacon and Dowson (1996).

Fig. 7 The house at Mowbray where Bleek and Lloyd lived with the |*xam* prisoners and the !kun boys

Fig. 8 Wilhelm Bleek (*left*) and Lucy Lloyd (*right*) photographed in Cape Town in the 1870s

married her a year later in St George's Cathedral in the city. Her sister Lucy (Fig. 8) arrived from Natal for the wedding (having been shipwrecked on the way) and later relocated to live with the couple. When Bleek became aware of a group of |xam Bushmen arrested on the northern Cape frontier and serving prison sentences at the Breakwater Convict Station in Cape Town he was keen to learn their language – a language virtually unrecorded – and create a |xam dictionary. He tried to interview these prisoners arrested, for the most part, for stock theft, some part of gangs that for

several years had evaded capture, one convicted of culpable homicide, but the prison was a grim environment and the convicts were taciturn. Bleek applied to the governor for some of the prisoners to be released into his custody and his request was granted. Over a period of time, Bleek gained the trust of these men, and the result was the opening of a floodgate of stories that spoke of a life of an extraordinarily rich connection to the land and the longing to return to it. The creation of the dictionary became the focus of the last part of Bleek's short life, and his sister-in-law Lucy Lloyd, possessed of a more acute ear, assisted him. Together, they recorded thousands of pages of |xam texts, translated into English, and stories which gave a view into hunter-gatherer life in a frontier zone unparalleled anywhere on earth. Bleek had a special interest in the stars and engaged her Majesty's astronomer at the Cape, Sir Thomas Maclear, to help identify the stars that the |xam San discussed. Bleek and Lloyd's archive has become a national treasure, a wellspring of ideas about and knowledge of a way of life that has vanished and a source for the construction of the new South African identity. Indeed the motto of South Africa is inscribed in |xam on the country's coat of arms, despite the fact that not a single living person speaks the language.

In the middle of the project to record San languages, Bleek died suddenly one night of the consumption that had plagued him all his married life. Jemima and Lucy, whose two sisters had also joined them, were left with five young daughters, one just a baby, and several San convicts in the home in Mowbray. Lacking support and adequate income they nevertheless continued the work that Bleek had begun, recording many thousand more pages and adding to the dictionary and concordance. Furthermore, the policy to bring labour from Namibia into the Cape in the late 1870s provided an opportunity for Lucy Lloyd to learn yet another San language and thus it was that !nanni and Tamme were placed in her care in September 1879.

Over the decade and a half that Lucy Lloyd listened to and transcribed |xam and !kun stories, she heard many accounts of the transformation of beings, the interpenetration of the world of the spirits with the world of the living, and of the Early Times in which all animals, the wind, the sun, the moon and the stars were people. At that time, in the very beginning, the southern African landscape was empty of animals and the dark sky was unpunctured by the light of stars. Only people inhabited the country, the ancestors of the San, making their way in a dim glow that emanated from the armpit of an old man who, in later, brighter days, would become the sun. These Early People embodied a principle of transformation, the prototypes of all that would come to symbolise change and transubstantiation – shrouding and disclosure – for although the first people were people, they were also animals before animals lost their humanity and within them were the qualities their animal heirs would one day embody. In this gloomy world, |kaggen, later known as Mantis, was the foolhardy master, young maidens possessed the power to transform a man into a tree and to create the stars, the rain beast could change a family into frogs and the features of the landscape were not only sites of stories but also themselves had agency.

In time, the Early Times gave way to a new order in which animals lost their humanity, the sun and moon were ordered to walk the sky, the hare was forced to

scamper on the earth with a split lip because she questioned the moon's ability to restore life and each animal mated with its own kind. But the Early Times were still present, and the spirit of transformation was alive in the world. Sentience belonged to the stars and the rain, and in the pools at night, the stars and all things could be seen including the loveliest of flowers which were once young girls transformed into the water's wives. Stories, too, were not mere moments of a storyteller's performance, they literally lived on the wind.

In many of the stories Lloyd's narrators conveyed a sense of storytelling around the campfire, where the snapping of burnt kindling, the sparks of light and the reflected flash of a predator's eye beyond the camp were an earthly mirror of celestial activity. Stories of the stars seemed to exist at that threshold between daylight and darkness, and between the living and the dead, a place in which the San were alert to the dangers of the spirit world and the beasts of the night. *"Brother, brother,"* said a |xam man, *"Why is it that the stars are shining above me? For it was the lion's eyes that shone as they approached him, for there were two lions."*[8]

In the early days, a young woman – the first maiden – had created the Milky Way by tossing wood-ashes into the sky. Scented roots of different kinds had burnt in the fire of the first maiden. Red roots – the young ones – made the red stars and the older white roots made the white stars. In the later days, the stars in Orion's Belt were named for the hunter who startled the springbok, the hunter who drove them forward and the hunter who lay in wait for them. Indeed the Milky Way, said the |xam San, resembled the massive herds of springbok which, like the stars drifting across the sky, were never still and moved tirelessly across the landscape.

As they sat around the campfire, the |xam could hear the sounds of the stars, and in the early times the Great Star !gaunu had named each star in song. In the later times, the San would watch for shooting stars, for these were the signals of the sorcerers' work and of the death of a friend. A star falling made the sound of the dying heart, for a shooting star would fall away from its usual place in the cosmos, going elsewhere, just like the heart no longer beating the rhythm of life. The magicians and sorcerers were the one's whose hearts, after death, then once again became stars but when a sorcerer died, he took his thoughts of the living with him and continued to exert his influence over the living. These sorcerers and sorcerer-stars hovered both above the campfire and at the edge of darkness around the camp and could snatch away a life with the shooting of invisible arrows.

One of the stories that Lucy Lloyd heard was related by a man named Dia!kwain (Fig. 9). It was a story told by a father, worried about the precariousness of life on the colonial frontier, to his son when, at any time, he could become orphaned and forced to fend for himself. *"Father taught me about the stars,"* said Dia!kwain, *"he used to say to me, that if I were sitting waiting by a porcupine's hole, I must watch the stars . . . watch the place at which the stars fall, for this is where the porcupine is."* The porcupine (Fig. 10) was alert to danger, it listened for death with its ears

[8] From the story: Men who hunted lions with bones by |han≠kass'o. http://lloydbleekcollection.cs.uct.ac.za

Fig. 9 A portrait photograph of Dia!kwain, who provided Bleek and Lloyd with almost 170 stories in the collection of over 130 |xam notebooks

and its nose, and it watched the stars and knew, with its thinking strings, to turn back as the Milky Way turned back and the Dawn's Heart Star arose. Bats would also, at times, inhabit the porcupine's hole at the place where the stars fell, and they too were signals of when the porcupine would return to the waiting hunter.[9]

The lore of the stars and the sidereal stories of the |xam and the !kun that were told to Lucy Lloyd offered ways to speak of both the origin of things, the fear of death, a relationship with home and country and to negotiate the changing life of the colonial frontier. In both the earth-bound world (Fig. 11) and in the stars, the San invested the lore and the history of their people. Nothing was quite as it seemed and little stayed the same. The world was protean and the features of the landscape were sites of meaning. In the telling and the hearing of their stories the shroud of ignorance that cloaked the world of the San (Fig. 12) began to give way to a vision

[9] The story is published in Bleek and Lloyd (1911) 246–53.

Fig. 10 A drawing of footprints of the porcupine around its burrow by another of the |xam informants, |han≠kass'o, made on 4 September, 1878

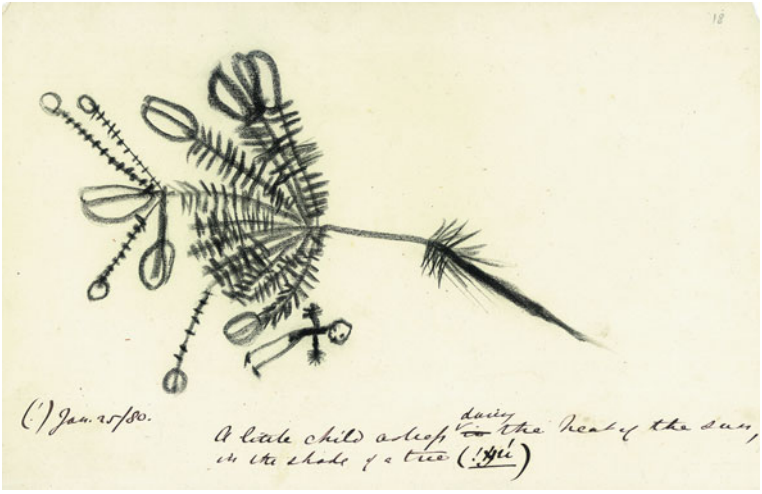


Fig. 11 A little child asleep during the heat of the day in the shade of the !yu tree, drawn by !nanni on 25 January, 1881

of lives lived in a web of connection with the environment, and the world of both the physical and the spiritual. In Lloyd's notebooks, we begin to understand that when the |xam saw the stars, they saw at once the wood ashes of the first maiden, the springbok hunters, the signals that guided their nocturnal prey and the spirits of the dead. For the !kun boys, the stars were both a means to influence the living and an embodiment of things that died. Stars died as they fell to earth – but then they also appeared again to return to their celestial perambulation. For !nanni, the star in

Fig. 12 A detail of a cloaked and transformed figure from one of George Stow's copies which included several figures of this nature, in various stages of transformation

its day-time burrow was a dead thing, a dark-coloured, hornless being to be dug up and disposed of, not a star at all. Yet its presence in the following night sky was a marker of endurance, survival and continuity (Fig. 5).

This shifting form of things is best imagined in !kun thought through a shape-shifting, therianthropic being named by !nanni and Tamme as |xue (Fig. 13). |xue was a Bushman, and !nanni had apparently seen his house and his child called |xum-ta when he was with the Herero. |xue took on many identities. As himself he was a Bushman who wore a front apron of a monkey's head with the bones still inside and a back apron of four monkey's feet. He wore a monkey's tail wound round his head. In his many transformations he resembled a lizard one day, a kid (|ou) the next, a little tsema and a ||ha and a python (≠ni) all on different days. When |xue died he resembled the |ou (and, unlike the dying star, had horns). |xue became a lion and spoke to lions (who understood him) and a dead elephant. |xue became water in the shadow of a tree and the wood pigeons at the edge of a pool. He became an Omuherero crying of thirst and a fungus on a tree of which men were afraid. Once he was a little fly and a black butterfly, later a dead owl on which flies feasted. |xue was a spirit who called on the darkness and murdered a child. He inhabited the trees and became the sweet fruits on which the people feasted. He stalked prey wearing the head and body of a dead ostrich and became a biting ant that stung his wife. At sunset, and at the time the stars appeared in the sky |xue would become himself again and lie down to sleep.

|xue was a master of change and transformation (Fig. 13). He could engage with the world as himself, and he could become invisible in the landscape – shrouded in

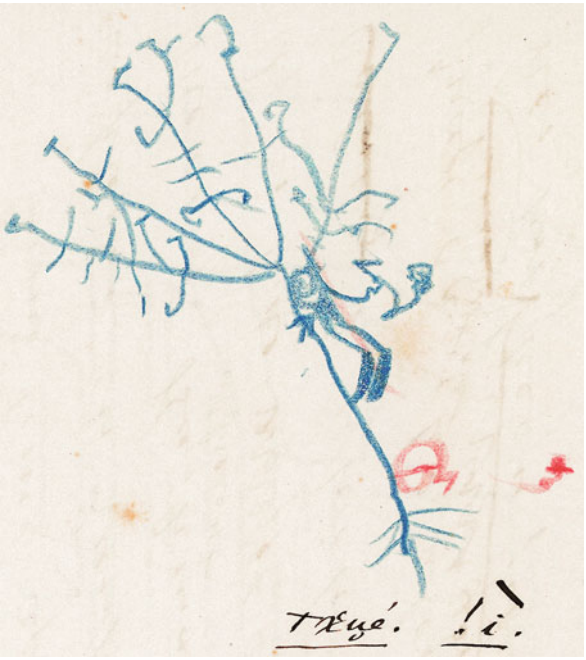


Fig. 13 A drawing by !nanni made on 27 April 1880, depicting the shapeshifter |xue as a tree

it – and so disguised effect and influence his relationship to it. |xue was a hunter, a parent, a child. He was powerful and weak, brave and afraid, vindictive and forgiving. As a hunter he was part of the landscape, invisible to his prey,[10] as a parent he was loving and cruel and as a child he was full of longing, fear and grief. |xue was in many ways the embodiment of what it meant to be a !kun boy in the violent and changing world in which !nanni and Tamme grew up. He was a way of being at once at one with and at odds with, but always a part of, all that was around him.

For the San, being an empathetic part of the world was also a means to survival. In the ritual dances performed all over southern Africa and reflected in some paintings, at which the dancer collapsed into a state of trance, he or she would take on the identity of the eland, feeling its body, its hair standing on end or the lion, feeling cornered in a hunt. The dead would inhabit the body of an owl or a jackal and visit the sleeping form of a loved one. Rain sorcerers would capture the Rain Bull, kill it and drag its body over the land, its blood attracting water and bringing rain to the dry country.[11] Or the dancer's blood shed from his nose and the sweat from his

[10] I am grateful for this insight into Marlene Winberg who interviewed Kapilolo Mahongo, a storyteller speaking a related !kun dialect who said that although he did not know the stories of |xue the word meant "hunter" – a person who had to be invisible in the landscape.

[11] For a discussion of the Rain Bull in rock art and San thought, as well as perspectives on interpretation of rock art see Anne Solomon (2006, 2007, 2008).

body would be rubbed on the sick as a healing ritual.[12] Hunting too, was an act of transformation. The |xam hunter stalking the springbok would literally feel the black stripe along his side and the place on his head where the horns emerged. He would become the springbok or the eland, and stagger as the animal died, mimicking its gait and mirroring its collapse. This immersion in the being of the prey enabled the hunter to use his own form to determine the outcome of the hunt and to avoid the forces that might assist the recovery of his prey.

If we return to the paintings and copies of George Stow, we see in his selection these therianthropes and this rich and deeply immersed world of transformation – and in the spirit of David Block's and Ken Freeman's metaphor – shrouding and masking (Fig. 14). There are paintings that represent the healing dances of blood, where the blood of the dancers is used to heal the sick. There are images of the spirit people, the transformed dead who have taken on the form of an eland, the animal heads and hooves, shrouding their sorcerer's identity or a lion. Some human figures appear frog-like or striped, and others with the sweat of the trance streaming from their armpits. There are paintings that represent the rain animal, a creature resembling a bull whose young are described as striped like a quagga, and who is sometimes horned, sometimes hippo-like, or with long tail and white teeth. This beast, called !khwa by the |xam, is a water-being of great power, transformed into the embodiment of the rain. Other paintings depict human figures with

Fig. 14 Another of Stow's copies, this one creating a sense of other-worldliness, from the banks of the Zwart Kei River

[12] Stow describes the dance of blood in Stow (1905). The so-called trance hypothesis of David Lewis-Williams suggests all rock art is related to trance experiences. This thesis has been very controversial in the study of southern African rock painting. See Lewis-Williams (1981 and 2006), Skotnes (1994), Le Quellec (2004) and Solomon (2008), for example.

horned caps, kaross-cloaked or with elongated legs, suggesting a feeling of bodily distortion experienced by the dancer who collapses into a trance state. Others still reveal the complex interactions of dancers, hunters, shamans or sorcerers and their stock-keeping neighbours. In one painting framed by Stow so as to emphasise its circular composition two large oxen occupy the central part of the image (Fig. 15). Around them are painted baboons, a leaping feline, skin aprons and small buck. Rising into the centre of the image are three white figures, an animal-headed man with spear and stick, a headless figure with skin chest-band, and a tall figure with hugely elongated legs, rising into the centre of the composition. The skin aprons and colours of the men are identified with the colours of the cattle and the distorted and hybrid figures suggest that only through the processes of transformation and the conjuring of other identities can the San compete with the changing world brought about by the settlement of the country by stock-keepers and farmers.

Fig. 15 George Stow's copy of a painting in which he sees the intersection of the San world with that of intruding stock-keeping peoples

In another painting (Fig. 14), from the banks of the Zwart Kei River, Stow's copy has given us the impression of a ghostly spirit world in which eland, hartebeest and other buck hover about and among two groups of transformed figures. One group – hunters – wear karosses and hold their sticks and arrow bags, another is horned and cloaked in dark skins and seems to emerge from the grey rock. A stooped figure in front of them suggests a dancer, falling into unconsciousness, or a figure from the spirit world. A later visitor to the site from which this painting came said that there were several layers of painting still faintly visible underneath these figures. They included a large yellow animal and under that a huge faded red snake, each layer shrouding the layer beneath, and indeed the rock surface itself appearing to

Fig. 16 The controversial copy of a putative ostrich hunt by George Stow, which Stow said was copied from rocks in the Wittebergen Reserve

be a shroud, lifted for a moment, for this vision of another world to be made partly visible[13].

There is one painting, a controversial one (Fig. 16), in which Stow has copied an image that represents an ostrich hunt. His daughter had described how he had found it in an inaccessible part of the Witteberg mountains but it has never been found again. In it a group of ostriches have turned to gaze at the approach of another, which is really a hunter in disguise, his legs and bow protruding from the enormous mask of the ostrich head, neck and body. In fact, this appears to be one of the very few paintings that could, without doubt, be described as a hunting scene in which the hunter disguised himself as his prey, and those rock art scholars – the orthodoxy today – who believe that San rock art in general has nothing to do with hunting, are at pains to try to prove the painting a fake. But for me this strategy represents a fundamental misprision. In the first place, it misunderstands George Stow and the complexity of his project and his approach to the art he copied, and in the second place it fails to recognise that hybrid identities, transformations and the shrouding of both hunter, trancing dancer and figures of the spirit world make up a deeply interrelated complex of ideas that are all potentially resident in the rock paintings.

3 Concluding Thoughts

To end, let us return to the stars. Imagine, for a bit, the coming dusk. All around the desert absorbs the scarlet of the setting sun, and then the dunes are bleached and the sky is drained of colour. It is 2 days from the new moon, so soon the landscape will disappear and turn to blackness and the sky will become punctured

[13] Lewis-Williams has productively described the rock surface as a veil. See Lewis-Williams (1990).

with starlight. Imagine the small !kun boys on the trek south along the trade road with their masters, and a |xam family, gathered around a campfire watching the sparks of the embers drift up to meet the heavens. The night will witness the death of a relative, the fall of a sorcerer's heart, the image of springbok moving in vast herds across the sky, the flicker of the eyes of lions on the edge of darkness and the call of the hammerkop. Stories of the early people will explain the place of the stars and, in the dead of night, a star will fall to earth and die, a dark-coloured thing in a small earthen burrow. For George Stow, out in the landscape for so many months making his copies of paintings (Figs. 17 and 18), he must have felt a deep bond with

Fig. 17 George Stow's copy of a painting of a rain animal, from a farm near the Orange River

Fig. 18 Details from two separate paintings copied by George Stow in which human figures are depicted in a transformed state: a bird-headed figure, possibly carrying a fly whisk on the left and a buck-headed figure which was part of a dance of similar figures. The paintings are, today, very badly preserved

the San as he watched the Milky Way and looked out for the bright stars Sirius and Canopus. He may even have felt he saw, just beyond the light of his own fire, the shape shifting figures of hunters, and transformed dancers and spirits moving into the darkness.

Acknowledgments All works by George Stow are from the Iziko South African Museum collection. All other images reproduced in this chapter are from the Digital Bleek and Lloyd Collections held at the CCA from originals held by Iziko, the National Library and the University of Cape Town.

References

1. Bleek, W.H.I. 1873, Report of Dr Bleek Concerning His Researches Into the Bushman Language and Customs. Presented to the Honourable the House of Assembly by Command of His Excellency the Governor, House of Assembly, Cape Town
2. Bleek, W.H.I. 1875, Second Report Concerning Bushman Researches, with a Short Account of Bushman Folklore. Presented to Both Houses of Parliament, by Command of His Excellency the Governor, Saul Solomon & Co, Cape Town
3. Bleek, W.H.I., Lloyd, L.C. 1911, Specimens of Bushman Folklore, George Allen, London
4. Deacon, J., Dowson, T. 1996, Voices from the Past: |xam Bushmen and the Bleek and Lloyd Collection, Witwatersrand University Press, Johannesburg
5. Gordon, R. 1992, The Bushman Myth: The Making of a Namibian Underclass, Westview Press, Boulder
6. Lewis-Williams, J.D. 1981, Believing and Seeing: Symbolic Meanings in Southern San Rock Paintings, Academic Press, London
7. Lewis-Williams, J.D. 1990, Through the Veil: San Rock Paintings and the Rock Face. S Afr Archaeol Bull, 45(151), 5–16
8. Lewis-Williams, J.D. 2006, Debating Rock Art: Myth and Ritual, Theories and Facts. S Afr Archaeol Bull 61(183), 105–114
9. Lloyd, L.C. 1889, A Short Account of Further Bushman Material Collected. Third Report Concerning Bushman Researches, Presented to Both Houses of the Parliament of the Cape of Good Hope, by Command of His Excellency the Governor, David Nutt, London
10. Schoeman, K. 1997, A Debt of Gratitude: Lucy Lloyd and the Bushman Work of G.W. Stow. South African Library, Cape Town
11. Skotnes, P. 1994, The Visual as a Site of Meaning: San Parietal Painting and the Experience of Modern Art. In: Contested Images: Diversity in Southern African Rock Art, eds. T. Dowson, J.D. Lewis-Williams, Witwatersrand University Press, Johannesburg, pp. 315–329
12. Skotnes, P. 2007, Claim to the Country: The Archive of Wilhelm Bleek and Lucy Lloyd, Jacana Press/Ohio University Press, Cape Town/Athens, OH
13. Skotnes, P. 2008, Unconquerable Spirit: George Stow's History Paintings of the San, Jacana Press/Ohio University Press, Cape Town/Athens, OH
14. Solomon, A. 2006, Roots and Revolutions: A Critical Overview of Early and Late San Rock Art Research. Afrique et histoire, 2(6), 77–110
15. Solomon, A. 2007, Images, Words and Worlds: The |xam Testimonies and the Rock Arts of the Southern San. In: Claim to the Country: The Archive of Wilhelm Bleek and Lucy Lloyd, eds. P. Skotnes, Jacana Press/Ohio University Press, Cape Town/Athens, OH, pp. 148–59
16. Solomon, A. 2008, Myths, Making and Consciousness: Differences and Dynamics in San Rock Arts. Curr Anthropol, 49(1), 59–86
17. Stow, G.W. 1964 [1905], The Native Races of South Africa, Struik, Cape Town
18. Stow, G.W., Bleek, D.F. 1930, Rock Paintings in South Africa from Parts of the Eastern Province and the Orange Free State, Methuen, London

Shrouds of the Night – Galaxies and René Magritte

David L. Block, Kenneth C. Freeman, and Ivânio Puerari

Abstract In this review, we focus on a word which appears in the title of these conference proceedings: the term "galaxy". We highlight the treachery of photographic or digital images of galaxies, by introducing the work of the Belgian surrealist artist René Magritte. We discuss his famous painting "The Treachery of Images" (La trahison des images, 1929), which portrays a pipe. Below it, Magritte painted these words, "Ceci n'est pas une pipe" ("This is not a pipe"). The painting is not a physical pipe, but rather a representation of a pipe. And so it is with galaxies and their masks: what we see is only a representation of what actually constitutes a galaxy. The definition of a galaxy given in the Hubble Atlas beautifully portrays the treachery of photographic images of galaxies, being then defined as closed dynamical systems, or island universes.

1 Galaxies and René Magritte

The Treachery of Images ("La trahison des images") is a painting by the Belgian René Magritte. The picture shows a pipe. Below it, Magritte painted, "Ceci n'est pas une pipe" which, translated from the French, is "This is not a pipe." The painting is indeed not a pipe, but rather an image of a pipe, which was Magritte's point.

Magritte's work frequently displays a juxtaposition of ordinary objects in an unusual context, giving new meanings to familiar things. The representational use of objects as other than what they seem is typified in his painting *The Treachery of Images*, which shows a pipe that looks as though it is a model for a tobacco store advertisement (Fig. 1). Below the pipe appears these words "Ceci n'est pas une pipe" ("This is not a pipe"), which seems a contradiction, but is actually true: the painting is not a pipe, it is an image of a pipe.

In 1968, Michel Foucault wrote a book entitled *Ceci n'est pas une pipe* ("This Is Not a Pipe"), inspired by René Magritte's painting. What is so captivating about

D.L. Block (✉)
School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa
e-mail: igalaxy@iafrica.com; david.block@wits.ac.za

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_2,

Fig. 1 "La trahison des images" – The Treachery of Images, by René Magritte. "Ceci n'est pas une pipe" (This is not a pipe) provides the viewer with an apparent contradiction. An interpretation and discussion of this remarkable painting appears in Section 1. Photo credit: 2009 Museum Associates/ LACMA / Art Resource, NY. Reproduced with permission from DALRO, South Africa © 2010 on behalf of ADAGP, France

this work by Magritte is the highly realistic depiction of a pipe on the one hand and the legend that Magritte writes below it, which echoes these words loud and clear: "This is not a pipe."

Foucault [30] argues that "[neither words nor the visible] can be reduced to the other's terms: it is in vain that we say what we see; what we see never resides in what we say. And it is in vain that we attempt to show, by the use of images, metaphors, or similes, what we are saying . . ." As we cast our eyes carefully upon Fig. 1, we are gripped by a sense of surprise – a surprise which "inaugurates a play of transferences that run, proliferate, propagate, and correspond within the layout of the painting, affirming and representing nothing."

In the beginning of his book, Foucault sets the stage for the treachery of images – the curtain doors figuratively unfold, and we behold "a carefully drawn pipe, and underneath it (handwritten in a steady, painstaking, artificial script, a script from the convent, like that found heading the notebooks of schoolboys . . .), this note: 'This is not a pipe.' "

It is, as Foucault notes: "as simple as a page borrowed from a botanical manual: a figure and the text that names it. Nothing is easier to recognize than a pipe, drawn thus – our language knows it well in our place – than the 'name of a pipe.' "

Magritte's work is a brilliant interplay between all pervasive power of language and that of image. Do we not, as young children, conjure up images in our minds by means of linguistics and the act of naming?

"Magritte names his paintings in order to focus attention upon the very act of naming," Foucault writes. "And yet in this split and drifting space, strange bonds

are knit, there occur intrusions, brusque and destructive invasions, avalanches of images into the milieu of words, and verbal lightning flashes that streak and shatter the drawings."

The neurophysiology of vision and the centre stimulus. Certainty is lost. Foucault explains this in terms of "similitude" – where ideas are henceforth cast amidst a tumultuous ocean as it were, to "multiply of themselves, to be born from their own vapor and to rise endlessly into an ether where they refer to nothing more than themselves . . ."

No matter how vividly, in other words, the picture conjures the thought of a real pipe, the viewer is aware of the fact that it isn't a real pipe, bur rather an analogue of a real pipe.

The pipe so depicted (Fig. 1) does not "satisfy emotionally." When Magritte once was asked about this image, he replied that of course it was not a pipe, just try to fill it with tobacco.

And so it is with galaxies and their masks, or shrouds. What *is* a galaxy? Do we, at best, merely see representations of a galaxy, or do we see the galaxy itself? Without a futuristic telescope dubbed by us a "mass telescope" [15], the failure of current arrays to capture dynamically active dark matter haloes poignantly highlights the treachery of images in an astronomical context.

The Oxford English Dictionary gives the established definition of a galaxy as follows:

> "a luminous band or track, encircling the heavens . . .; the Milky Way", and states that the transferred and figurative use is "now chiefly applied to a brilliant assemblage or crowd of beautiful women or distinguished persons."

The definition becomes much more complex. For example, globular clusters such as ω Centauri may actually be nucleated dwarf ellipticals which have been tidally stripped [33]. In the context of Magritte, we may say of ω Centauri, *Ceci n'est pas un amas globulaire* being translated, "This is not a globular cluster" but an ancient dwarf galaxy.

A detailed criterion that separates dwarf galaxies from extreme globular clusters is needed to establish the divergence between actual observations and cosmological simulations. Many more dwarf galaxies have been predicted around the Milky Way, than is actually observed [37, 44].

2 The Treachery of Conventional Images of Galaxies

What *is* the morphology of the Sombrero Galaxy, for example? We all know its optical morphology from the *Hubble Atlas*. But that is only a representation of what there is. The deep images which follow in Figs. 2, 3, and 4 were made by David Malin stacking several plates, invariably all from the UK Schmidt Telescope. The represented wavelengths are those in a slightly broadened B-band: Kodak's IIIa-J emulsion, behind a Schott GG395 filter. The passband is thus 3950–5500 Å, closely equivalent to the SDSS g band. (Occasionally, David Malin used a GG385 filter, yielding a slightly wider passband). The limiting surface brightness in these images is $\sim$ 28mag arcsec^{-2}.

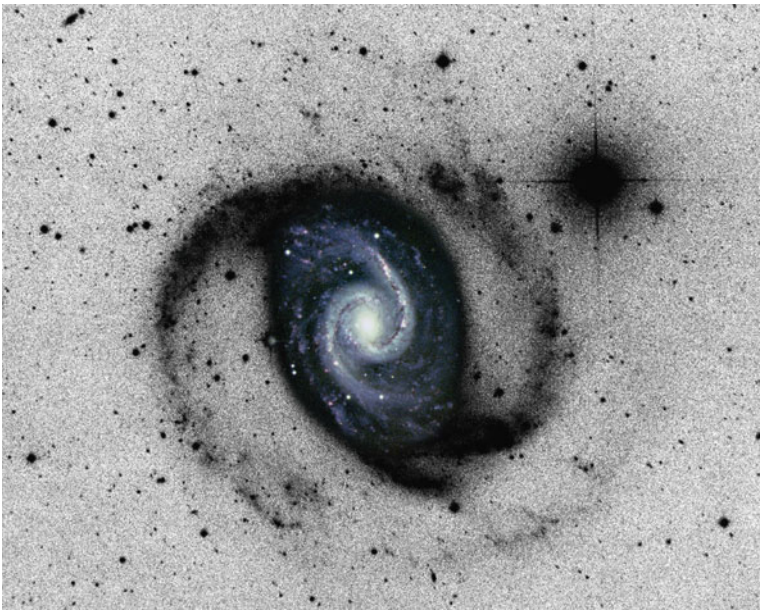


Fig. 2 A deep image of NGC 1566, reaching a surface brightness of ~ 28 mag arcsec^{-2}, produced by David Malin in the stacking of photographic glass plates. An optical photograph of the galaxy is overlayed, for comparative purposes. Although the dust content in the outer disk must be low and almost devoid of bright young OB stars, the morphology seen here remains too faint to be recorded by conventional infrared detectors. Photograph courtesy David Malin

Most of the light from stars in these deep Malin images comes from red giant branch (RGB) stars older than about 2 Gyr, with masses < 1.6 M$_\odot$. The stars near the tip of the RGB are K to M giants, depending on metallicity [Fe/H].

What is emphasized in astonishing detail is that none of these stellar haloes can be predicted from conventional images of these galaxies, shown as insets in Figs. 2, 3, and 4. *La trahison des images* looms large. What *is* a galaxy such as the Sombrero Messier 104 or Messier 83?

Halos of HI gas, too, can be very extended and show remarkable spiral arm structure in the gas. One is reminded of NGC 2915 (Fig. 5), where the HI envelope extends for over 22 disk scalelengths (see [34], and references therein). Prominent spiral arms in the outer HI envelopes of spiral galaxies have recently been the focus of attention by Bertin and Amorisco [10].

3 The Duality of Spiral Structure

> We're now looking at a transition to a possible change in the way we look at galaxies. Sometimes . . . we see disks that have a spiral structure that we couldn't have dreamt existed from looking at the optical picture . . . we've got a possibility here of applying the morphology to

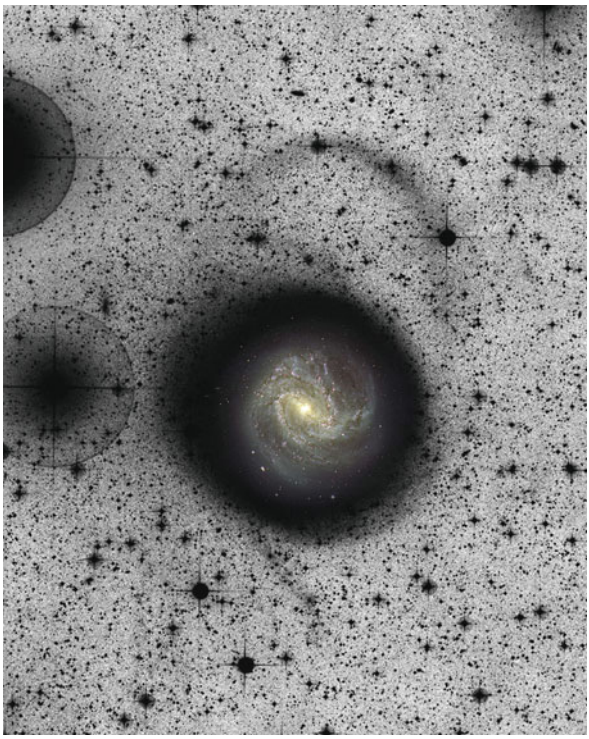


Fig. 3 A deep image of NGC 5236 (Messier 83) shows two remarkable outer "handles". A conventional, optical image of M83 is overlayed. Old stars in the faint outer disk must be responding to the dark matter halo in which the galaxy is embedded. Conversely, within the optical boundary, the dark matter distribution must be responding to the gravitational field of the stars. Photograph: David Malin

> a physical framework, perhaps in a way that none of us could have dreamt of before we had the capability of sweeping the dust away from the galaxy in a figurative sense. – R.J. Allen [2].

Optically thick dusty domains in galactic disks can completely camouflage or disguise underlying stellar structures. *Cosmic dust grains act as masks or shrouds* [15]. The dust masks obscure whether or not the dust lies in an actual screen or is well intermixed with the stars. The presence of dust and the morphology of a galaxy are inextricably intertwined: indeed, the morphology of a galaxy can completely change once the gaseous Population I disks of galaxies – the shrouds – are dust penetrated (e.g. [12, 16]).

The classification of galaxies has traditionally been inferred from photographs or CCD imaging shortward of the 1 μm window, where *stellar* disks are not yet dust penetrated. Images through an I (0.83 μm) filter can still suffer from attenuations by dust at the 50% level (e.g. [50]). The NICMOS and other near-infrared camera arrays offer unparalleled opportunities for deconvolving gaseous and stellar mor-

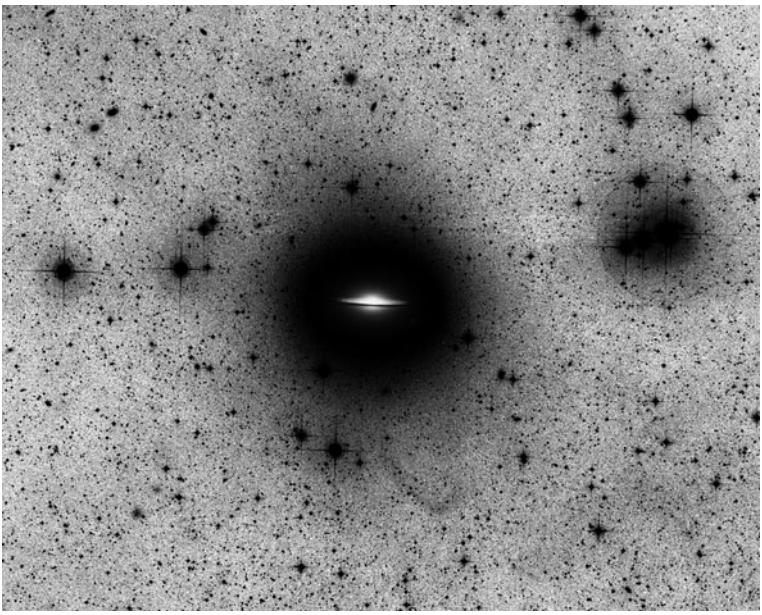


Fig. 4 A deep, photographically stacked image of the Sombrero Galaxy shows extraordinary detail in its outer domains. Halos of dark matter and of stars are dynamically "live" and interactions of stars in the outer disk and of dark matter must continuously be taking place. Photograph: David Malin

phologies, because the opacity at K (2.2 μm) – be it due to absorption or scattering – is always low. The extinction (absorption+scattering) optical depth at K is only 10% of that in the V-band [41].

Many years before the advent of large format near-infrared camera arrays, it became increasingly obvious from rotation curve analyses that optical Hubble type is not correlated with the evolved stellar morphology. This was already evident in the pioneering work of Zwicky [55] when he published his famous photographs showing the "smooth red arms" in M51. In the *Hubble Atlas* and other atlases showing optical images of galaxies, we are looking at masks. At the gas. Not the stars, to which the properties of rotation curves are inextricably tied. Burstein and Rubin [22] demarcate three principal types of mass distribution, with Hubble type a and b classes being found *among all three types* more or less equally. Clearly, one needs to probe what lies behind the dusty Shrouds of the Night.

The distribution of dust grains in the bulges and disks of galaxies may be very widespread as illustrated in the magnificent hand-drawn dust atlas by Lynds [40]. Combining optical and near-infrared images with radiative transfer models yields a 90% increase in dust masses determined by IRAS and betrays the presence of a population of cold dust grains (T $\sim$ 20 K and colder). Full details appear in Block et al. [17].

Furthermore, dust need not be confined to the disk itself. To cite but one striking example: the 2 kpc segment of bulge light scattered off the inner dust lane in NGC

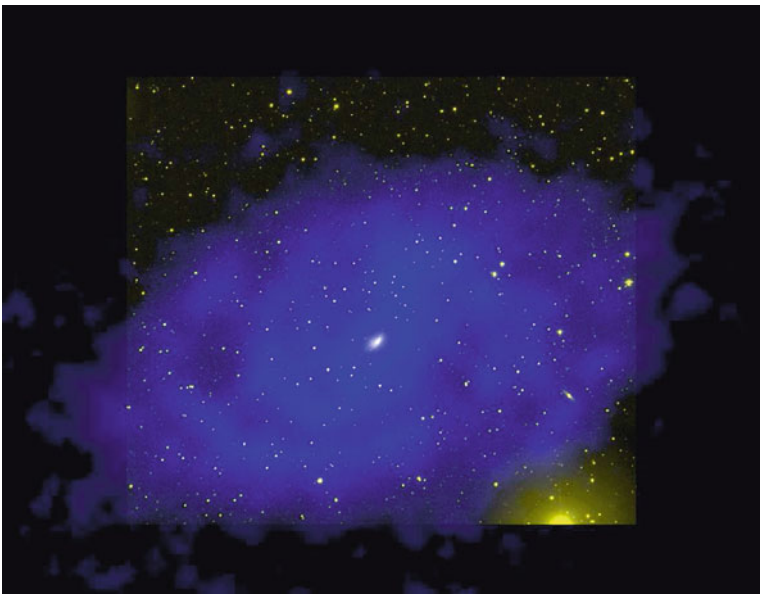


Fig. 5 The *blue* compact galaxy NGC 2915. This is a superposition of an optical image and an HI image from the ACTA. Foreground stars and the inner optical component of the galaxy appear *yellow*, while the HI distribution is *colour-coded blue*. The HI arms extend out to ~ 22 disk scale-lengths. The HI data is from Meurer et al. [42] and the image was kindly prepared by Gerhardt Meurer. See also Freeman [34] for further discussion

2841, forms the largest reflection nebula yet identified, and lies ~ 0.5 kpc above the plane of the disk of NGC 2841 (see [14]).

The tracing of dust grains in nature can be very elusive: high levels of dust extinction do *not* necessarily imply that the effects of dust attenuation (i.e. observed reductions in surface brightness profiles) will also be large, because scattering by dust grains may fill in at least part of the lost surface brightness. The effective albedo of dust grains in the near-infrared can actually be higher than in the optical (for a determination of the near-infrared albedo of dust grains in M51, see Fig. 4 in Block [11]).

Furthermore, large amounts of dust do not necessarily imply red V−K colours. The dust column density on the far side of an inclined spiral with relatively *blue* V−K colours can be just as high as on the near side, where the V−K colour distribution can be much redder [28].

From a dynamical viewpoint, the disk of a spiral galaxy can be separated into two distinct components: the young *gas-dominated* Population I disk and the older *star-dominated* Population I disk. (The usage of Population I and Population II has truly evolved since the 1957 Vatican Conference. The term Population II now would refer to a metal-weak halo population, whereas the metal-rich disk stars are classed as belonging to the old Population I component). The *gaseous* Population I component contains features of spiral structure (OB associations, HII regions, dust and cold

interstellar HI gas), which are naturally fast evolving; dynamically, it is very active and responsive, because, being characterized by small random motions (a "cool disk"), it fuels Jeans instability. In contrast, the stellar component in the Population I disk, which is dynamically "warmer" because of the larger epicycles, contains the *old population of stars*, betraying the underlying mass distribution – it is the 'backbone' of the galaxy [39].

Two notes of caution here:

Firstly, Rhoads [49] has considered the contribution of young, massive red supergiant stars to near-infrared emission. These stars become red supergiants when they evolve off the main sequence at $\sim 10^7$ years. Rhoads [49] concludes that "while the overall 2 μ m emission is primarily due to old stellar populations, local features in the near-infrared light of disk galaxies can be due primarily to *young stars*" (emphasis, ours).

Second, studies of the backbones of galaxies at wavelengths longer than K must also be treated with care. Abraham (this volume) emphasizes the contamination at 3.6 μm by PAHs, as well as by circumstellar dust. This could explain why the morphologies of a sample of optical and 3.6 μm images have been reported by Buta and collaborators (2010, submitted) to be similar: the ideal wavelength to study galactic backbones longward of K would be 4.5 μm and not 3.6 μm.

It is important to stress that from a physical (dynamical) point of view, one requires *two* classification schemes – one for the gas and a separate one for the stars. A near-infrared classification scheme can never *replace* an optical one, and vice-versa, because the *current* distribution of old stars strongly affects the *current* distribution of gas in the Population I disk.

The dynamic interplay between the two components – via a *self-regulation* mechanism – is crucial and has been studied extensively (e.g. [7]), who term this mechanism a "dynamical thermostat"; self-regulation may actually fail when the two components are not sufficiently well coupled. To quote Pfenniger et al. [47], "The interesting aspect of [feedback] is that the systematic global properties of galaxies are then no longer necessarily determined by the initial conditions of collapse."

We find a duality in spiral structure (Fig. 6), and this has a sound theoretical framework within the modal theory for spiral structure – as developed by Bertin, Lin and collaborators.

A typical turbulent speed associated with the cold gas Population I component is ~ 6 km s^{-1}; in contrast, the velocity dispersion for old stars would typically be six times larger. While a gas cloud may be constrained to move in a thin annulus $\Delta r \sim 300$ pc, an old star may wander in an annulus 2 kpc thick [7]. To derive a *coherent* physical framework for the excitation of spiral structure in galaxies, one must consider the co-existence of these *two* dynamical components.

K band images are invariably (but not always) dust penetrated, dependent upon $\tau(V)$, the optical depth. A few galaxies are so optically thick that even at K (2.2 μm) or K' (2.1 μm) one still sees long arms of dust: an example is NGC 2,841 [13].

We believe there is a fundamental limit in predicting what an evolved stellar disks might look like. The greater the degree of decoupling, the greater is the uncertainty.

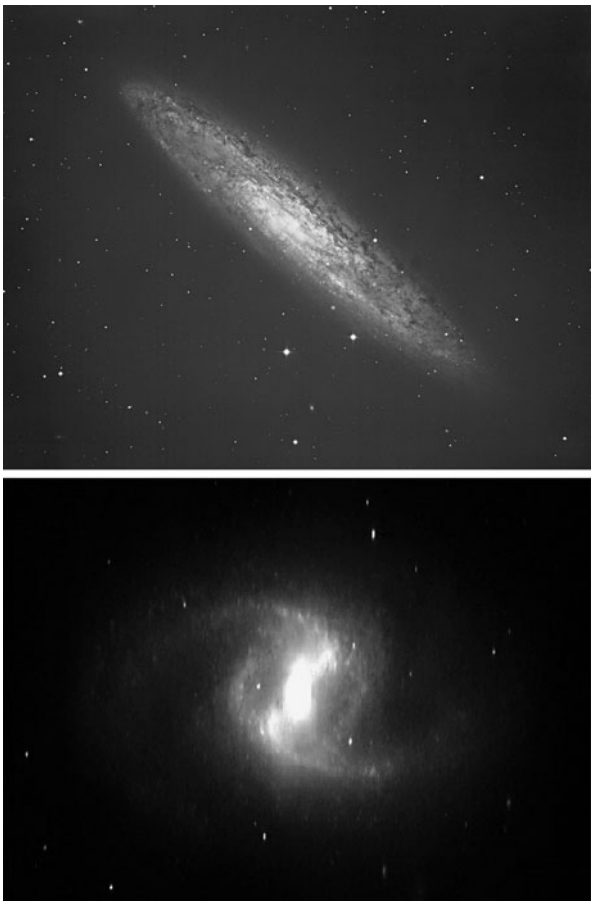


Fig. 6 *Top*: An optical photograph of the spiral galaxy NGC 253 shows a plethora of short, fleece-like (flocculent) spiral arms, with no apparent regularity whatsoever. *Bottom*: A dust penetrated, near-infrared image of NGC 253 shows a bar, an inner ring, and two grand design spiral arms. The infrared image is deprojected to "face-on". *Top*: courtesy D. Malin; *bottom*: courtesy O.K. Park and one of the authors (KCF)

The fact that a spiral might be flocculent in the optical is very important, but it is equally important to know whether or not there is a decoupling with a grand design old stellar disk. The optically flocculent galaxies NGC 5,055 and NGC 3,521, for example, have "grand design" m=2 near-infrared counterparts [54], as does the famous optically flocculent specimen NGC 253 ([15]; see also Fig. 6 in this chapter).

The fact that a spiral might be flocculent in the optical is crucial, but it is also crucial to know whether that morphology is retained in the near-infrared. No prediction on what the near-infrared appearance of an optically flocculent spiral can, a priori, be made. The converse is also true: given a stellar backbone such as in NGC 253,

one cannot predict what the gaseous component will actually look like, as Fig. 6 clearly emphasizes.

An analogy is in order. We are reminded of the duality of matter itself, as pioneered in the seminal work "Ondes et Mouvements" by L. de Broglie. Clouds of probability are associated in both the microscopic and the macroscopic worlds. In the world of quantum mechanics, the wave function has probabilities associated with it. In the world of imaging stellar backbones, one might assign a probability as to what the optical (gaseous) component might look like, but the uncertainty remains. One can never tell for sure [15]. At the foundation of such observations lies the theoretical predictions of the modal theory of spiral structure.

Shu [53] addressed this uncertainty in the coupling/decoupling of gaseous and stellar disks by drawing an analogy with the weather. To quote:

> Make all the measurements you can, still no one can predict the weather seven days from now. You can guess the weather tomorrow with some precision, but you really cannot guess well for a week later, no matter how fine are your observations because of the chaos in the system. We need to be prepared for this in our subject . . .

The 2D Fast Fourier decomposition of the near-infrared images in our studies over the years has employed a program developed by one of us (IP) (see also [52]).

Following Danver [26], logarithmic spirals of the form $r = r_o \exp(-m\theta/p_{\max})$ are utilized in our decompositions.

The amplitude of each Fourier component is given by [52]:

$$A(m, p) = \frac{\Sigma_{i=1}^{I} \Sigma_{j=1}^{J} I_{ij}(\ln r, \theta) \exp\left(-i(m\theta + p \ln r)\right)}{\Sigma_{i=1}^{I} \Sigma_{j=1}^{J} I_{ij}(\ln r, \theta)}$$

where r and θ are polar coordinates, $I(\ln r, \theta)$ is the intensity at position $(\ln r, \theta)$, m represents the number of arms or modes, and p is the variable associated with the pitch angle P, defined by $\tan P = -\dfrac{m}{p_{\max}}$.

On the basis of the tightness of the winding of the spiral arm pattern in the near-infrared, we identify three dust penetrated classes α, β, and γ, depending on the pitch angle of the dominant m-spiral at K or K' (see Fig. 7). Typical values for these bins are 10, 25, and 40° respectively. On the y-axis of Fig. 7 is plotted the gravitational bar torque, as elucidated in Buta and Block [23].

Many galaxies show the presence of a significant m=1 component in the near-infrared (often in the form of a lopsidedness of the spiral). The linear modal theory of Bertin et al. [8, 9] predicts that m=1 modes should generally be dominated by m=2 modes when available, since the latter are more efficient in transporting angular momentum outwards. However, modes with m>2 are generally suppressed in the stellar disk by the Inner Lindblad Resonance (ILR). While the disk mass participating in the mode is crucial, the gas content of the galaxy is important: gas-rich spirals can generate modes m>2. It had earlier been predicted [7, 12] that infrared images should show *a ubiquity of global one and two-armed structures in the underlying stellar disk* and we believe that any classification of dust-penetrated

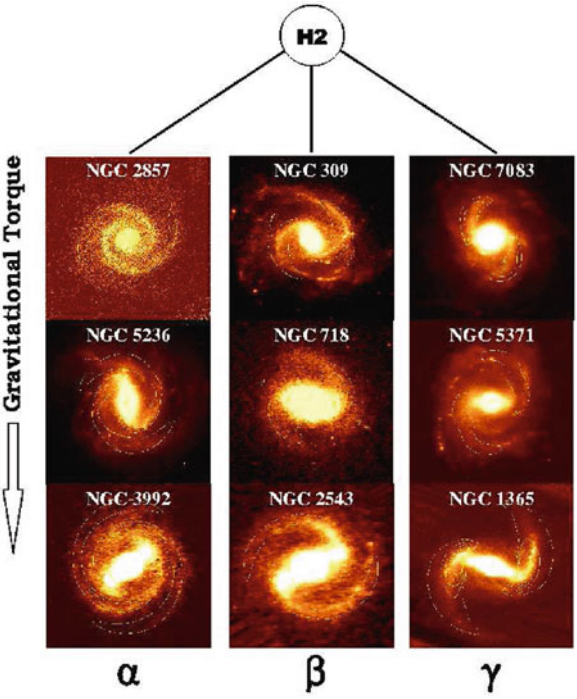


Fig. 7 Spiral galaxies in the dust penetrated regime are binned according to three quantitative criteria: first, Hm, where m is the dominant Fourier harmonic (illustrated here are the two-armed H2 family). Next, follows the dust-penetrated pitch angle families α, β, or γ (for class α, the pitch angles range from $\sim$ 4 to 15°; for class β, the deprojected pitch angles range from $\sim$ 15 to 30°, while open-armed class γ spirals have pitch angles ranging between $\sim$ 35 and 75°). Finally, we compute the gravitational torque, which is identical to the bar torque in galaxies presenting a bar. Bar torques are not derived from bar ellipticities but exploit the full gravitational potential of the disk within which it is embedded. Note that early type b spirals (NGC 3992, NGC 2543, NGC 7083, NGC 5371, and NGC 1365) are distributed within all three families (α, β, and γ). Hubble type and dust penetrated class are uncorrelated

disks at K should take full account of this result. Theory and observations are in excellent agreement.

In contrast, the dynamics of the cold gaseous Population I disk, characterized by different scalelengths, velocity dispersion, thickness, and behaviour at the relevant Lindblad resonances, explains why spiral galaxies are optically so often overwhelmed by higher m modes and other more irregular fast evolving features, supported by the cold interstellar gas [5, 6].

The redistribution of angular momentum by large-scale spiral torques will be stronger for stellar arms which are more open; some authors have postulated that such a redistribution may lead to rapid changes in the disk and even modify the properties of the rotation curve. This is the concept of secular evolution of a galaxy, from an open to a more tightly wound morphology, within one Hubble time.

4 The Building of Galactic Disks: Active Evolution in the Triangulum Spiral M33

> That galaxies evolve is a trivial statement: galaxies are made of stars, stars evolve, and therefore galaxies evolve. But over and above such minimal evolution of the older stellar content of a galaxy, which is often called *passive* evolution, there exists the possibility of *active* galactic evolution . . . Augustus Oemler, Jr.

Disks of galaxies appear to form from the inside out [18] and evolve in an *active* sense. When using the terminology *active galactic evolution*, we include evolution of carbon stars in a galaxy – specifically in the outer domains. Carbon stars are important tracers of intermediate age stellar populations; they also make excellent test particles for investigating the kinematics of galaxies [46]. Carbon stars output such prodigious fluxes of light in the near-infrared that the K-band luminosity in galaxies containing a significant population of these intermediate age stars can be enhanced by upto a factor of 2 [45].

Relative to the inner regions of spiral galaxies, the mean ages of the outer regions are known to be somewhat younger and more metal-poor [4]. We thus expect the contribution from intermediate-age stars to be stronger in these outer regions, and the near-IR surface brightness of the outer disk of a galaxy will be preferentially enhanced by the presence of thermally pulsing – asymptotic giant branch (TP-AGB) stars. This brightening of the outer disk in the near-IR may well contribute to the apparent sharp radial truncation observed in the disks of many spiral systems [38]. Astronomers routinely use near-IR observations to minimize the effects of dust extinction, but it is precisely in this band that TP-AGB stars are so bright; the actual contribution of carbon stars to the near-IR light of spiral galaxies beyond our Local Group is a field which remains to be fully exploited, especially with forthcoming very large telescopes. It was Aaronson [1] and others who, many years ago, urged astronomers to remember the importance of the asymptotic giant branch for understanding the stellar content of nearer galaxies.

In the context of exploring the growth of a nearby disk, we present deep near-infrared images of the Triangulum Spiral Messier 33 in our Local Group (see Fig. 8 and the discussion in [21]). M33 was probably first discovered by Giovanni Hodierna in 1654, before being independently discovered by Messier almost a century later, in 1764. The question beckons: Is Messier 33 an isolated, dynamically closed system of stars, as indeed suggested by early photographs, such as those secured using the Crossley reflector at the Lick Observatory? It was Heber D. Curtis who in 1918 described the Triangulum Galaxy as "a close rival to the Nebula of Andromeda as the most beautiful spiral known".

The galaxy presents somewhat of an optically flocculent (fleece-like) appearance, a fact already attested to by Lord Rosse in his early drawings ("the whole nebula in [is] flocculi"). In a range of 1–12, Elmegreen and Elmegreen [29] place the Triangulum Galaxy in their class 5 bin. Sandage and Humphreys [51] note that ten spiral arms may be identified in deep optical images, although the galaxy is most famous for its two bright inner set of spiral arms. At longer, near-infrared wavelengths, the

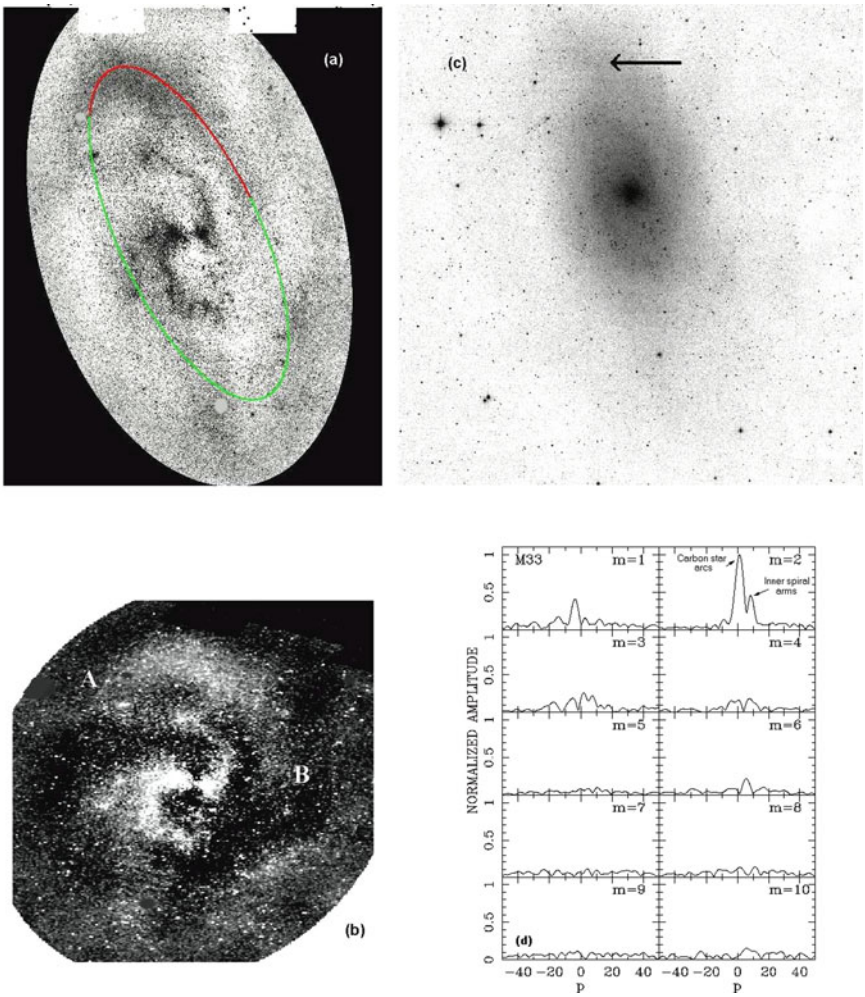


Fig. 8 **a** A deep JHK_s mosaic of M33. A gargantuan plume-like ring of red stars stretches in a swath (up to $5'$ in width) for over 120°, commencing at $\sim 14'$ north of the centre of M33. In order to best reveal the plumes and inner arms, we have subtracted an axisymmetric model of the disk. **b** A deprojected JHK_s mosaic of M33, with an axisymmetric disk model subtracted. The northern plume is labelled **a** to **b**. **c** A deep 2MASS H band image of M33 (non-deprojected) reveals the northern arc (*arrowed*) and other red arcs in the outer disk of M33, **d** Fourier spectra generated from the deprojected near-infrared mosaic of M33. The dominant $m = 2$ mode in M33 arises from the giant outer ring or arcs of carbon stars

galaxy appears much smoother as the older underlying population dominates the light [47]. M33 extends over 1° along the major axis and it has a plethora of star clusters, emission regions, and supernova remnants (see [35] for a complete atlas of these). Its distance modulus is 24.64^m, corresponding to a linear distance of 840 kpc [31]. Its linear diameter is $\sim$ 17 kpc, approximately one-half that of the Milky Way.

4.1 Farewell to Monolithic Collapse of M33

Deep 2MASS images of M33 reveal remarkable arcs of red stars in the outer disk, spanning over 120° in azimuth and delineating a swath up to 5′ (see Fig. 8). The northern arc is dominant although a faint southern counterpart arc, forming a partial ring, can be seen in Fig. 8. In order to better see these plume features visually, we subtract off an axisymmetric disk model. The $J-K_s$ colours of stars in these arcs are so red (1.2 mag and cooler) that an M-giant population is excluded as the dominant light source. The arcs would be completely missed in conventional near-infrared surveys such as 2MASS.

To generate our Fourier spectra, we carefully determined the barycentre of the light (mass) distribution and deprojected about that point. As in Deul and van der Hulst [27] and Regan and Vogel [48], we adopted the de Vaucouleurs values of log R = 0.23 (R is the ratio of the major–minor axis) and a position angle of 23°.

In M33, we comment that the only dominant modes are $m = 1$ and $m = 2$; the galaxy does not present 10 arms in the near-infrared, as it does optically. The theme of the duality of spiral structure is ever recurrent. We immediately note that the two inner "grand-design" trailing spiral arms seen optically show up rather strikingly in the near-infrared regime. The computed pitch angle is $\sim$ 14°. This inner set of near-infrared arms may also be seen in Regan and Vogel [48] (see their Figs. 1 and 3).

Far more important, however, is the *dominant* and very high peak for m=2; the peak does not occur at $p_{\max}$ of zero (as expected for bars, for example, with pitch angles $\sim$ 90°, corresponding to spectral m=2 signatures with a peak at zero), as noted in Fig. 8. Rather, the dominant m=2 peak corresponds to the set of giant arcs (very open two-arm structures) on both the northern and the southern sectors of the disk (Fig. 8, panel d). The pitch angle of the dominant m=2 mode is $\sim$ 58°. *Surprisingly, it is these outer arcs (and not the inner grand design structure for which M33 is so famous) which dominates the m=2 Fourier spectra.*

In Block et al. [20], the Low-Resolution Imaging Spectrograph (LRIS) attached to the Keck I telescope and the Echellette Spectrograph and Imager (ESI) on Keck II was used to secure spectra of stars in the arcs. To resolve individual stars in the arcs, the Hale 5-m reflector, equipped with the 2048 × 2048 near-IR camera WIRC, was employed to provide photometrically calibrated images at JHK. Keck spectra confirm that stars in the arcs are indeed carbon stars, being readily identified by the C_2 Swan bands and CN bands.

Where have these carbon stars in the outer disk formed? Where is accretion of gas in M33 most efficiently expected to take place? M33 has a huge warped envelope of neutral hydrogen gas, with a notable asymmetric extension toward the North-West (e.g. [25]). These authors suggest that this NW extension may be a betrayal of tidal interaction between M33 and M31, which are only $\sim$ 200 kpc apart.

In their modelling of the HI envelope, Corbelli and Schneider find that the phase changes at a radius of $\sim$ 20 arcmin (2–3 disk scalelengths), which is precisely the outer domain of the arcs reported in our study. It seems highly plausible, therefore, that fresh, low-metallicity gas is being fed to the host galaxy M33, via external accretion. The outer disk from which mass is accreted is inclined to the inner disk

of M33 and has a different angular momentum. We believe that it is the signature of gas flowing inward and accreting at $\sim$ 20 arcmin, from which the very red, and relatively metal-poor stars, have been formed.

In summary, we propose that intermediate age (0.6–2 Gyr) carbon stars have formed as a result of accretion at the HI warp. If true, we bid farewell to monolithic collapse in our Local Group: galaxy disks such as that in Messier 33 grow from the inside, outward.

Carbon star diagnostic for the evolution of outer disks hold enormous promise, due to prodigious output in K-band fluxes.

It is tempting to liken the ring in M33 to the one recently reported in the outer disk of our Milky Way [36]. For the Milky Way ring, the rotation period is $\sim$ 600 Myr. For M33, the rotation timescale at the radius of the ring would be $\sim$ 240 Myr. Any carbon-bearing clusters of age 0.6 Gyr would have only undergone $\sim$ 2 orbits.

5 Lessons from the Andromeda Spiral

Although regarded for decades as showing little evidence of a violent history, the Andromeda Spiral Messier 31 has a well-known outer ring of star formation at a radius of 10 kpc whose centre is offset from the galaxy nucleus. Block et al. [19] report the presence of a second, inner dust ring with projected dimensions 1.5 by 1 kpc and offset by $\sim$ 0.5 kpc from the centre of the galaxy. The two rings are interpreted to be density waves propagating in the disk of M31. We suggest that both rings result from a companion galaxy (presumably M32) plunging in a near head-on collision through the disk of M31. That an almost head-on collision has occurred only $\sim$ 210 million years ago suggests the treachery of optical images of the Andromeda Spiral, wherein one is tempted to believe that its morphology has existed quiescently for billions of years.

6 Conclusion – "La condition humaine"

> Lo, there, quod he, cast up thine eye. Se yonder, lo, the Galaxye, Which men clepeth the Milky Wey, For hitt is whytt, and some parfey, Callen hit Watlinge Strete. – Geoffrey Chaucer (The House of Fame)

> What are galaxies? No one knew before 1900. Very few people knew in 1920. All astronomers knew after 1924 . . . Each galaxy is a stellar system somewhat like our Milky Way, and isolated from its neighbors by nearly empty space.

The above definition (from the *Hubble Atlas*, 1961) is not true for the Andromeda Spiral, neither for the Triangulum Galaxy (as also clearly emphasized by Ferguson at our Conference).

As far as dwarf galaxies are concerned, the title of a forthcoming paper "Willman 1 in X-rays: Can you tell a dwarf galaxy from a globular cluster?" [43] emphasizes the critical need for a clean separation between dwarf galaxies and stellar clusters.

The pioneering research of Arnaboldi and collaborators (e.g. [3]) highlights the ongoing formation of galaxies within clusters, by tracing diffuse intracluster light by means of the identification and radial velocity measurements of intracluster planetary nebulae, using their [OIII] 5007 Å emission.

Why bother about the definition of a galaxy?

Prior to the inference of dark matter haloes in spiral galaxies from rotation curve studies [32], dark matter was not regarded as a crucial component of a galaxy. Today we know otherwise, although we know that tidal dwarf galaxies show no evidence for dark matter. (Presumably these have formed from baryons stripped during an encounter.)

In defining a galaxy, there are clearly different levels of complexity; here are a few:

Should magnetic fields be incorporated into the definition of a galaxy? We are reminded of a sobering thought by Shu [53] in addressing this specific point at a Conference:

> We haven't yet heard a single talk that had Maxwell's equations in it. It [magnetic fields] is the only other long range force that we know about, right? When did magnetic fields become important in galaxy formation and evolution? The fact that we see synchrotron radiation in radio sources certainly says that in some parts of galaxies, magnetic fields were generated at early times. Magnetic fields are important because they are the dominant way, other than gravitational torques, by which you can transport angular momentum. So if angular momentum is a big problem in galaxy formation, maybe we should look at magnetic fields before we think of exotic forms of dark matter!

What about accretion of gas in the outer domains? The duality of spiral structure? What are the implications of the deep images of David Malin (Figs. 2, 3, and 4), in defining the term "galaxy"? New galaxy classification schemes are called for. Vera Rubin writes: "One look at the stunning images produced by astrophotographer David Malin is sufficient to convince any skeptic."

Without a clear understanding of what constitutes a galaxy, astronomers walk in a land of shadows and of masks.

In the spirit of Magritte's "La trahison des images" seen in Fig. 1, we provide the reader with three astronomical equivalents, shown Figs. 9, 10, and 11. The legend in Fig. 9 is translated "This is not a globular cluster", which might at first sight seem rather unusual for ω Centauri. It has been argued that what we see in Fig. 9 is rather a representation of an ancient galaxy ([33] and subsequent papers).

The legend in Fig. 10 reads: "This is not a galaxy". Of course not. What we see is the optical mask only – a representation of the galaxy at optical wavelengths. How amazingly different is the morphology of the Sombrero in deep images (Fig. 4).

Likewise, Fig. 11 is an optical representation of the galaxy NGC 5084, with one of the highest mass-to-light ratios known (as studied by Freeman and collaborators. For a detailed discussion, see [24]). Figure 11 fails to yield the minutest hint of its massive dark matter halo; *this photograph* – that which we see represented in Fig. 11 – is *not* NGC 5084, with its massive dark matter halo.

> In place of an astronomer who frames the Triangulum Galaxy or the Andromeda Spiral Galaxy or ω Centauri to picture it, Nature produces its own image without a necessary frame.

Fig. 9 Omega Centauri, captured by the Wide Field Imager at the European Southern Observatory, Chile. "Ceci n'est pas un amas globulaire" is translated "This is not a globular cluster". Photograph: ESO

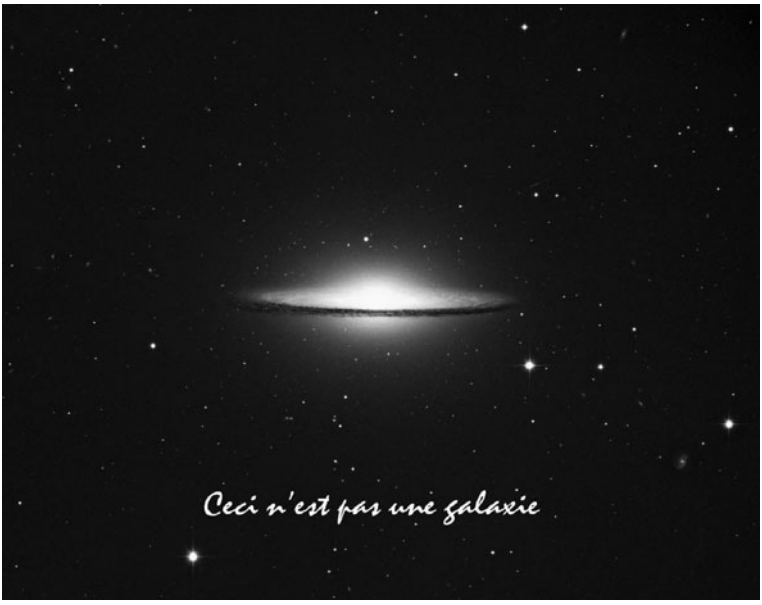


Fig. 10 Below the optical representation of Messier 104 appears these words: "Ceci n'est pas une galaxie" being translated "This is not a galaxy". Photograph: David Malin

Fig. 11 NGC 5084, the most massive disk system known, has a mass–to-light ratio of ~ 200, and an inferred mass up to 1×10^{13} $M_{\odot}$. *This* – that which is represented in this optical photograph – is not *the galaxy NGC 5084* – with its massive, undetected dark matter halo. Ceci n'est pas une galaxie. It remains a representation thereof, in the optical regime. Photograph: Anglo Australian Telescope

This is, after all, our human condition.

We tend to "frame reality" in our minds, as captured in a poignant and astute tone by René Magritte in his exquisite rendition of the human condition, entitled "La condition humaine" (1934).

Acknowledgments We are indebted to all our collaborators with whom it has been so enriching to conduct research with, over the many years. The title of this chapter is drawn from a book by two of us (DLB and KCF), published in New York in 2008. It is a great pleasure to thank Giuseppe Bertin and Bruce Elmegreen for their careful reading of our chapter. We warmly thank Françoise Combes for her expert and most willing assistance in translation. It remains a great pleasure to thank AVENG and Mr. F. Titi for their continued sponsorship. DLB fully acknowledges the visionary insight of Roger Jardine, Kim Heller, and Fani Titi in facilitating our ongoing research into masks of cosmic dust. IP acknowledges support from the Mexican Foundation Conacyt.

References

1. Aaronson, M. 1986, In: Stellar Populations, Cambridge University Press, Cambridge, p. 45
2. Allen, R.J. 1996, In: New Extragalactic Perspectives in the New South Africa, eds. D.L. Block, J.M. Greenberg, Kluwer, Dordrecht, p. 585

3. Arnaboldi, M., et al. 2006, IAU Symposium, 234, 337
4. Bell, E.F., de Jong, R.S. 2000, MNRAS, 312, 497
5. Bertin, G. 1991, IAU 146. Dynamics of Galaxies and Their Molecular Cloud Distributions, Kluwer, Dordrecht, 93
6. Bertin, G. 1993, PASP, 105, 604
7. Bertin, G., Lin, C.C. 1996, Spiral Structure in Galaxies: A Density Wave Theory, MIT Press, Cambridge, MA
8. Bertin, G., Lin, C.C., Lowe, S.A., Thurstans, R.P., 1989a, ApJ, 338, 78
9. Bertin, G., Lin, C.C., Lowe, S.A., Thurstans, R.P., 1989b, ApJ, 338, 104
10. Bertin, G., Amorisco, N.C. 2010, A&A, 512, 17
11. Block, D.L. 1996, In: New Extragalactic Perspectives in the New South Africa, eds. D.L. Block, J.M. Greenberg, Kluwer, Dordrecht, p. 1
12. Block, D.L., Bertin, G., Stockton, A., Grosbøl, P., Moorwood, A.F.M., Peletier, R.F. 1994a, A&A, 288, 365
13. Block, D.L., Elmegreen, B.G., Wainscoat, R.J. 1996, Nature, 381, 674
14. Block, D.L., Stockton, A., Elmegreen, B.G., Willis, J. 1999, ApJ, 522, L25
15. Block, D.L., Freeman, K.C. 2008, Shrouds of the Night, Springer, New York, NY
16. Block, D.L., Wainscoat, R.J. 1991, Nature, 353, 48
17. Block, D.L., Witt, A.N., Grosbøl, P., Stockton, A.1994b, A&A, 288, 383
18. Block, D.L., Bournaud, F., Combes, F., Puerari, I., Buta, R. 2002, A&A, 394, L35
19. Block, D.L., Bournaud, F., Combes, F., et al. 2006, Nature, 443, 832
20. Block, D.L., Combes, F., Puerari, I., Freeman, K.C., et al. 2007, A&A, 471, 467
21. Block, D.L., et al. 2004, A&A, 425, L37
22. Burstein D., Rubin, V. 1985, ApJ, 297, 423
23. Buta, R.J., Block, D.L. 2001, ApJ, 550, 243
24. Carignan, C., Cotê, S., Freeman, K.C., Quinn, P.J. 1997, AJ, 113, 1585
25. Corbelli, E., Schneider, S.E. 1997, ApJ, 479, 244
26. Danver, C.G. 1942, AnLun, 10
27. Deul, E.R., van der Hulst, J.M. 1987, A&AS, 67, 509
28. Elmegreen, B.G., Block, D.L. 1999, MNRAS, 303, 133
29. Elmegreen, D.M., Elmegreen, B.G. 1987, ApJ, 314, 3
30. Foucault, M. 1982, This Is Not a Pipe, University of California Press, Berkeley, CA
31. Freedman, W.L., Wilson, C.D., Madore, B.F. 1991, ApJ, 372, 455
32. Freeman, K.C. 1970, ApJ, 160, 811
33. Freeman, K.C. 1993, ASP, 48, 608
34. Freeman, K.C. 1999, In: Towards a New Millenium in Galaxy Morphology, eds. D.L. Block, I. Puerari, A. Stockton, D. Ferreira, Kluwer, Dordrecht, p. 119
35. Hodge, P.W., Skelton, B.P., Ashizawa, J. 2002, An Atlas of Local Group Galaxies, Kluwer, Dordrecht
36. Ibata, R.A., Irwin, M. J., Lewis, G.F., Ferguson, A.M.N., Tanvir, N. 2003, MNRAS, 340, L21
37. Klypin, A., et al. 1999, ApJ, 522, 82
38. Kregel, M., van der Kruit, P.C., de Grijs, R. 2002, MNRAS, 334, 646
39. Lin, C.C. 1971, In: Highlights of Astronomy, ed. C. de Jager, Reidel, Dordrecht, 88
40. Lynds, B.T. 1974, ApJS, 28, 391
41. Martin, P.G., Whittet, D.G.B. 1990, ApJ, 357, 113
42. Meurer, G., Carignan, C., Beaulieu, S., Freeman, K.C. 1996, AJ, 111, 1551
43. Mirabal, N., Nieto, D. 2010, MNRAS, in press
44. Moore, B., et al. 1999, ApJ, 524, L19
45. Mouhcine, M., Lancon, A. 2003, A&A, 402, 425
46. Mould, J., Aaronson, A. 1986, ApJ, 303, 10
47. Pfenniger, D., Martinet L., Combes, F. 1996, In: New Extragalactic Perspectives in the New South Africa, eds. D.L. Block, J.M. Greenberg, Kluwer, Dordrecht, p. 291
48. Regan, M.W., Vogel, S.N. 1994, ApJ, 434, 536

49. Rhoads, J.E. 1997, In: Extragalactic Astronomy in the Infrared, eds. G.A. Mamon, T.H. Thuân, J.R.T. Vân, Editions Frontières, Gif-sur-Yvette, France, 45
50. Rix, H.-W., 1993, PASP, 105, 999
51. Sandage, A., Humphreys, R.M. 1980, ApJ, 236, L1
52. Schröder, M.F.S., Pastoriza, M.G., Kepler, S.O., Puerari, I. 1994, A&AS, 108, 41
53. Shu, F.H. 2004, In: Penetrating Bars Through Masks of Cosmic Dust, eds. D.L. Block, I. Puerari, K.C. Freeman et al., Springer, Dordrecht, p. 813
54. Thornley, M.D. 1996, ApJ, 469, L45
55. Zwicky, F. 1957, Morphological Astronomy, Springer, Berlin


Nikon

Twin Masks of Spiral Structure? A Local Perspective

Thomas Y. Steiman-Cameron

Abstract We examine models for the spiral structure of the Milky Way proposed over the past half century. Many approaches have been pursued to decipher the geometry of the Galaxy's spiral arms, often with conflicting results. While a general consensus exists that a global pattern exists, considerable disagreement remains in the details. Arm geometries, orientations, and even the number of arms are still debated. Close examination of the literature reveals a clear division between four- and two-arm spiral arms. Four-arm models follow naturally from observations of classical tracers of spiral arms – enhanced gas densities and associated star formation – while two-arm models primarily flow from observations linked to the distribution of cool evolved stars. We examine the dichotomy between two-arm and four-arm models and discuss its implications.

1 Introduction

The classical tracers of spiral arms include enhanced gas and dust densities, star-forming regions, OB stars and associations, supernovae, and photodissociation regions. These physical characteristics lead to enhanced radiation at specific wavelengths whose observations can be used to trace the arms. For example, neutral atomic gas is characterized by 21 cm HI emissions, molecular gas by CO and H_2 emissions, OB stars and young stellar associations by enhanced UV fields, relativistic particles characterized by synchrotron emission and emission from HII regions, and dust by FIR continuum radiation. In principal, the geometry of the Galaxy's arms can be obtained from determinations of the spatial distributions of these emissions and/or the sources of these emissions themselves. Difficulties in determining the precise location of these tracers, and limits on the classically defined tracers, underlie much of the uncertainty in the Milky Way's spiral structure.

T.Y. Steiman-Cameron (✉)
Indiana University, Bloomington, IN, USA
e-mail: tomsc@astro.indiana.edu

D.L. Block et al. (eds.), *Galaxies and their Masks*,

DOI 10.1007/978-1-4419-7317-7_3,

While it had long been suspected that the Milky Way is a spiral galaxy, the first direct evidence of spiral structure in the Galaxy did not exist before the early 1950s when Morgan and coworkers [37, 38] determined that blue giants in the solar neighborhood fall along three well-defined arcs. They interpreted these arcs as portions of local spiral arms. By the late 1950s, Mills [30, 36] reported that the intensity profile of nonthermal radio emissions along the Galactic plane displayed a "stairstep" profile. He argued that the locations of these steps corresponded with tangents of spiral arms. During the following decade, local structures continued to be outlined but progress in pinning down the global structure of the spiral arms was slow in coming. By 1970, only two models had been presented for the Galaxy's spiral arms [31, 63]. However, despite using essentially the same 21 cm observations, they came to very different conclusions: Kerr and Kerr [31] proposed a four-arm model for the Galaxy while Weaver [63] advocated two spiral arms. Their areas of agreement were limited: there is global structure to the Galaxy's spiral arms, a tangent direction to a global arm exists at near 50° Galactic longitude (the Sagittarius Arm), and the local arm (Orion spur) is not a global feature.

Through the early 1970s, the only means available for exploring the global structure of the arms was 21 cm observations [51]. Tangent directions obtained from nonthermal radio observations could assist 21 cm observations, but were not reliable. Existing 21 cm models did not match up with local arm structures derived from optical observations. New technology was expected to bring broader HII region distance determinations within reach in the near future. This expectation was finally met in the latter part of the 1970s.

2 Approaches to Modeling

Since the studies of Kerr and Kerr [31] and Weaver [63], approximately 100 models have been proposed for the spiral structure of the Milky Way. Many approaches have been pursued to decipher the geometry of the Galaxy's spiral arms, often with conflicting results. While a general consensus exists that a global pattern exists, considerable disagreement remains in the details. In particular, the number of arms is still debated. Here we discuss the dichotomy between two- and four-arm models and explore the situation from historical and physical perspectives.

Approaches to determining the distances and distribution of spiral tracers can be grouped into seven general categories. These include (1) spectrophotometric distances, (2) kinematic distances, (3) velocity-field mapping, (4) tangent point fitting, (5) emissivity mapping, (6) B-Field mapping, and (7) electron density mapping.

Spectrophotometric Distances: This provides the most direct means of determining distances to spiral tracers. It has been applied to OB associations, the exciting stars of HII regions, long-period Cepheids, Wolf-Rayet stars, carbon N-stars, and Be stars. Its principal sources of error include misclassification of a star's spectral type and distance scale uncertainty (distance to the Hyades). While the most accurate of all approaches, it is limited to relatively nearby spiral structures. For this reason, all *global* models of the Galaxy's spiral structure are based upon other methods.

Kinematic Distances: When velocity information is available, kinematic distance determinations are often utilized. This is frequently used in studies of HII regions. Kinematic distance determinations require the assumption of a kinematic model for the Galaxy (rotation curve, characteristics of the Local Standard of Rest, and peculiar motions of the Sun relative to the LSR). External to the solar circle, unique distances are provided by the observations. Interior to the solar circle, two solutions exist. In some cases, additional OH or H_2CO observations are used to resolve the ambiguity. Possible sources of error include random dispersions about the local systemic velocity, streaming motions, uncertainties in the Galactic kinematic model, non-uniqueness interior to solar circle, and velocity crowding. Collectively, these can provide 10–20% uncertainties in derived distances [46].

Velocity-Field Mapping: This approach is related to kinematic distances, except in this case the velocity field of gas, in the form of longitude-velocity maps (typically CO or HI) serve as the observational data to be fit. Self-consistent dynamical models of gas flows in an assumed Galactic gravitational potential are followed and the complex velocity fields predicted by these models are compared with observed HI and CO kinematics to find reasonable matches between model predictions and observations. Models for streaming motions have also been examined in the light of velocity field observations to determine the presence and characteristics of spiral arms [45]. Distances obtained with this method are subject to the same sources of error cited above for kinematic distances. In addition, uniqueness problems provide a major source of uncertainty.

Tangent Point Fitting: Tangent directions to spiral arms have been evaluated using intensity, star count, and velocity data. Intensity maps at many wavelengths exhibit distinct local maxima along the Galactic Plane. The directions along which these local maxima are seen are commonly interpreted as tangent directions to spiral arms. Tangent directions have also been assigned to those longitudes where the slope suddenly changes in plots of terminal-velocity versus longitude. In addition, maxima in line-of-sight stellar densities in the IR (star counts) have interpreted as tangent directions [11, 16, 17, 62]. Geometries of spiral arms are obtained by appropriately pairing peaks on opposite sides of the Galactic center along with an assumed spiral form (the logarithmic spiral form is commonly assumed). Sources of uncertainty include the possibility that a specific tracer may be lacking at or near a tangent point.

Emissivity Mapping: In situations where a strong contrast exists between the emissivity of a tracer in arm and inter-arm regions, emissivity mapping has shown itself to be a promising approach. This approach is related to tangent point fitting, except that intensity information is now available at Galactic longitudes other than just at tangent locations. A mathematical model is created for the volume emissivity of the tracer, as a function of position within the Galaxy. The emissivity model is convolved with the response function of the detector associated with observations and integrated over the intercepted volume of the Galaxy. The resultant synthetic intensity is then compared with observed maps. By varying the free parameters of the emissivity model (which include the geometry of the arms) values for these parameters are found which maximize the agreement with the observations. This

approach has been used with synchrotron emissions, FIR dust emission, gamma-ray observations, and [CII] and [NII] intensities. An analogous approach has also been used with stellar population surveys. Nonuniform emissivity and uniqueness issues are sources of uncertainty.

B-Field Mapping: This approach examines the structure of the Galaxy's magnetic field. It has been proposed that a bisymmetric spiral field exists in the Galaxy and this field reverses its direction from one arm to the next. To examine this structure, a number of authors have utilized Faraday rotation measures (RM) of the polarization plane of extragalactic radio sources and pulsars. The observed RM, which is a strong function of the B-field along the line of sight, is fit with predictions from models with assumed magnetic field and thermal electron distributions. The validity of the underlying assumptions concerning the properties of the Galaxy's spiral magnetic field is disputed and provides a major source of uncertainty.

Electron-density Mapping: This approach utilizes dispersion measures of pulsars combined with models of the electron density field of the Galaxy. It has principally been used to extend the spatial range of existing models [55].

Complementing these approaches are the meta-analyses of Vallée [58–61]. Vallée surveyed models published in the literature and derived new models by performing appropriately weighted averages to these models.

3 Model Statistics and the Two-Arm / Four-Arm Dichotomy

Since the first detailed models of the Galaxy's spiral structure were presented, approximately 100 models have been proposed, with $\sim$ 60 being fully independent and global. Of the latter, 17% have two arms, 68% have four arms, 5% involve a superposition of two and four arms, and 10% have complex arm structures. In addition, a three-arm model was found to fit COBE ISM cooling line data, but was considered to be unrealistic on other grounds [53, 54]. Slightly more than half of the models are logarithmic spirals. Proposed pitch angles range from 5° to 31°. Kinematic (HII Regions), tangent fitting, and emissivity mapping approaches have provided the most consistent models. Figure 1 shows the names traditionally applied to arm segments of the four-arm Milky Way. All four-arm models qualitatively resemble this figure, but generally differ in detail. Graphical depictions of a subset of published models are presented in Figs. 2, 3, 4, and 5. To provide an easy means of inter-comparing models, each is superposed on a background depicting the four-arm model of Steiman-Cameron et al. [53, 54].

While the majority of published models have four arms, a number of observations, including several of recent origin [3, 11, 16, 17, 45], have been interpreted as consistent with two arms. To gain insight into this dichotomy, we will briefly examine representative four- and two-arm models.

Steiman-Cameron et al. [53, 54], hereafter SWH, obtained a four-arm model for the Milky Way by fitting volume emissivity models of spiral arms to observations of [CII] 158 μm and [NII] 205 μm line intensities along the Galactic plane obtained

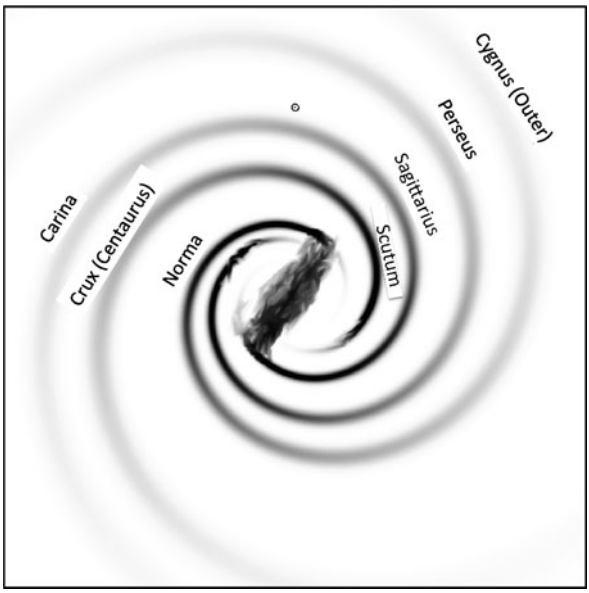


Fig. 1 Traditional and alternate names of arm segments in a four-arm Milky Way, superposed on the model of Steiman-Cameron et al. [53, 54]. Not labeled is the 3-kpc arm, the fourth quadrant arm lying interior to Norma and connected with the Cygnus arm

with the FIRAS instrument of COBE. These lines are very important coolants of the ISM and their emissions are strongly concentrated in spiral arms. They offer perhaps the strongest arm/interarm contrast of any spiral tracer. Comparisons between intensity maps generated from these models and the COBE observations provide very tight constraints on the physical and spatial characteristics of the spiral arms. While finding a four-arm model that provided an excellent match to the COBE/FIRAS data, SWH also determined that no two-arm model is able to fit the FIRAS data. To do so would require arms to cross each other.

The geometry of the SWH model is that shown in Fig. 1. While details differ for individual arms, the SWH model is in good agreement with many four-arm models, including the classic HII region model of Georgelin and Georgelin [25] (Fig. 2) and its later modifications and extensions by Downes et al. [15] (Fig. 2) and Caswell and Haynes [10] (Fig. 2). It also matches well the models of Wainscoat et al. [62] (Fig. 3), Nakanishi and Sofue [39] (Fig. 4), Levine et al. [33] (Fig. 4), and, with the exception of the Crux portion of the Scutum-Crux arm, Vallée [61] (Fig. 5). Inspection of Figs. 2, 3, 4, and 5 demonstrates that the majority of the four-arm models differ only in their details, with the largest regions of disagreement interior to ~ 4 kpc.

In contrast, there is a fair bit of scatter among the two-arm models. Most of them, however, have in common their foundation on a nonclassical spiral tracer – cool evolved stars, predominantly K and M giants. Drimmel and Spergel [16, 17] analyzed the J, K, and 240 μm bandpass data obtained with the *DIRBE* instrument

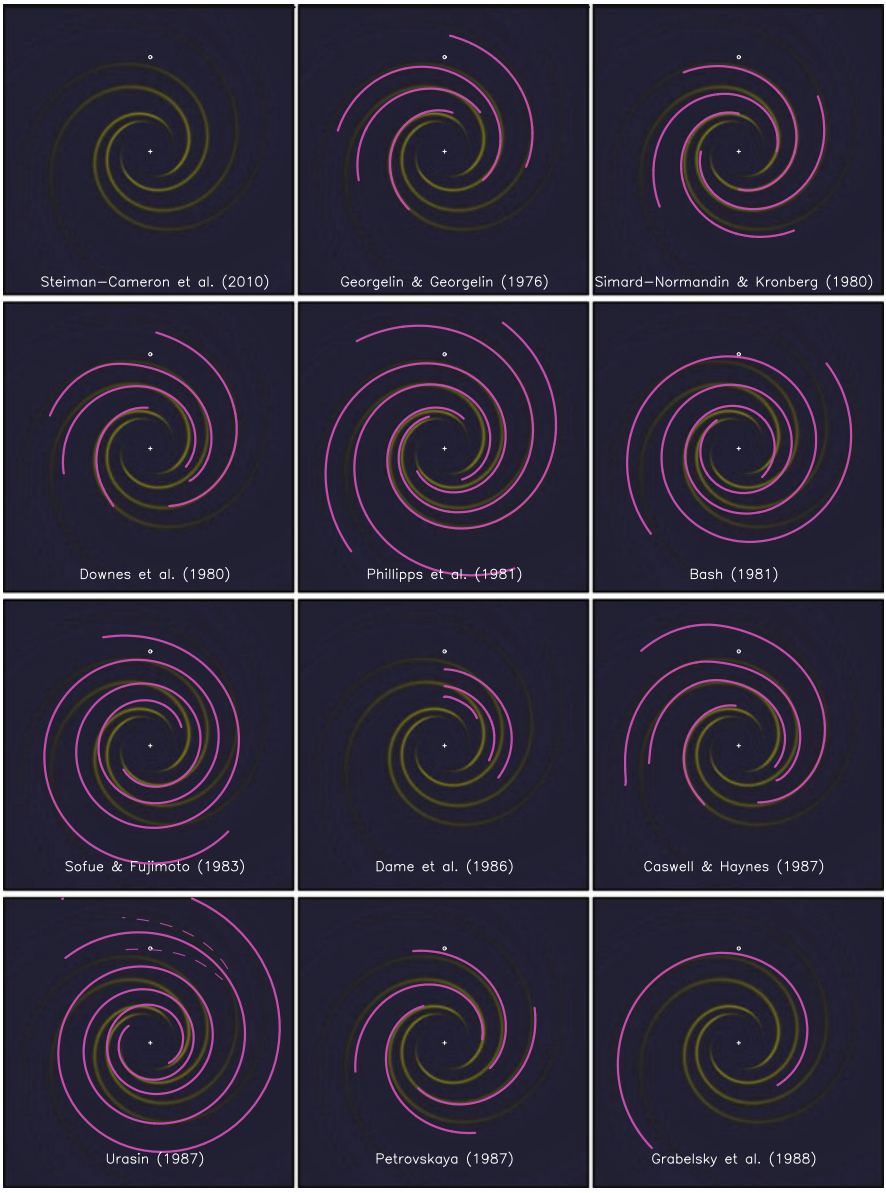


Fig. 2 A subset of models from the literature (1976–1988). To provide an easy means of inter-comparing models, each is superposed on a background depicting the SWH model [53, 54], shown in the first frame. *Top Row*: Steiman-Cameron et al. [53, 54], Georgelin and Georgelin [25], Simard-Normandin and Kronberg [50]; *Second Row*: Downes et al. [15], Phillipps et al. [44], Bash [2]; *Third Row*: Sofue and Fujimoto [52], Dame et al. [14], Caswell and Haynes [10]; *Fourth Row*: Urasin [57], Petrovskaya [43], Grabelsky et al. [28]

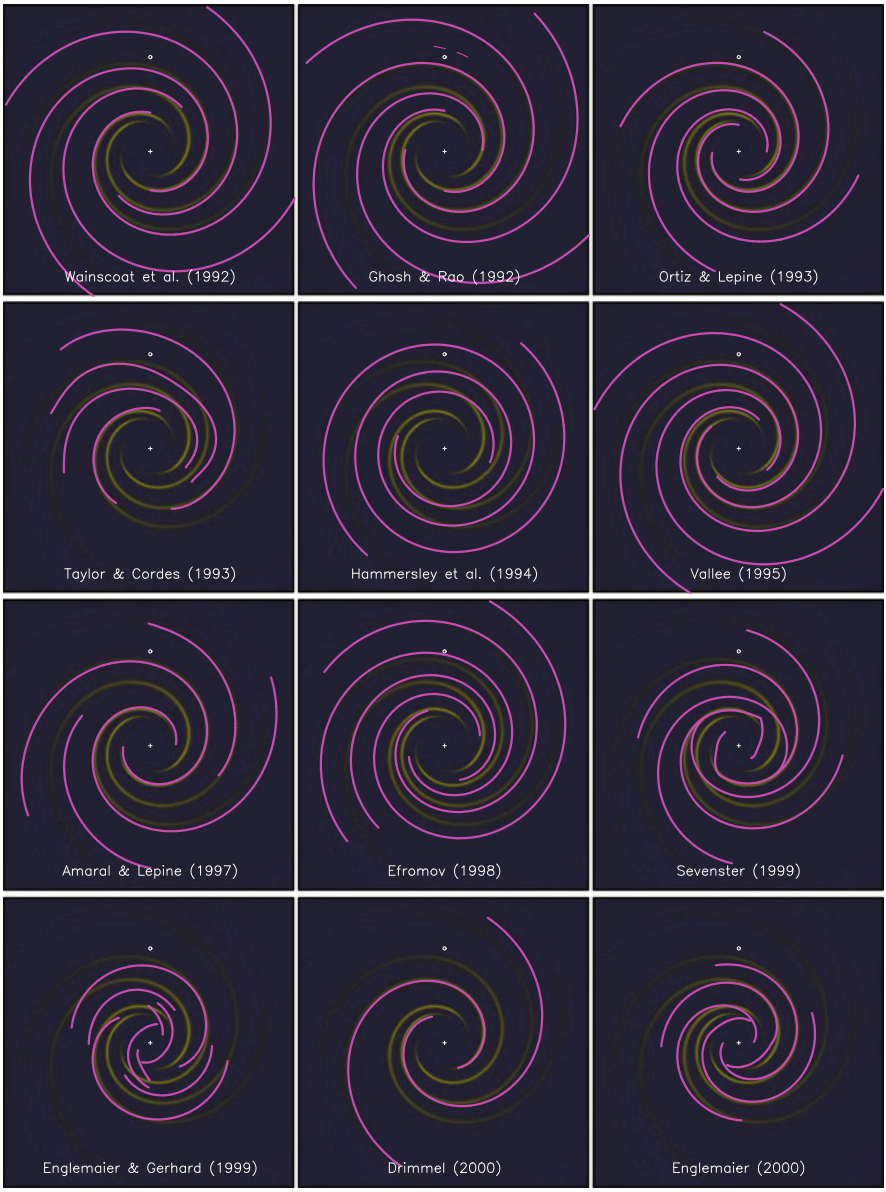


Fig. 3 A subset of models from the literature (1992–2000). *Top Row*: Wainscoat et al. [62], Ghosh and Rao [26], Ortiz and Lépine [42]; *Second Row*: Taylor and Cordes [55], Hammersley et al. [29], Vallée [58]; *Third Row*: Amaral and Lépine [1], Efremov [18], Sevenster [48]; *Fourth Row*: Englmaier and Gerhard [21], Drimmel [16], Englmaier [22]

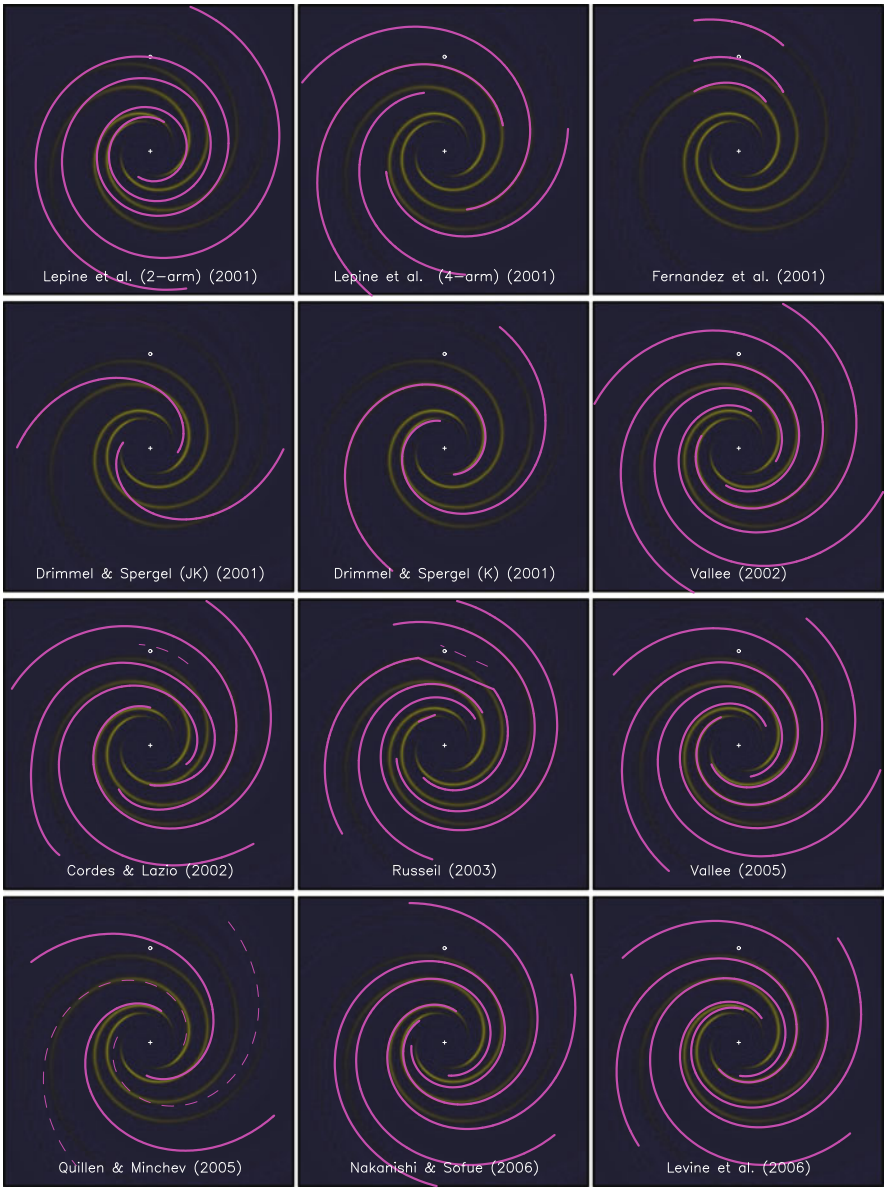


Fig. 4 A subset of models from the literature (2001–2006). *Top Row*: Lépine et al. (2-arm) [32], Lépine et al. (4-arm) [32], Fernandez et al. [23]; *Second Row*: Drimmel and Spergel (J + K bands) [17], Drimmel and Spergel (K-band) [17], Vallée [59]; *Third Row*: Cordes and Lazio [12], Russeil [46], Vallée [60]; *Fourth Row*: Quillen and Minchev [45], Nakanishi and Sofue [39], Levine et al. [33]. The *solid lines* in the Quillen and Minchev model represent the spiral distribution of stars while the *dashed lines* depict the gaseous response to the spiral stellar structure

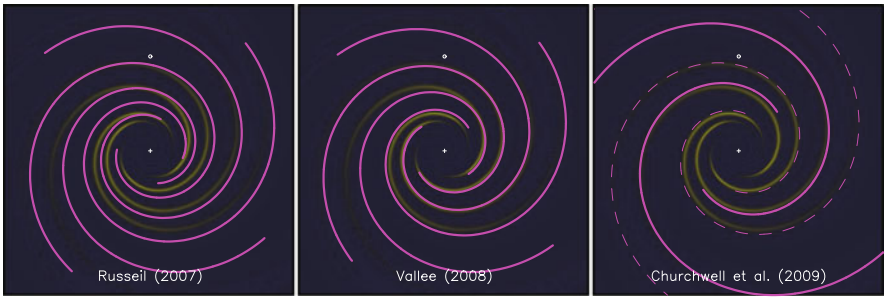


Fig. 5 A subset of models from the literature (2007–2009). Russeil et al. [47], Vallée [61], Churchwell et al. [11]. The *solid lines* in the Churchwell et al. model represent the spiral distribution of K & M giants while both the *dashed* and *solid lines* depict the gaseous component of the spiral structure

of *COBE*. The two near-IR bandpasses are sensitive to cool evolved stars while the 240 μm bandpass is sensitive to dust. Using a tangent point approach, they found evidence of four arms in the 240 μm data, with geometries similar to those described above. On the other hand, only two arms were detected in the J and K data, representing an older evolved stellar component. Two different two-arm configurations were presented, one based solely upon the K-band data and another upon the combined J and K data (Fig. 4). These models differ in that the J plus K-based model arms have much larger pitch angles than the K-band model. There is excellent agreement between their K-band model and the SWH model for the Scutum-Crux arm (see Fig. 4).

Supporting evidence for the view that there exists a two-arm pattern in the old stellar population comes from Quillen and Minchev's investigation into the effect of spiral structure on the stellar velocity distribution in the Solar neighborhood [45]. They performed numerical integrations of particle orbits in the solar neighborhood to determine that clumps in the solar neighborhoods stellar velocity distribution could be caused by spiral density waves. Comparisons with observations lead them to conclude that a two-armed spiral density wave could account for the observed kinematics and that these two arms are essentially stellar in nature. They further argued that two additional gaseous arms may exist as a dynamical response to the gravitational perturbation brought about by the stellar arms (Fig. 4), as previously proposed by Martos et al. [34] (see below).

Churchwell et al. [11] (Fig. 5) also used a tangent fitting approach in their examination of star counts in the *GLIMPSE* Point Source Catalog of Spitzer. The GLIMPSE Point Source Catalog is similar to the DIRBE J and K bandpass data in that both are sensitive to cool evolved stars. Churchwell et al. proposed that the Galaxy is essentially a two-arm system, with two of the classically identified arms (Scutum-Crux and Perseus) representing a spiral density wave enhancement in older evolved star that originates at the ends of strong bar, and two of the four classically defined arms (Sagittarius-Carina and Norma-Cygnus) becoming minor structures. The tangent to the Sagittarius arm is not visible in the projected densities of K and M giants and supergiants, as derived from GLIMPSE point sources.

The two-arm models of both Drimmel and Spergel and Churchwell et al. are based upon the observations sensitive to the spatial distribution of K and M giant stars, a tracer that is not classically associated with spiral arms. Both works, along with that of Quillen and Minchev, suggest that the spatial distribution of these stars represents the spiral pattern that underlies the density wave enhancement of matter within the Galaxy and that the four arms that are classically defined represent the gas response to this underlying density wave.

A statistical examination of published models is consistent with this view. Forty-four of these have four arms and ten have two arms. Five of the two-arm models, and all but two two-arm models proposed since 1985, are based upon observations of K and M giants. At the same time, four-arm models are overwhelmingly based upon sources associated with gas, young stellar populations, and star formation. Only two of the four-arm models are based upon the distribution of evolved stars, and one of these [62] examined only a four-arm fit to the *IRAS* point source catalog; they did not rule out an equally good fit with a two-arm model.

Thus we are left with the conclusion that observations of young population spiral tracers (gas, dust, star-forming regions, OB associations, etc.) yield four spiral arms while observations of older, evolved stars lean heavily toward two-arm models.

Given its apparent lack of visibility in an older stellar population, and the key role this absence plays in the construction of two-arm models, it is important to examine the evidence for the Sagittarius arm in other spiral tracers. The *classic* defining property of spiral arms is the presence of star-forming regions. The large star-forming complex W51 is located at/near the commonly accepted tangent point of the Sagittarius arm. The W51 GMC is one of the most luminous and active regions of massive star formation in the Galaxy and falls among the top 1% of clouds in the Galactic disk by size and the top 5% by mass [9]. Anomalous gas velocities seen in the W51 region have long been attributed to streaming motions in the Sagittarius spiral arm [6, 49]. The Sagittarius arm was the first "arm" accepted as a large-scale spiral feature of the Galaxy; by the late 1960s there was little agreement concerning anything about the Galaxy's spiral structure, except for the existence of the Sagittarius arm [51]. Since that time, the Sagittarius arm has been seen in the spatial distribution of HII regions, maps of HI and CO, CO analyses of individual molecular clouds, synchrotron emission, OB stars and associated diffuse nebulae, molecular clouds, FIR emission from warm dust, Cepheids, Hα, H109α, radio continuum analyses of star-forming complexes, and [CII] and [NII] emissions. In addition, attempts to map the large-scale structure of the Galaxy's magnetic field by observations of the Faraday rotation measure of extragalactic sources are consistent with a reversal of the Galactic B-field associated with the Sagittarius-Carina arm.

In short, the Sagittarius arm is seen in essentially every tracer of spiral structure associated with a young population. Thus its existence as a major spiral arm, at least among those spiral tracers traditionally associated with spiral structure, i.e., star formation and enhanced gas densities, is not in question. Nonetheless, its signature is absent in the projected stellar densities of *Spitzer* and *COBE/DIRBE*. This presents an interesting puzzle for which there are two obvious solutions: (1) there is no enhanced stellar density in the tangency region because there is no stellar

spiral corresponding with the Sagittarius arm and (2) these stars are present in the tangency region but unusual absorption renders them difficult to see. Benjamin et al. [3] acknowledged both possibilities but settle on the former solution. Here we briefly examine each possibility.

It may be that older evolved stars in our Galaxy have a different spiral distribution than the gas rich spiral arms. In this case, two stellar arms may represent the true spiral density wave in the disk, while the four gas-rich arms represent the response of gas in the disk to the density perturbation presented by the stellar spirals. If so, the Milky Way's spiral structure would look different when young and old (blue versus red) objects are viewed separately. Indeed, a number of external galaxies reveal differences in their spiral structure when observed in the blue (young objects) representing the gas distribution, and in the near IR (older objects) representing the true mass distribution [4, 5].

Gas-dynamical simulations of the response of gas disks to an imposed two-arm spiral potential led Martos et al. [34] to suggest that it is possible for a four-arm gaseous spiral to coexist with a two-arm stellar spiral for specific regions of parameter space (e.g., pattern speed, magnetic field strength and geometry, and large-scale turbulence) [34, 35, 64]. While it is not clear that the Milky Way falls within the appropriate range of parameter space, this scenario offers a possible solution to the historical pattern of four-arm spiral models being derived from observations of spiral tracers associated with gas and two-arm models being derived from observations of the older stellar distribution.

On the other hand, this dichotomy may arise from unusual absorption in the Sagittarius tangency direction. The "star-count excesses" on which both the *DIRBE* [16, 17] and *Spitzer* [3, 11] two-arm models are based are small signals that could suffer appreciably from unexpected obscuration. Indeed, Benjamin et al. [3] note a gap in *GLIMPSE* Point Source Catalog-derived star-count excesses nearer the center of the Galaxy and they attribute this gap to obscuration.

A clear anticorrelation exists between star counts in the *GLIMPSE* Point Source Catalog and counts of unusually red sources [3, 13], which can be due to increased extinction or high diffuse background [3]. Both of these are associated with star-forming regions and molecular clouds. Given the presence of W51 at the Sagittarius tangency point, this issue warrants more close attention. In addition to its prominence as a star-forming region, W51 is a region of heavy obscuration [7, 24, 27, 40, 41]. Indeed, W51 is too obscured to be seen at visible wavelengths. The large size and unique location of W51 with respect to the Sun means that the line of sight toward W51 intersects the spiral arm over several kiloparsecs of path length [6, 49]. Suggestions of excessive obscuration in the near IR come from observations of infrared dark clouds (IRDCs) in the directions of W51. IRDCs are cold, below 13 K [7, 19, 56] and preferentially located toward spiral arms [8, 20, 41]. Much of the extinction toward the Sagittarius tangent may come from material lying in front of the W51 cloud complex [27].

It is clear that the possible effects of extinction and diffuse emission on source counts in the W51 region deserve future scrutiny. It is equally clear that the possibility that classic spiral arms associated with star formation may not represent the

underlying spiral form of the mass distribution provides a possible solution to a classic puzzle and also merits additional study.

4 Central Bar

Although not a spiral feature, the central bar of the Galaxy can have its physical characteristics constrained by the observations used in constructing spiral models or by the models themselves. Since the first suggestion of a bar-like structure in the Milky Way, debate has ranged over its size, shape, strength, and orientation. Some contend that the bar is weak (or even nonexistent), while others posit a bar structure that dominates the appearance of the Galaxy [3, 11].

In a precursor to the Churchwell et al. spiral model, analysis of *GLIMPSE* Point Source Catalog data by Benjamin et al. [3] led them to conclude that the Galaxy has a strong bar extending 4.4 kpc from the Galactic center and inclined by 40° relative to the Sun-Galactic center line. The conclusion is based on the Glimpse star-count densities plotted in their Figure 4.

While the *COBE/FIRAS* data used in constructing the SWH spiral model provide no direct measure of the Galaxy's central bar, the strong signal provided by the [CII] and [NII] emissions can provide constraints on the bar's size and orientation. Star formation is expected to occur at the ends of the bar, as seen in external galaxies and in numerical simulations. [CII] and [NII] emissions associated with these star-forming regions should show up in those directions aligned with the ends of the bar. [CII] and [NII] intensities both show significant enhancements at $+28°$ and $-25°$ Galactic longitude [54]. If these correspond with the approximate ends of a bar, then the half-length of the bar coupled with its orientation cannot project to a greater longitude than $+28°$ and $-25°$. This result is consistent with Glimpse star-count densities seen in figure 4 of Benjamin et al. [3].

An upper limit on the *physical* size of the bar follows from the four-arm spiral spiral model of SWH. Arms in this model arise $\sim 3 \times (R_{\odot}/8.5)$ kpc from the Galactic center. SWH argues that the bar does not extend beyond this physical radius. Thus, while in agreement on the projected limits of the bar, the SWH limit on the physical size of the bar is significantly smaller than that proposed by Benjamin et al. [3] and Churchwell et al. [11] based upon *Spitzer* observations.

Acknowledgments This work was supported in part by NASA grant NAG5-13112.

References

1. Amaral, L.H., Lépine, R.D. 1997, MNRAS, 286, 885
2. Bash, F.N. 1981, ApJ, 250, 551
3. Benjamin R.A., et al. 2005, ApJ, 630, L149
4. Bertin, G., Lin, C.C. 1996, Spiral Structure in Galaxies, a Density Wave Theory, MIT Press, Cambridge, MA

5. Block, D., Elmegreen, B., Wainscoat, R. 1996, Nature, 381, 674
6. Burton, W.B. 1970, A&A, 2, 291
7. Carey, S.J., Clark, F.O., Egan, M.P., et al. 1998, ApJ, 508, 721
8. Carey, S.J., Egan, M.P., Kuchar, T.A., et al. 2000, BAAS, 197, 516
9. Carpenter, J.M., Sanders, D.B. 1998, AJ, 116, 1856
10. Caswell, J.L., Haynes, R.F. 1987, A&A, 171, 261
11. Churchwell, E., Babler, B.L., Meade, M.R., Whitney, B.A., Benjamin, R., Indebetouw, R., Cyganowski, C., Robitaille, T.P., Povich, M., Watson, C., Bracker, S. 2009, PASP, 121, 213
12. Cordes, J.M., Lazio, T.J.W. 2003, preprint (astro-ph/0207156)
13. Cutri, R.M., et al. 2001, Explanatory Supplement to the 2MASS Second Incremental Data Release, Caltech, Pasadena
14. Dame, T.M., Elmegreen, B.G., Cohen, R.S., Thaddeus, P. 1986, ApJ, 305, 892
15. Downes, D., Wilson, T.L., Bieging, J., Wink, J. 1980, A&A Supp, 40, 379
16. Drimmel, R. 2000, A&A, 358, L13
17. Drimmel, R., Spergel, D.N. 2001, ApJ, 556, 181
18. Efremov, Y.N. 1998, A&AT, 15, 3
19. Egan, M.P., Shipman, R.F., Price, S.D., et al. 1998, ApJ, 494, L199
20. Egan, M.P., Carey, S.J., Price, S.D., et al. 1999, In: ESA SP-427: The Universe as Seen by ISO, ESA/ESTEC, Noordwijk, p. 671
21. Englmaier, P., Gerhard, O. 1999, MNRAS, 304, 512
22. Englmaier, P. 2000, RvMA, 13, 97
23. Fernandez, D., Figueras, F., Torra, J. 2001, A&A, 372, 833
24. Figuerêdo, E., Blum, R.D., Damineli, A., Conti, P.S., Barbosa, C.L. 2008, AJ, 136, 221
25. Georgelin, Y.M., Georgelin, Y.P. 1976, A&A, 49, 57
26. Ghosh, T., Rao, A.P. 1992, A&A, 264, 203
27. Goldader, J.D., Wynn-Williams, C.G. 1994, ApJ, 433, 164
28. Grabelsky, D.A., Cohen, R.S., Bronfman, L., Thaddeus, P., May, J. 1988, ApJ, 331, 181
29. Hammersley, P.L., Garzon, F., Mahoney, T., Calbet, X. 1994, MNRAS, 269, 753
30. Kerr, F.J., Hindman, J.V., Carpenter, M.S. 1957, Nature, 180, 677
31. Kerr, F.J., Kerr, M. 1970, ApJ, 6, L175
32. Lépine, R.D., Mishurov, Y.N., Dedikov, S.Y. 2001, ApJ, 546, 234
33. Levine, E.S., Blitz, L., Heiles, C. 2006, Science, 312, 1773
34. Martos, M., Hernandez, X., Yáñez, M., Moreno, E., Pichardo, B. 2004, MNRAS, 350, L47
35. Martos, M.A. 2008, BAAS, 40, 269
36. Mills, B.Y. 1959, PASP, 71, 267
37. Morgan, W.W., Sharpless, S., Osterbrock, D. 1952, AJ, 57, 3
38. Morgan, W.W., Whitford, A.E., Code, A.D. 1953, ApJ, 118, 318
39. Nakanishi, H., Sofue, Y. 2006, PASP, 58, 847
40. Nanda Kumar, M.S., Kamath, U.S., Davis, C.J. 2004, MNRAS, 353, 1025
41. Ormel, C.W., Shipman, R.F., Ossenkopf, V., Helmich, F.P. 2005, A&A, 439, 613
42. Ortiz, R., Lépine, J.R.D. 1993, A&A, 279, 90
43. Petrovskaya, I.V. 1987, Pisma AZh, 13, 474
44. Phillipps, S., Kearsey, S., Osborne, J.L., Haslam, C.G.T., Stoffel, H. 1981, A&A, 98, 286
45. Quillen, A.C., Minchev, I. 2005, AJ, 130, 576
46. Russeil, D. 2003, A&A, 397, 133
47. Russeil, D., Adami, C., Georgelin, Y.M. 2007, A&A, 470, 161
48. Sevenster, M.N. 1999, MNRAS, 310, 629
49. Shane, W.W., Bieger-Smith, G.P. 1966, BAN, 18, 263
50. Simard-Normandin, M., Kronberg, P.P. 1980, ApJ, 242, 74
51. Simonson III, S.C. 1970, A&A, 9, 163
52. Sofue, Y., Fujimoto, M. 1983, ApJ, 265, 722
53. Steiman-Cameron, T.Y., Wolfire, M., Hollenbach, D.J. 2008, BAAS, vol. 40, p. 213
54. Steiman-Cameron, T.Y., Wolfire, M., Hollenbach, D.J. 2010, ApJ, submitted

55. Taylor, J.H., Cordes, J.M. 1993, ApJ, 411, 674
56. Teyssier, D., Hennebelle, P., Pérault, M. 2002, A&A, 382, 624
57. Urasin, L.A. 1987, SvAL, 13(5), 356
58. Vallée, J.P. 1995, ApJ, 454, 119
59. Vallée, J.P. 2002, ApJ, 566, 261
60. Vallée, J.P. 2005, AJ, 130, 569
61. Vallée, J.P. 2008, AJ, 135, 1301
62. Wainscoat, R.J., Cohen, M., Volk, K., Walker, H. J., Schwartz, D.E. 1992, APJS, 83, 111
63. Weaver, H. 1970, In: The Spiral Structure of Our Galaxy, IAU Symp. 38, eds. W. Becker, G. Contopoulos, Reidel, Dordrecht, p. 126
64. Yáñez, M.A., Norman, M.L., Martos, M.A., Hayes, J.C. 2008, ApJ, 672, 207

The Mask of Complexity in Disk Galaxies

Daniel Pfenniger

Abstract One of the main objectives for a scientist is to sort out the complications of nature and to find out ways to describe its mechanisms with simple comprehensible terms. Disk galaxies belong certainly to the class of complex systems which, however, display regular patterns and correlations across their variety of morphologies. On the one hand, one needs to describe galaxies with complex models adapted to the huge amount of existing data and to the possibilities provided by computers, but on the other hand one needs also to develop deliberately simple models capturing the essential characteristics of galaxies; only with such a complexity reduction we can claim to understand what are galaxies. Thus complexity acts as a mask which can be sometimes removed, unveiling the hidden simplicity of galaxies.

1 Introduction

The more we study galaxies, the more complex they appear to us. Hierarchical galaxy formation over cosmic times and star formation in a multiphase interstellar medium in which dust grains play an important rôle, as well as black holes injecting huge energy amounts in their surroundings appear as an ensemble of extraordinary complicated phenomena acting at widely different time and spatial scales, for which most of the physics remains by far out of control.

Despite this and the fact that a substantial part of the mass, the dark matter, remains still to be effectively identified, on the other hand, we have the impression that we *understand* increasingly better galaxies, in the sense that we can better and better summarize their most essential nature by only a few sentences, capture the most general properties observed to occur in most of them.

For example, a one-sentence description of spiral galaxies could be "Disk galaxies are the largest rotating matter concentrations in the Universe hold by gravity, forming stars from gas, and able to keep most of the stellar activity ejecta and to recycle them in subsequent star generations."

D. Pfenniger (✉)
Geneva Observatory, University of Geneva, Sauverny, Switzerland
e-mail: daniel.pfenniger@unige.ch

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_4,

Actually, this process of sorting out simplicity from the nature complexity lies at the heart of the scientific activity. It is one of the major aims of science, because it increases the level of understanding of nature by humans, in other words it educates us. When we say that we *understand* a phenomenon, we mean that we have been able to forge a mental representation which discards most of the anecdotal complexity, while keeping its most relevant aspects. This simplified, abstracted mental representation of complex phenomenons can then be replayed in our brains as we think about them.

A theory, or a model, is essentially a scheme compressing complexity to a level that our brain (or the brain of at least another human) can manage. In such a view, a theory is not an absolute truth (because not containing the exact description of nature), but a best match simpler representation between our finite mental capacity to handle complex processes, and the much more complex reality. Over centuries we have come to realize that no theory is an absolute truth, since we know of no physical theory which is an exact representation of nature, including quantum mechanics and general relativity. Theories are not fundamentally different from models, only their scope is broader than one of models.

As an example illustrating the previous statements, for astronomical purposes the Earth is often described as a point mass, or as a perfect sphere, or as a spheroid. Such a drastically simplified description of course ignores that if we want to understand the Earth over its history, we should include not only geological processes such as plate tectonics but also biological phenomenons, since they have played a major rôle over gigayears in determining the structure and composition of the outer rock layers and atmosphere of our planet. But since such a detailed model is not necessary for calculating, say, the solar system ephemeris over million years, it is perfectly appropriate and sound to omit such irrelevant aspects, but to retain the essential ones. The choice of keeping and discarding parts of reality leading to a model or a theory is partly subjective and depends ultimately also on our particular finite mental capacities.

In some other situations, like describing a particular system such as the Milky Way, we do not need to really understand all the details, but to make detailed predictions. Formal tools like computer programs do allow us to extend the usual possibilities provided by mathematics to much larger complexity levels. Both better instruments and better computers make the collection and processing of huge amounts of information possible. In such cases, complex models are useful, but insufficient to improve understanding. Therefore, simultaneously we must work out a synthesis of this wealth of information, and find ways to compress most of this information to a much smaller size, in order to reach a level where we can pretend to *understand* what a galaxy is.

2 Build-Up of Complication

Before the computer age disk galaxies were described with simple models. Typically, they would be described as being steady, isolated, and axisymmetric smooth rotating mass distributions made of permanent point masses, point masses supposed

to mimic long-lived stars. The collisionless fluid approximation following Boltzmann's equation and Newtonian dynamics were the primary tools thought to be able to grasp most of the physics of galaxies.

This picture has progressively changed. Let us mention here a few illustrative works having shown over the years that the above simple model of disk galaxies as steady axisymmetric disks of stars is insufficient, since some of its assumptions could be shown to be invalid after a shorter than expected timescale.

The kinematic of Milky Way stars, as revealed by the Hipparcos satellite (e.g., [5]), showed in more details what had been slowly emerging since several decades: the velocity distribution of neighboring stars to the Sun, reduced to a local reference frame, was not a smooth ellipsoid, as expected from an axisymmetric well-relaxed stellar dynamical model of the Milky Way, but a clumpy structure in velocity space, including clumps of old stars. One could not invoke some remaining initial condition memory of common star formation to explain all these streams. Non-axisymmetric structures in the Milky Way and time-dependence could no longer be discarded. The Milky Way bar had been shown to exist [3], but the degree of time dependence was still to be appreciated.

Bars had been more and more recognized in the previous decades to be the normal state of a disk galaxy, instead of the converse. Sellwood and Sparke [18] showed that time-dependence was also to be expected as normal even in a frame rotating with the bar, because the pattern speed of a bar in simulations was higher than the pattern speed of the surrounding spiral arms. Necessarily the phase difference between the bar and the typical four surrounding arms varies in time. When the bar is in phase with a pair of arms, these arms are reinforced due to the superposition of the bar and arm potential wells, and conversely the other pair of arms is weakened. When the bar is out of phase with both arm pairs, the four arms have similar, intermediate amplitudes. The consequence that can be expected from this configuration is that star and molecular cloud formation in arms should follow a periodic modulation with a frequency corresponding to four times the difference between the bar and spiral arms angular pattern frequencies.

The spiral arms were initially believed to be small density perturbations, until near-infrared observations revealed that the true stellar mass contrast, mostly hidden by dust, is actually substantial (e.g., [7]). In typical spiral galaxies the arm–interarm density contrast with respect to the mean is of order of 1, which indicates that spiral arms are strong nonlinear perturbations where the collective mass participates to building the bar/spiral structure. A spiral galaxy as understood now is not strictly time-independent, but necessarily a secularly evolving disk due, at the least, to this bar–spiral arms interactions.

Mergers and galaxy interactions also add further time-dependence of a more erratic kind. The discovery of streams in the Milky Way and nearby galaxies halos (e.g. [11]) showed that stellar halos are not virialized and that each galaxy can follow substantially different histories. The interesting point is that, despite different histories, galaxies manage to build up strong correlations in their asymptotic dynamical and stellar properties, which indicate that some attractors must be present and searched for.

Fux [9, 10] made Milky Way N-body models reproducing as well as possible the rich features seen in the HI and CO gas. Since the gas distribution is extremely clumpy, especially the molecular gas (e.g., [4]), the fits of the highly structured gas made by Fux represent an extremely transient situation lasting, as seen from the Sun position, only a few tens of Millionyear at the level of detail shown by both the simulations and the observations. At the level of the gas details, it is now clear that the clumpy structures in the interstellar medium evolve fast with Myr timescales, they must be seen like the eddies in a turbulent fluid, i.e., highly transient structures, with the peculiarity that self-gravity, with its anti-thermodynamical behavior leading to negative specific capacity, is never completely negligible from spiral arms to star forming regions, which contributes to make such a medium behaving unlike fluids commonly observed on Earth.

The continuous growth of computer capacity over 60 years, where every 5 years the performance of computers has grown by an order of magnitude (a variant of Gordon E. Moore's law, [15]), is a historically unprecedented technological revolution. The performance of the top computers of the moment becomes easy to access for average scientists 5–15 year later, and for the general public 20–30 year later. This revolution leads to major changes throughout the society. For sciences demanding complex models, like biology and astronomy, computer programs, a formal description of a sequence of deterministic logical operations, play more and more the rôle that mathematics has fulfilled over last centuries for physics. Both mathematics and computer models allow in complementary ways to represent and manipulate in a deterministic way (hopefully) the part of reality than we wish to describe.

The continuous exponential growth of computers leads to a paradigm change concerning how to represent the dynamics of galaxies. The collisionless fluid description with Boltzmann's equation is indeed unable to describe the force fluctuations induced by the different mass concentrations in a real galaxy, like molecular clouds or even the stars. A number of stars like 10^{11} are indeed not very large for a six-dimensional phase space in order to represent a smooth, differentiable flow. Namely, if phase space is binned by cells containing each of the order of 100 stars to represent a reasonably smooth fluid, then the number of bins per phase space coordinate is rather low: $(10^{11}/100)^{1/6} \approx 32$. Today N-body simulations reach $N \sim 10^{10}$ (e.g., [2, 19]). So following G. Moore's law, in the next 5 years a number like $N = 10^{11}$ of individual mass points should be increasingly accessible. We conclude that the best formal representation of galaxy dynamics will shift from the traditional Boltzmann continuum description to the full granular N-body description where each body represents a galaxy mass fraction comparable to the one of a star. We see also that in the future gravitational physics will be less and less the main difficulty preventing to understand galaxies, but the physics associated with either gas, stars, or black holes, and the physics associated with dark matter.

In conclusion, we see that a disk galaxy is a much more complex system than thought initially. We need complex computer models to reproduce some of its main characteristics. However, despite the simplistic description of earlier models, they were and still represent useful reductions of complexity. We now know that galaxies are never strictly isolated, but the isolation concept is useful to abstract the ideal

galaxy that would only depend on internal processes, the most relevant ones for understanding the galaxy global properties. We know that time-dependence is necessary to understand the formation and long-term evolution of galaxies, but time-independent models capture the concept of equilibrium, for which the virial theorem follows, where gravitational and twice the kinetic energy are in balance. From that we deduce the useful rule that the equilibrium of more complex systems is determined to first order by the balance of the dominant forms of energy densities and that any imbalance of these energies provides a timescale of evolution.

3 Modeling

Thus the modeling work may be divided in at least three phases.

1. For a poorly understood problem at early stage of understanding, like galaxies in the first half of the twentieth century, it is appropriate to make as simple as possible models to see if they capture at least some basic properties of the studied objects. The model simplicity allows to easily identify which aspects are too rough and should be refined.

   Here the early papers of Ken Freeman [6–8], included in his PhD thesis, provide an excellent example of such an exploratory modeling work on objects that were not understood at all at the time, barred galaxies. The objective was to understand the structure and evolution of barred spiral galaxies using analytical models of self-gravitating collisionless systems. In the first paper [6], barred galaxies were described as rotating self-gravitating homogeneous elliptical cylinders of infinite height. In the second paper [7], the bar geometry was modified as triaxial homogeneous ellipsoids, while in the third paper [8] as thin inhomogeneous elliptic sheets. Exploring bar models of widely different geometries, even if looking crude with respect to real barred galaxies, was useful to delineate if conclusions were robust, independent of the considered geometry variations. These analytical models remain among the rare ones able to capture the essential characteristics of a bar of a barred galaxy, i.e., a self-gravitating system where gravity is balanced not only by the centrifugal and pressure forces but also by the Coriolis force. It is remarkable that in such an early work, from the insight gained by the study Freeman was already able to argue: “if there is an evolution pattern inherent in the Hubble classification, the direction of evolution would be SBc → SBa.” The by-product of such a ground work, despite its simplicity was to give an early insight into the physics of barred galaxies and about their possible evolution.

   So we see that simple models have their own virtue, they are closer to what our brain is able to manage well and allow to develop intuition.
2. In a second phase, better observations and more powerful computers allow to refine the models. As the number of constraints increases, the models grow in complexity and hopefully become closer matches of reality. These models allow to better manipulate and test our understanding of complex systems, such as galaxies. As the model sophistication grows more and more these models

resemble the natural systems. As we do with natural systems we observe the model outputs and build up some familiarity by finding out correlations. At any stage in this complexification process, the modeling work is incomplete if we do not gain more insight about the modeled systems.

3. Therefore, in a third phase we need to use what has been learned in the more complex models to discard anecdotal features and to keep the most essential aspects. This complexity reduction process has for purpose to improve our understanding of the studied objects. We can claim to understand a system when we can figure out *without computer* how it behaves.

4 Toward Simplified Models of Disk Galaxies

Until about the second half of the twentieth century, the interstellar gas and the dust were either ignored or considered as negligible components. Therefore dynamics was considered as decoupled from the dissipative processes related to stars. Galaxy dynamics was therefore viewed as non-dissipative.

Soon after computers became available they have been used to model disk galaxies (e.g., [12]). For the next decades the gravitational physics was perceived as the most difficult aspect of galaxies to better understand. Since no tool like statistical mechanics exists for self-gravitating systems, numerical works, such as N-body simulations, were at the forefront of research and allowed to progress enormously.

Today we stand at a stage where further progress needs no longer to increase gravitational resolution (by increasing N), but in priority to improve the additional physics besides dynamics and its impact on the whole galaxy over time. Infrared observations have shown that a substantial part of the stellar light is thermalized and diverted by dust. Quantitatively this energy flux, a power of the order of a sizable fraction of the total luminosity, is large enough to modify the size of a galaxy over Gyr, since it amounts to a magnitude comparable to the galaxy gravitational energy divided by its rotational period [14], and this across the whole spiral sequence.

Once we understand that galaxies are not conservative structures over their lifetime, but actually dissipative structures which continue to dissipate energy at least as long as gas exists and star formation proceeds, we have ground to expect that *attractors* exist which explain the similarities and correlations of disk galaxies despite very different histories.

A basic example of attractor is disks, seen to form in widely different astrophysical conditions. Disks result from a better conservation of angular momentum in regard of energy for self-gravitating structures. Different conditions then determine the differences between large-scale disks like galaxies from small-scale disks such as proto-planetary disks.

Therefore to understand better galaxies, we need to describe attractors beyond the disk attractor and the factors making these attractors possible. If energy is dissipated, a natural way to look for attractors is to describe disk galaxies as systems tending asymptotically toward an energy minimum. However, a problem with minimizing

energy in gravitational systems is that gravitational energy is not bounded from below. In Newtonian dynamics massive points or curves are singular objects having exported an infinite amount of energy. So to keep energy minimization physical, processes that prevent gravitational energy to diverge must be included.

A standard trick in physics is to assume that the class of allowed functions is everywhere sufficiently differentiable, which automatically excludes wild local behaviors. The physical justification behind that is granted when microscopic physics takes care to prevent singularities to grow and smooth out irregularities. The exact way this smoothing occurs is irrelevant at macroscopic scale, provided it exists. Thus using differentiable functions is a convenient way to hide microscopic complications and to produce simple and efficient macroscopic models.

Thus if the energy of a thin disk with nonzero angular momentum balanced by rotation is minimized without constraining the boundary conditions, a disk with flat rotation curve (Mestel's disk) turns out to be the solution of this minimization problem [13]. The problem can be formulated as the total energy being a functional of the surface density, *assuming a differentiable surface density*. This is an encouraging result, but clearly insufficient, since real disks need to be both supported by rotation and kinetic pressure to be stable.

A dissipative disk tends to decrease its gas pressure of the star velocity dispersion. This has the effect of bringing the disk toward two kinds of dynamical instabilities before being infinitely cold and thin. A too cold disk in the radial direction is unstable with respect to radial instability [16, 20] and a too thin disk is unstable with respect to a fire-hose like instability [1]. When a disk is driven slightly beyond the verge of instability it becomes reactive to perturbations, and the reactions tend to counteract the dissipation effect. The natural state for a dissipative galaxy is therefore to keep a *marginal state* with respect to such instabilities, constraining the states of disks.

A marginally stable disk also in both the radial and the transverse directions *explains* why disk galaxies display regularly spiral arms and warps while they contain gas and form stars: disk galaxies are then particularly reactive to any kind of perturbations, internal or external. The common cause of spiral arms and warps is then identified with a single rule: dissipation in a self-gravitating disk leads to a marginally stable state.

We intend to study more the consequences of these constraints on an energy minimization of self-gravitating disks. The aim is to build simple models allowing to reduce the apparent complexity of disk galaxies.

5 Conclusions

The disk galaxy models have evolved through several phases. First models, such as the ones developed by Ken Freeman in his PhD thesis, were deliberately kept simple in order to learn the basic properties of these systems. Such works are useful to develop insight. Subsequently, models underwent a complexification made nec-

essary in view of the large amount of observational data and the increased computational power made available by the information technology revolution. Now, there is a need to reconsider these complex models in order to identify the most basic factors determining the common features of disk galaxies, the attractors. A self-gravitating disk is an example of attractor appearing in very different conditions independently of the initial conditions. Attractors in general allow to obtain simple asymptotic descriptions, or explanations, of apparently complicated natural systems. In such situations complexity acts as a mask which can be sometimes removed when the attractors are identified.

References

1. Araki, S. 1985, PhD Thesis, Massachusetts Institute of Technology
2. Boylan-Kolchin, M., Springel, V., White, S.D.M., Jenkins, A., Lemson, G. 2009, Monthly Not Royal Astron Soc 398, 1150
3. Blitz, L., Spergel, D.N. 1991, Astrophys J 379, 631
4. Dame, T. M., Hartmann, D., Thaddeus, P. 2001, Astrophys J 547, 792
5. Dehnen, W. 1998, Astron J 115, 2384
6. Freeman, K.C. 1996a, Monthly Not Royal Astron Soc 133, 47
7. Freeman, K.C. 1996b, Monthly Not Royal Astron Soc 134, 1
8. Freeman, K.C. 1996, Monthly Not Royal Astron Soc 134, 15
9. Fux, R. 1997, Astron Astrophys 327, 983
10. Fux, R. 1999, Astron Astrophys 345, 787
11. Ibata, R., Chapman, S., Ferguson, A. M. N., Lewis, G., Irwin, M., Tanvir, N. 2005, Astrophys J 634, 287
12. Lindblad, P. O. 1960, Stockholms Observatoriums Annaler 21, 4
13. Pfenniger, D., 1989, Astrophys J 343, 142
14. Pfenniger, D. 1991, in “Dynamics of Disc Galaxies”, Varberg Castle Sweden, 389
15. Moore, G.E. 1965, Electronica 38, 8
16. Safronov, V.S. 1960, Annales d’Astrophysique, 23, p.979
17. Seigar, M. S., James, P. A. 1998, Monthly Not Royal Astron Soc 299, 672, and 685
18. Sellwood, J. A., Sparke, L. S. 1988, Monthly Not Royal Astron Soc 231, 25
19. Teyssier, R., Pires, S., Prunet, S., Aubert, D., Pichon, C., Amara, A., Benabed, K., Colombi, S., Refregier, A., Starck, J.-L. 2009, Astron Astrophys 497, 335
20. Toomre, A. 1964, Astrophys J, 139, 1217

Cosmic Magnetic Fields – An Overview

Richard Wielebinski and Rainer Beck

Abstract Magnetic fields have been known in antiquity. Aristotle attributes the first of what could be called a scientific discussion on magnetism to Thales, who lived from about 625 BC. In China "magnetic carts" were in use to help the Emperor in his journeys of inspection. Plinius comments that in the Asia Minor province of Magnesia shepherds' staffs get at times "glued" to a stone, a lodestone. In Europe the magnetic compass came through the Arab sailors who met the Portuguese explorers. The first scientific treatise on magnetism, "De Magnete", was published by William Gilbert who in 1600 described his experiments and suggested that the Earth was a huge magnet. Johannes Kepler was a correspondent of Gilbert and at times suggested that planetary motion was due to magnetic forces. Alas, this concept was demolished by Isaac Newton, who seeing the falling apple decided that gravity was enough. This concept of dealing with gravitational forces only remains en vogue even today. The explanations why magnetic effects must be neglected go from "magnetic energy is only 1% of gravitation" to "magnetic fields only complicate the beautiful computer solutions". What is disregarded is the fact that magnetic effects are very directional (not omni-directional as gravity) and also the fact that magnetic fields are seen everywhere in our cosmic universe.

The Interstellar Matter (ISM) is ionized, hence cosmic magnetic fields are quite easy to generate. Magnetic fields have been measured in or around practically all celestial objects, either by in-situ measurements by spacecrafts in the Solar system or by remote sensing of cosmic objects. The Earth, the Sun, solar planets, stars, pulsars, the Milky Way, nearby galaxies, more distant (radio) galaxies, quasars, and even intergalactic space in clusters of galaxies have significant magnetic fields. Information on cosmic magnetic fields has increased enormously as the result of the rapid development of observational methods, especially in radio astronomy. In the Milky Way, a wealth of magnetic phenomena was discovered, which are only partly related to objects visible in other spectral ranges. The large-scale structure of the Milky Way's magnetic field is still under debate. The available data for external

R. Wielebinski (✉)
Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
e-mail: rwielebinski@mpifr-bonn.mpg.de

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_5,

galaxies can well be explained by field amplification and ordering via the dynamo mechanism. The measured field strengths and the similarity of field patterns and flow patterns of the diffuse ionized gas give strong indication that galactic magnetic fields are dynamically important. In spite of our increasing knowledge on magnetic fields, many important questions on the origin and evolution of magnetic fields, their first occurrence in young galaxies, or the existence of large-scale intergalactic fields remained unanswered. The present upgrades of existing instruments and several planned radio astronomy projects have defined cosmic magnetism as one of their key science projects.

1 Introduction

While the astronomical community developed the theory of gravitation to explain the observed Universe physicists and engineers pursued the synthesis of the electromagnetic theory. An understanding of the relationship between electricity and magnetism began in 1819 with work by Hans Christian Oersted, who discovered that an electric current could influence a compass needle. Several other experiments followed, with André-Marie Ampère, Carl Friedrich Gauss, Michael Faraday, and others finding further links between magnetism and electricity. James Clerk Maxwell synthesized and expanded these insights into Maxwell's equations, unifying electricity, magnetism, and optics into the electromagnetic fields. Another monumental discovery was the propagation of electromagnetic waves through space, an experiment of Heinrich Hertz, the basis for modern communication. On the engineering side the invention of the dynamo by Werner von Siemens must be mentioned.

The measurement of magnetic fields was not a simple matter. Carl Friedrich Gauss (the unit still in use in astronomy) and Nicola Tesla developed methods and gave rise to units named after them. An important discovery in the quest of a remote sensing methods to measure magnetic fields were the experiments of Pieter Zeeman who in 1896 showed the splitting of a line emission by strong magnetic fields. The first report of a cosmic magnetic field outside the Earth was the result of a direct measurement of the Zeeman effect by E. Hale in the magnetic fields in sunspots of the Sun in 1908. The discovery of radio emission by Karl Jansky in 1932 gave rise to speculations about their origin. Fermi, Alfvén, Kiepenheuer, Shklovsky, Ginsburg et al. realized that magnetic fields were responsible for this new window on the Universe. It was also clear that the observed cosmic rays would require magnetic fields for their creation and their containment within the Galaxy.

Optical polarization observations were first successful in 1949. Polarization of optical and infrared emission can also be caused by elongated dust grains which are aligned in magnetic fields as interpreted by the Davis–Greenstein mechanism in 1951. This interpretation was not accepted for a long time in the optical astronomy community. With the advent of radio astronomy this controversy could be resolved in favor of magnetic fields and an active study of magnetic fields could begin. The theory of synchrotron emission theory that was developed pointed out that synchrotron emission should be highly polarized. In fact, in homogenous magnetic fields, up to 75% linear polarization of the continuum emission can be expected. This suggestion was taken up by observers of optical radiation who found in 1954

that the Crab Nebula was highly polarized and hence emitting light through the synchrotron process. The radio confirmation of the polarization of the Crab Nebula followed in 1957. The first definite detection of the linear polarization of the Galactic radio waves was published by in 1962. At the same time the polarization of the bright radio galaxy Cygnus A and the Faraday rotation of the polarization angles of the linearly polarized radio emission in Centaurus A were detected. Observations at two frequencies of a section of the Milky Way showed that the interstellar medium of the Milky Way can also cause Faraday effect. During this exciting time of definite detections of interstellar and extragalactic magnetic fields by observations of linear polarization, the Zeeman effect of radio spectral lines proved to be more elusive. Several groups attempted to measure magnetic fields by this direct method. It was in 1968 that finally the Zeeman effect at radio wavelengths was successfully observed in the absorption profile of the H I line in the direction of Cassiopeia A. From this time onward considerable data were collected on the distribution of magnetic fields in the Milky Way.

The first suggestions about the presence of magnetic fields in nearby galaxies were made in 1958, based on observations of the polarization of stars in the Andromeda galaxy, M 31. In 1967 observations of the linear polarization of diffuse starlight were successful in several bright nearby galaxies. In 1970 the polarization of stars in the Magellanic Clouds implied the presence of magnetic fields in these neighboring irregular galaxies. Low-frequency radio observations of galaxies showed non-thermal spectra and hence indicated the presence of magnetic fields. The first detection of the linear polarization of the radio emission from nearby galaxies in 1972 led the way to massive improvement on our knowledge of the morphology of magnetic fields in galaxies. These early radio observations were in good agreement with the earlier optical polarization studies of galaxies.

The presence of magnetic fields in the distant Universe is suggested by observations of clusters of galaxies and of very distant, high-z sources. In fact nearly every high-z record (now at $z \sim 7$), invariably leads to a detection of non-thermal radio emission and hence proves the presence of magnetic fields. Gamma ray bursts, from the distant Universe, can be detected as weak non-thermal radio sources. Observations of nearby clusters of galaxies showed extended radio emission that must be due to the presence of inter-galactic magnetic fields.

In this talk, the status of our knowledge about the magnetic fields in our Milky Way and in nearby star-forming galaxies is summarized. We present recent results of cosmic magnetic fields. Also the present discussion about the origin of magnetic fields in galaxies will be briefly discussed. The reader is furthermore referred to more extensive reviews [6, 8, 62].

2 Observational Methods

As the methods of measuring of magnetic fields have been discussed widely in the literature, we present only a short summary of the methods to ensure that the present limitations become clear. Figure 1 shows the most used observational methods.

Elongated dust grains can be aligned with their major axis perpendicular to magnetic field lines, the Davis–Greenstein effect. This is the basis to measure magnetic

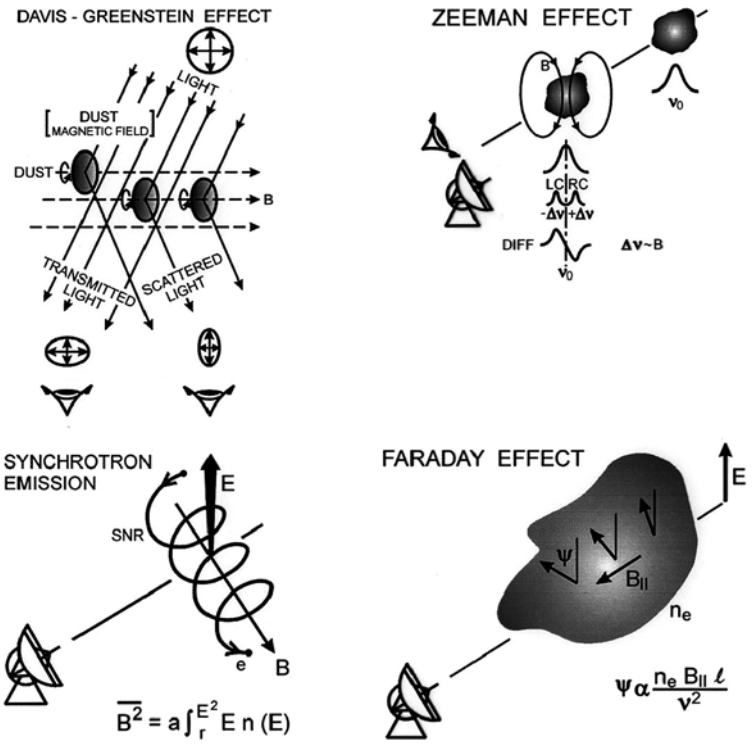


Fig. 1 The four major remote observational methods of studying magnetic fields

fields with optical and near-infrared polarization, by observing individual stars or of diffuse starlight. Light can also be polarized by scattering, a process unrelated to magnetic fields. This contamination is small when observing stars but needs to be subtracted from diffuse light, requiring multi-color measurements. In the far-infrared (FIR) and sub-millimeter wavelength ranges, the emission of elongated dust grains is intrinsically polarized and scattered light is negligible. If the grains are again aligned perpendicular to the magnetic field lines, the E-vectors point perpendicular to the field. Charged particles (mostly electrons) when moving at relativistic speeds (cosmic rays) around magnetic fields lines on spiral trajectories generate electromagnetic waves by the synchrotron process. Cosmic rays in interstellar magnetic fields are the origin of the diffuse radio emission from the Milky Way. The widely used minimum-energy (equipartition) estimate of the field strength is used, based on the measurement of total intensity, no direct method is available. Linear polarization is a distinct signature of synchrotron emission. The emission from a single electron gyrating in magnetic fields is elliptically polarized. An ensemble of electrons shows only very low circular polarization, but strong linear polarization with the plane of the E-vector normal to the magnetic field direction.

The linearly polarized radio wave is rotated by the Faraday effect in the passage through a magneto-ionic medium. This effect gives us another method of

studying magnetic fields – their regular component along the line of sight. As Faraday rotation angle is sensitive to the sign of the field direction, only regular fields can give rise to Faraday rotation, while anisotropic and turbulent fields do not. Measurements of the Faraday rotation angle from multi-wavelength observations allow determination of the strength and direction of the regular field component along the line of sight. Its combination with the total intensity and the polarization "vectors" yields in principle the three-dimensional picture of galactic magnetic fields and can separate the three field components: regular, anisotropic, and turbulent. The use of Rotation Measure synthesis [12] allows this.

The Zeeman effect is the most direct method of remote sensing of magnetic fields. It has been used in optical astronomy since the first detection of magnetic fields in sunspots of the Sun. The radio detection was first made in the H I line in absorption. More recent observation of the OH or H_2O lines used the larger frequency shifts of these molecular line species. At millimeter wavelengths further molecular lines can be used.

3 Magnetic Fields in the Milky Way

The earliest optical polarization observations by Hiltner and Hall in 1949 were interpreted to be due to dust alignment in magnetic fields and hence a tracer of magnetic fields in galaxies. The interpretation of dust alignment by magnetic fields was given by Davis and Greenstein [16]. It took some time to convince the optical community that the polarization was due to dust grains aligned in magnetic fields. However, the radio polarization observations confirmed the magnetic explanation. Based on an all-sky catalogue of Mathewson & Ford [41] with 1,800 entries a conclusion could be drawn: magnetic fields align along the Galactic plane with some spurs normal to the plane. These early observations were possible for nearby stars, a few at a maximal distance of 4 kpc. More recent compilation includes some stars out to $\approx$ 8 kpc. In view of these distance limitations it is not possible on the basis of optical polarization alone to model the magnetic field of the Milky Way. Polarization observations of the diffuse far-infrared or sub-mm emission in the Milky Way are restricted to dense molecular/dust clouds. New instruments on high sites hold the possibility of expanding such results using this spectral range.

The radio continuum emission of the Milky Way and star-forming galaxies at frequencies below 10 GHz mostly originates from the synchrotron process and hence traces the distribution of magnetic fields and cosmic rays. The contribution of thermal radio emission is generally small, except in bright star-forming regions. Only at frequencies higher than 10 GHz the thermal emission may dominate locally. At frequencies below about 300 MHz absorption of synchrotron emission by thermal gas can become strong. Hence the observation of total radio continuum intensity in the frequency range of about 300 MHz–10 GHz is a perfect method to investigate magnetic fields. Since the observed intensity is the integral from many emission areas along the line of sight, its interpretation is not always simple. Furthermore, the angular resolution of all-sky surveys is limited and hence cannot show the details of

extended sources. Numerous radio continuum surveys have been made in the early days of radio astronomy. These all-sky surveys showed the Galactic emission with a maximum toward the Galactic center, the band of emission along the Galactic plane, maxima in the tangential directions of the local spiral arm: Cygnus ($l \approx 80°$) in the northern and Vela ($l \approx 265°$) in the southern skies and some "spurs" of emission. In addition a few strong extragalactic sources were seen superposed on the Galactic emission. The analysis of total synchrotron emission gives an equipartition strength of the total field of $6 \pm 2\,\mu$G in the local neighborhood and $10 \pm 3\,\mu$G at 3 kpc radius. The radial scale length of the total field is about 12 kpc. These results are similar to those in external galaxies.

The angular resolution has improved so that at present all-sky surveys with resolution of under 1° are available. The WMAP satellite surveys at frequencies from 23 to 94 GHz gave us a new view of the radio continuum sky at high radio frequencies. At the highest WMAP frequencies we see mainly thermal emission originating in interstellar dust. An additional component due to spinning dust has been postulated be seen in the 10–100 GHz frequency range. This spinning dust component has recently been confirmed in the WMAP data set. There is a large gap between the lower frequency all-sky surveys and the high-frequency data. A 5 GHz all-sky survey with compatible angular resolution but also good sensitivity is badly needed.

Linear polarization of the continuum emission is a more direct indicator of magnetic fields, because there is no confusing thermal component. However, linear polarization is subject to Faraday effects. After the first detections of polarized Galactic radio waves in 1962 several all-sky polarization surveys were made. Major progress was achieved by Wolleben et al. [65] and Testori et al. [58] who mapped the whole sky in linear polarization at 1.4 GHz with an angular resolution of $36'$. Several polarization maxima are seen, e.g., toward the "Fan region" at $l \approx 140°$, $b \approx 10°$, where we look perpendicular to the local spiral arm. The "North Polar Spur" (NPS) emerges from the Galactic plane at $l \approx 30°$ as well as additional spur-like features are the results of magnetic fields compressed by expanding supernova remnants. In particular, the NPS can be followed, in polarization, to the southern sky. Toward the inner Galaxy (Galactic longitude $90° > l > 270°$, Galactic latitude $|b| < 30°$) strong turbulence in the polarized intensity is seen, due to Faraday effects on small scales. All-sky polarization data at 23 GHz were published by the WMAP team (e.g., [31]). There is good agreement between the 23 GHz and the 1.4 GHz polarization maps in the polarization features away from the Galactic plane, but the high-frequency map shows less Faraday depolarization toward the inner Galaxy and near the plane (see Figures 2 and 3).

Galactic plane surveys have been made from the earliest days of radio astronomy to delineate the extended Galactic sources like supernova remnants and H II regions, usually with no linear polarization data. Many of the published Galactic plane surveys between 22 MHz and 10 GHz cover only a narrow strip along the Galactic plane in the inner Galaxy. Total intensity surveys at several frequencies were used to separate the thermal H II regions (with a flat radio spectrum) from the steep-spectrum

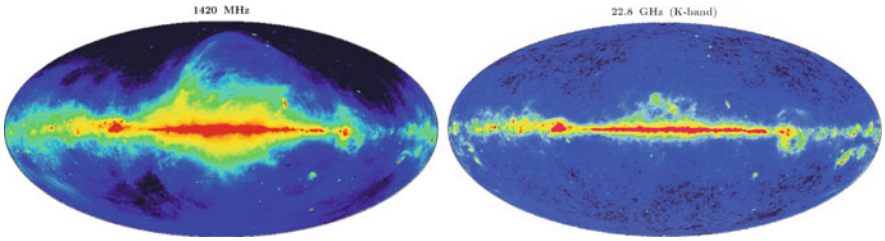


Fig. 2 Ground-based all-sky surveys exist from 45 to 1,420 MHz. Space-based all-sky surveys (WMAP) exist from 22.8 to 93.5 GHz. *Left*: 1420 MHz survey from P. Reich, MPIfR [48]; *right*: 22.8 GHz survey from WMAP project [36]

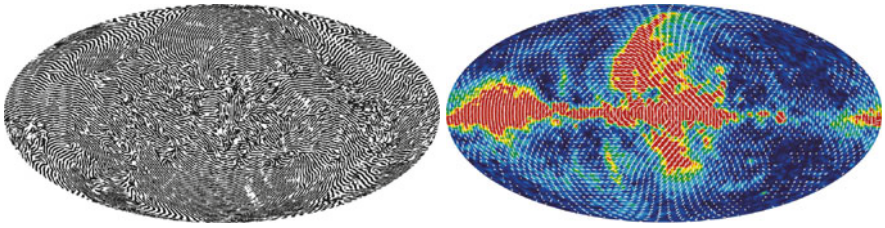


Fig. 3 All-sky surveys in polarized intensity (left B-vectors [64]) at 1.4 GHz and at 23.3 GHz (E-vectors right). Note the turbulence due to magnetic field toward the inner Galaxy. The 1.4 GHz diagram from M. Wolleben, DRAO, 23.3 GHz map from Hinshaw et al. [31]. The 1.4 GHz image was created by "alice" algorithm from David Larson, John Hopkins University

non-thermal sources (supernova remnants). Even at 5 GHz Faraday rotation plays an important role near the Galactic plane ($|b| < 5°$). With high enough angular resolution, Faraday rotation leads to complete depolarization at certain values of Faraday rotation measure (RM) showing up as "canals" in the maps of polarized intensity. However, a careful determination of the extended polarized background is necessary for a reliable determination of polarized intensity, polarization angles and RM. When this "absolute calibration" is done, most of the "canals" may disappear. An arc minute resolution survey of the Galactic plane made at DRAO at 1.4 GHz was combined with data from 25- to 100-m telescopes to achieve best possible absolute calibration [38]. As a second new phenomenon, "Faraday Screens" were discovered. These are foreground clouds of diffuse thermal gas and magnetic fields which Faraday-rotate or depolarize the extended polarized emission from the background. In addition to the well-known polarized SNRs and un-polarized H II regions, molecular clouds, pulsar-wind nebulae, and planetary nebulae were identified as Faraday Screens. Depending on the rotation angle and the polarization angle of the background emission, such screens may appear bright or dark. The strength and structure of regular fields can be estimated via the RM. Such observations can trace magnetic structures to sub-parsec scales. Sensitive surveys at higher radio frequencies are needed to allow a systematic study of the Faraday Screen phenomenon.

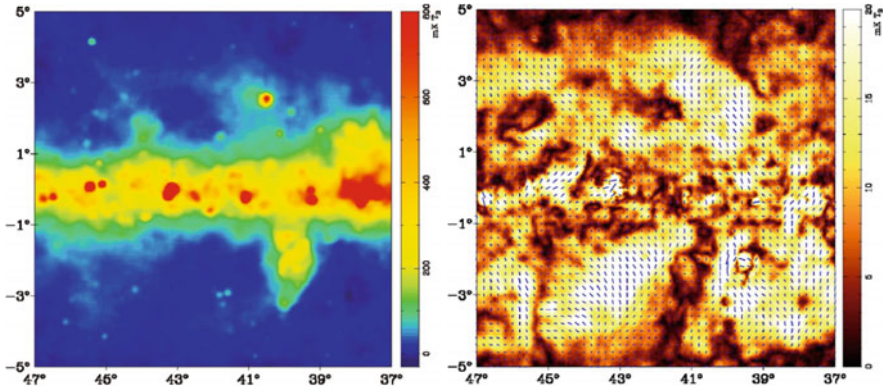


Fig. 4 A section of the Sino–German λ6 cm Urumqi survey, courtesy of X. Sun, NAOC, Beijing, China. *Left*: total intensity, *right*: polarized intensity [57]

This is being achieved, e.g., by a Sino–German survey of the Galactic plane at 5 GHz (see Figure 4) with the Urumqi telescope [24, 56].

The Galactic Center is unique source with unusual radio continuum features. Mapping of the Galactic Center region by Yusef-Zadeh et al. [66] showed several features vertical to the plane. The radio continuum emission is most intense and has a flat spectral index [49], for many years accounted to thermal emission. However, this intense emission is highly polarized (e.g., [54]) and has been interpreted to be due to mono-energetic electrons. Also the polarization "strings" imply vertical magnetic structures, much different from the azimuthal directions of the magnetic fields seen along the Galactic plane. The discussion about the intensity of the magnetic fields has yielded very high values (mG range) based on the high rotation measures. A recent interpretation of the magnetic field phenomena in the Galactic Center was given by Ferrière [20].

Faraday rotation (FR) is a powerful tool for studying magnetic fields (see Figure 5). First, ionospheric rotation, later Faraday effects due to the Galactic ISM were detected soon after the discovery of linear polarization of the Galactic radio waves. At first the FR of diffuse emission was studied. Later with increasing samples of Extragalactic Radio Sources (EGRS) modeling of the magnetic field was attempted. Finally pulsars, most of which are concentrated to the Galactic plane were used to model the Galactic magnetic fields. Pulsars are the ideal sources to probe the magnetic fields through the Faraday effect. Since pulsars have no measurable angular structure and they are highly polarized and they are the ideal probes. Pulsars are Galactic objects and hence their distribution is close to the Galactic plane toward the inner Galaxy. In fact very few pulsars are known toward the anti-center of the Galaxy. Hence a combination of pulsars and EGRS is optimal for studies of the Galactic magnetic field. With pulsars we can also measure the Dispersion Measure (DM) and hence the value of the average regular magnetic field in the line of sight can be deduced.

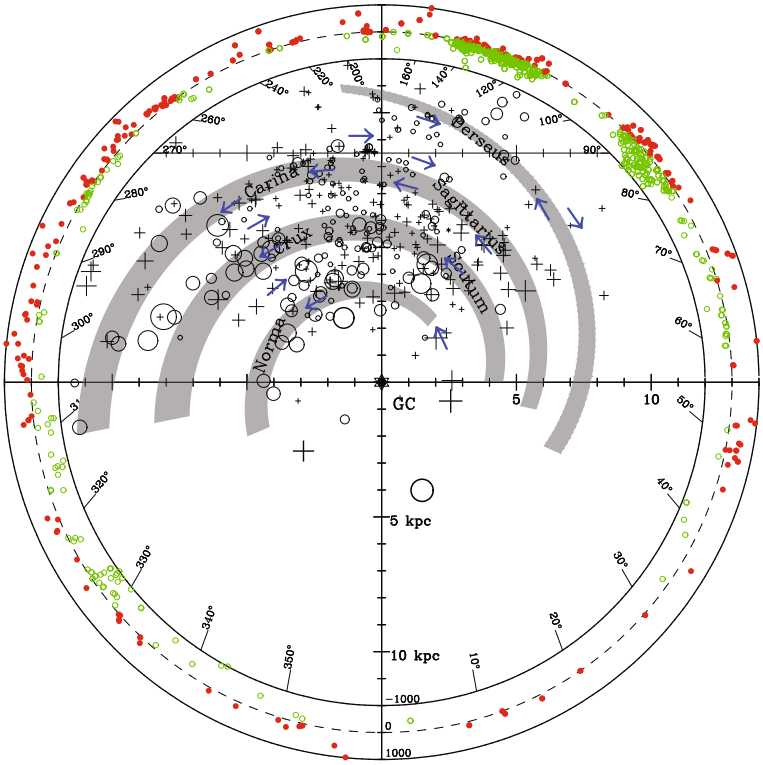


Fig. 5 Faraday rotation measures (RM) of pulsars in the Milky Way (within inside circle) and of extragalactic radio sources (between *inside* and *outside circles*). *Plus signs* indicate positive RM toward pulsars, *small circles* negative RM. *Red symbols* indicate positive RM toward extragalactic sources, *green symbols* negative RM. *The blue arrows* suggest large-scale magnetic fields along a model of spiral arms in the Milky Way. Our sun is located at the upper crossing of coordinate lines [25]

The major compilation of pulsar rotation measures, also using already published data, are given in Han [25], Han et al. [27] and references therein. Additional results are found in Mitra et al. [42] and Noutsos et al. [44]. The distribution of rotation measures, as given by [25], shows a huge variation of signs and magnitudes. This may indicate a large-scale regular magnetic field with multiple reversals. Large RMs are often due to local H II regions. Recent studies of pulsars in the fourth Galactic quadrant by Nota and Katgert [43] also concluded that magnetic fields have considerable structure on small scales, but that average values in arm and interarm are well defined. The RM values increase for distant objects, but very few pulsars were found beyond the Galactic Center. The limit of $|\mathrm{RM}| \approx 1000\,\mathrm{rad\,m^{-2}}$ for EGRS holds also for pulsars. This seems to indicate that the RM toward EGRS is partly averaged out in passage through the Galaxy. The large-scale regular field of the outer Milky Way is either weak or frequently reversing its direction.

The Zeeman effect is the most direct method of measuring magnetic fields. It has been used in the optical range for detecting magnetic fields in the Sun and in stars. At radio wavelength the use of the Zeeman effect proved to be more difficult. For one, the frequency shifts caused by the weak magnetic fields are minute and require sophisticated instrumentation. The H I line gave the first definitive detections, usually in absorption toward strong Galactic sources [60]. The technique was refined so that at present magnetic fields as weak as $\approx 5\,\mu$G can be detected with the Arecibo telescope [29]. The observation of the Zeeman effect in the OH molecule (e.g., [14, 15]) advanced the field further. It became clear that many of the positive detections were in molecular clouds with maser sources. Strong magnetic fields (≈ 80 mG) were detected in interstellar H_2O maser clouds [21]. Millimeter-wavelength astronomy gave us additional results for high recombination lines (e.g., [59]).

4 Magnetic Fields in Galaxies

Magnetic fields in external galaxies can be observed with the same methods as in the Milky Way. Naturally the spatial resolution of the telescopes is much worse in galaxies, and the detailed structure of extragalactic fields on scales below about 100 pc are still invisible. On the other hand, the large-scale field properties, like the overall pattern and the total extent, can be best measured in external galaxies. Observations in the Milky Way and in external galaxies are complementary.

Weak optical polarization (generally below 1%) is the result of extinction by elongated dust grains in the line of sight which are aligned in the interstellar magnetic field (the Davis–Greenstein effect). Optical polarization surveys yielded the large-scale structure of the field in the local spiral arm of our Milky Way. The first extragalactic results by Hiltner [30] were based on starlight polarization of globular clusters in M 31 and showed that the magnetic field is aligned along the galaxy's major axis. Polarization of starlight in the LMC also gave evidence for ordered magnetic fields [53].

Polarization from diffuse optical light was used to search for large-scale magnetic fields in nearby galaxies (e.g., [1]). Indications for ordered fields along the spiral arms were found in several nearby galaxies, e.g., M 51, M 81, M 82, NGC 1068, and in NGC 6946. The pattern in M 51 (Fig. 6) agrees well with the radio polarization results (Fig. 7) in the inner spiral arms but differences are seen in the outer arms and in the companion galaxy, which is un-polarized in the radio image. Radio continuum polarization, on the other hand, traces the magnetic fields in the diffuse medium which are mostly oriented parallel to the plane and have significant vertical components only beyond some height above the plane.

At wavelengths $\lambda \leq 6$ cm, Faraday rotation of the polarized synchrotron emission is generally small (except in central regions), so that the B-vectors (E-vectors rotated by 90°) directly trace the orientations of the ordered field. Spiral patterns were found in almost every galaxy [4], even in those lacking optical spiral structure like the

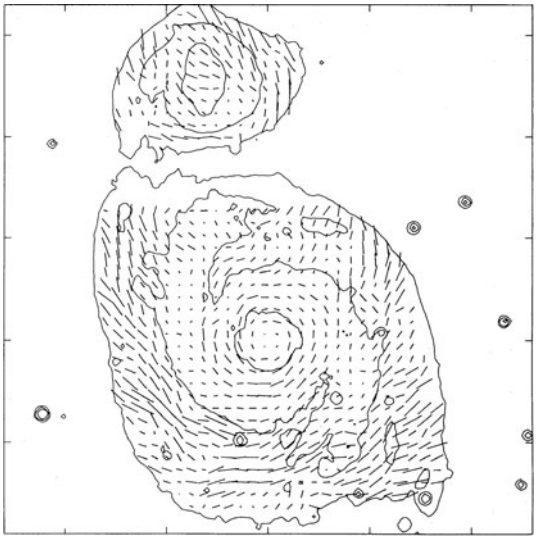


Fig. 6 Spiral galaxy M 51. E-vectors of the optical polarization of diffuse light which trace the spiral magnetic field orientation [51]. Compare with the radio polarization map in Fig. 7

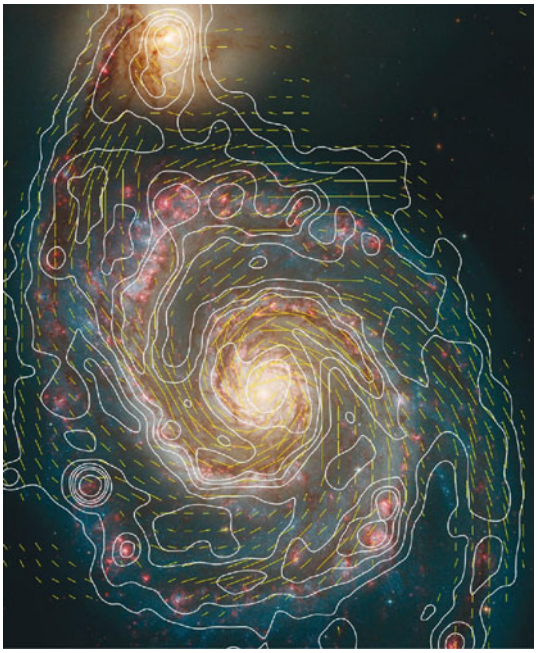


Fig. 7 Spiral galaxy M 51. Total radio intensity (contours) and B-vectors at 6 cm wavelength, combined from observations with the VLA and Effelsberg 100 m telescopes (Fletcher et al. [23]). The background optical image is from the HST (Hubble Heritage Team). Graphics: Sterne und Weltraum

ringed galaxies and flocculent galaxies, while irregular galaxies show at most some patches of spiral structure. Spiral magnetic fields are also observed in the nuclear starburst regions of barred galaxies (Fig. 10).

The Andromeda Nebula M 31 was an early target for magnetic field observations [7, 22]. The spiral galaxy M 51 also has been a great favorite in the early investigations because of its luminosity. M 51 shows a clear spiral magnetic field morphology (Fig. 7). The field lines generally follow the spiral arms with a displacement to the inside of the arms. The galaxies M 83 and NGC 2997 [26] are cases similar to M 51, with enhanced ordered (anisotropic) fields at the inner edges of the inner optical arms, ordered fields in interarm regions, and ordered fields coinciding with the outer optical arms.

The galaxy NGC 6946 was also an early favorite [33] with magnetic arms clearly between the optical ones (Fig. 8). For this galaxy considerable data are available in other spectral ranges. The radial distribution of synchrotron intensity in many spiral galaxies is well described by an exponential decrease of the scale length. In case of equipartition between the energy densities of magnetic fields and cosmic rays, the scale length of the magnetic field is $l \approx 16$ kpc (see Fig. 9). These are still lower limits because energy losses of cosmic-ray electrons increase with increasing distance from their origin in the galaxy's star-forming regions, and a lower density of cosmic-ray electrons needs a stronger field to explain the observed synchrotron intensity. Fields in the outer disk of galaxies can be amplified by the mean-field dynamo even without star formation activity because turbulence can be generated

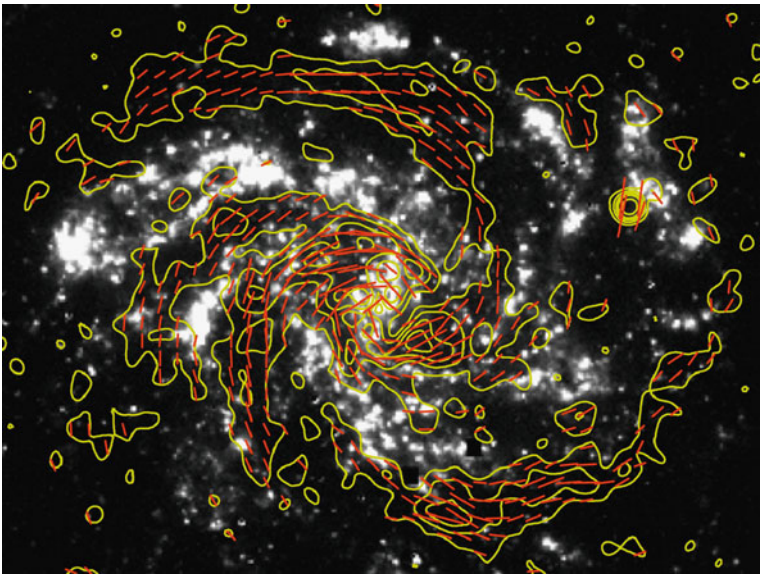


Fig. 8 Spiral galaxy NGC 6946. Polarized radio intensity (contours) and B-vectors at 6 cm, combined from observations with the VLA and Effelsberg 100 m telescopes (Beck [5]). The background Hα image is from Anne Ferguson. Graphics: Sterne und Weltraum

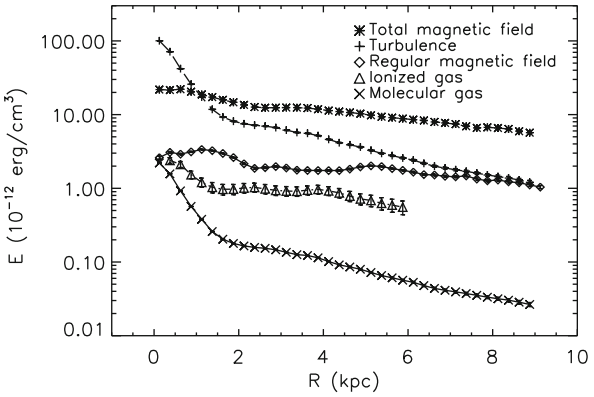


Fig. 9 Spiral galaxy NGC 6946. Radial variation of the energy densities of the total magnetic field, ordered magnetic field, turbulent motion of the neutral gas, thermal energy of the ionized gas, and thermal energy of the molecular gas (Beck [5])

by the magneto-rotational instability. The typical scale lengths of the density of neutral and ionized gas are only about 3 kpc, so that the magnetic field energy dominates over the turbulent energy in the outer region of galaxies if a constant turbulent velocity is assumed. The speculation that magnetic fields may affect the global gas rotation [3] needs testing by future radio observations with higher sensitivity and at low frequencies where energy loss of electrons is smaller.

Gas and stars in the gravitational potential of strongly barred galaxies move in highly noncircular orbits. Numerical models show that gas streamlines are deflected in the bar region along shock fronts, behind which the cold gas is compressed in a fast shearing flow [2]. The warm, diffuse gas has a higher sound speed and is not compressed. According to simulations, the shearing flows around a bar should amplify magnetic fields and generate complicated field patterns changing with time [45].

NGC 1097 (Fig. 10) is one of the nearest barred galaxies and hosts a huge bar of about 16 kpc length. The total radio intensity (not shown in the figure) and the polarized intensity are strongest in the downstream region of the dust lanes (south-east of the centre). This can be explained by a compression of turbulent fields in the bars shock, leading to strong and anisotropic fields in the downstream region. The pattern of field lines in NGC 1097 is similar to that of the gas streamlines as obtained in numerical simulations. This suggests that the ordered (partly regular) magnetic field is aligned with the flow and amplified by strong shear. Remarkably, the optical image of NGC 1097 shows dust filaments in the upstream region which are almost perpendicular to the bar and thus aligned with the ordered field. Between the region upstream of the southern bar and the downstream region the field lines smoothly change their orientation by almost 90°. The ordered field is probably coupled to the diffuse gas and thus avoids being shocked in the bar. The magnetic energy density in the upstream region is sufficiently high to affect the flow of the diffuse gas.

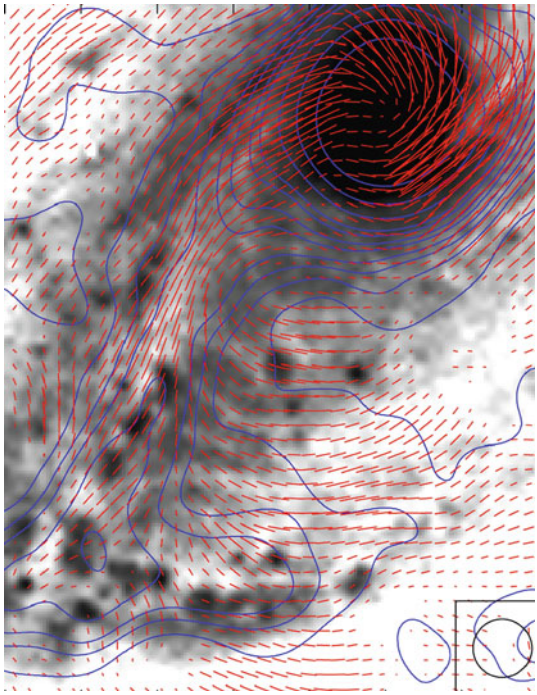


Fig. 10 Southern half of the barred galaxy NCC 1097. Total radio intensity (contours) and B-vectors at 3.5 cm, observed with the VLA (Beck [4]). The background optical image is from Halton Arp (Cerro Tololo Observatory)

Flocculent galaxies have large disks but no prominent spiral arms. Nevertheless, spiral magnetic patterns are observed in all flocculent galaxies, indicative that the mean-field dynamo works without assistance from density waves. Radio continuum maps of irregular, slowly rotating galaxies of the Local Group may reveal strong total equipartition magnetic fields of more than 10 μG. In NGC 4449 for instance a fraction of the field is ordered with about 7 μG strength and a spiral pattern [13]. Faraday rotation shows that this ordered field is mostly regular and the mean-field dynamo is operating in this galaxy. Dwarf irregular galaxies with almost chaotic rotation do not have any regular fields and only spots of faint polarized emission. Their turbulent field strengths are still comparable to those in large spiral galaxies (5–10 μG), because star formation activity is sufficiently high for the operation of the small-scale dynamo. This value may indicate a sensitivity limit of present-day observations or a threshold for small-scale dynamo action.

The Magellanic Clouds are the closest irregular galaxies and deserve special attention. Polarization surveys with the Parkes single-dish telescope at several wavelengths had low-angular resolution and revealed weak polarized emission. Two magnetic filaments were found in the LMC south of the 30 Doradus star formation complex [35].

Radio halos are observed around the disks of most edge-on galaxies, but their radio intensity and extent varies significantly. Possibly the best investigated edge-on

galaxy is M 82 [34]. In spite of the different intensities and extents of radio halos, their exponential scale heights at 5 GHz are about 1.8 kpc [18, 28]. A prominent exception is NGC 4631 with the largest radio halo observed so far. With large-scale heights, the magnetic energy density in halos is much higher than that of the thermal gas while still lower than that of the dominating kinetic energy of the gas outflow.

Polarization "vectors" do not distinguish between halo fields which are sheared into elongated loops or regular dynamo-type fields. A large-scale regular field can be measured only by Faraday rotation measures (RM). The detailed analysis of the multi-frequency observations of the highly inclined galaxy NGC 253 [28] allowed to identify an axisymmetric disk field with even symmetry and an X-shaped halo field with a quadrupolar pattern, also of *even* symmetry. Evidence for quadrupolar fields in galaxy halos around all galaxies observed so far with less than 60° inclination comes from the polarization asymmetry along the major axis.

Elliptical galaxies with active nuclei are among the brightest known radio sources. Their jets and radio lobes are generated by magneto-hydrodynamic processes. Radio emission from quiet elliptical and S0 galaxies is also associated with their nuclei [19]. Apart from the nuclear activity, elliptical galaxies are radio-faint because star-formation activity is very low and cosmic-ray electrons are rare. A few elliptical galaxies form stars in their inner regions, but synchrotron emission and hence magnetic fields were not yet detected.

5 More Distant Magnetic Fields

Magnetic fields have been observed in numerous distant objects. Radio galaxies are basically magnetic objects where a jet from the nucleus powers the outer lobes, distinct magnetic fireworks. The most distant quasars with $z \sim 7$ invariably are detected as a non-thermal radio source, implying the presence of magnetic fields. Study of polarized quasars with up to $z \sim 3$ by Kronberg et al. [37] and Bernet et al. [9] implied the presence of strong intervening magnetic fields.

Intergalactic magnetic fields have been detected in clusters of galaxies. The magnetic fields are seen as weak diffuse haloes at low radio frequencies. The most studied cluster is Coma A where in addition to the weak halo emission Rotation Measure of EGRS beyond the cluster could be used to determine the magnetic filed within the cluster. The radio halo of Coma A was studied by Deiss et al. [17] implying magnetic fields of $B \approx 2\,\mu$G. A similar result was derived by Kim et al. [32] using the rotation of EGRS beyond the Coma cluster.

6 Field Origin and Amplification

The origin of the first magnetic fields in the Universe is still a mystery [61]. The generation of the very first "seed" fields needs a continuous separation of electric charges, e.g., by the Biermann battery or the Weibel instability [39, 52]. The Biermann battery may generate a field of $\leq 10^{-20}$ G in the first galaxies or stars. A

large-scale intergalactic field of $\leq 10^{-12}$ G may be generated in the early Universe and could also serve as a seed field in protogalaxies. However, such a field is hard to maintain because the galaxy rotates differentially, so that field lines get strongly wound up, in contrast to the observations [55]. Moreover, a coherent large-scale field as observed, e.g., in M 31 cannot be explained by the primordial field model. The same is true for kinematical models of field generation by induction in shearing and compressing gas flows, which generate fields with a few kiloparsce coherence length and frequent reversals [40].

Origin of seed fields in protogalaxies has been discussed by Rees [47]. More promising is the magnetization of protogalaxies to $\geq 10^{-9}$ G by field ejection from the first stars or the first black holes followed by dynamo action. The dynamo transfers mechanical into magnetic energy. It amplifies and/or orders a seed field. The small-scale or fluctuation dynamo does not need general rotation, only turbulent gas motions (e.g., [10]). The source of turbulence can be thermal virialization in protogalactic halos or supernovae in the disk or the magneto-rotational instability (MRI) (e.g., [50]). Within less than 10^9 yr even weak seed fields are amplified to the energy density level of turbulence and reach strengths of a few μG.

The mean-field or α–Ω dynamo is driven by turbulent gas motions from supernova explosions or cosmic-ray driven Parker loops (α) and by differential rotation (Ω), plus magnetic diffusivity (η) (e.g., [8, 46, 63]). It generates a large-scale ("mean") regular field from the turbulent field in a typical spiral galaxy within a few 10^9 year. If the small-scale dynamo already amplified turbulent fields of a few μG in the protogalaxy, the mean-field dynamo is needed only for the organization of the field ("order out of chaos"). The field pattern is described by modes of different azimuthal symmetry in the disk and vertical symmetry or antisymmetry perpendicular to the disk plane. Several modes can be excited in the same object (e.g., [63]). In almost spherical, rotating bodies like stars, planets or galaxy halos, the strongest mode consists of a toroidal field component with a sign reversal across the equatorial plane (vertically antisymmetric or "odd" parity mode A0) and a dipolar poloidal field component with field lines crossing the equatorial plane. The halo mode can also be oscillatory and reverse its parity with time (e.g., causing the cycle of solar activity). The oscillation timescales (if any) are very long for galaxies and cannot be determined by observations.

In flat, rotating objects like galaxy disks, the strongest mode consists of a toroidal field component, which is symmetric with respect to the equatorial plane and has the azimuthal symmetry of an axisymmetric spiral in the plane without sign reversals (vertically symmetric or "even" parity mode S0), and a weaker poloidal field component of quadrupolar structure with a reversal of the vertical field component across the equatorial plane. The next higher azimuthal mode is of bisymmetric spiral shape (see Fig. 11) with two sign reversals in the plane, followed by more complicated modes. The pitch angle of the spiral field is determined by the rotation curve of the galaxy, the turbulent velocity, and the scale height of the warm diffuse gas [55].

Regular dynamo-type magnetic fields generate large-scale patterns of Faraday rotation measures (RM). Axisymmetric fields were detected in several galaxies (e.g., M 31, IC 342, LMC) and dominate, in agreement with dynamo models. However,

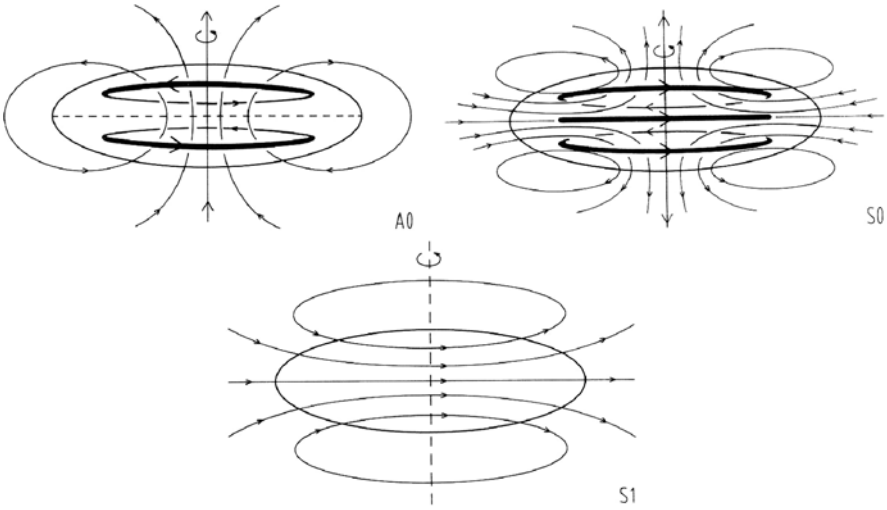


Fig. 11 Possible modes of a dynamo (Wielebinski and Krause [63])

in many galaxies no clear RM patterns were found, either due to limited sensitivity and to distortions of the field structure, e.g., by interactions. The mean-field dynamo generates large-scale helicity with a non-zero mean in each hemisphere. As total helicity is a conserved quantity, the dynamo is suppressed by the small-scale fields with opposite helicity, unless these are removed from the system. Hence, outflow with a moderate velocity or diffusion is essential for an effective mean-field dynamo. This effect may relate the efficiency of dynamo action to the star-formation rate in the galaxy disk. Mean-field dynamo models including outflows with moderate velocities can also generate X-shaped fields [11, 28]. For fast outflows the advection time for the field becomes smaller than the dynamo amplification time, so that the dynamo action is no longer efficient.

There are several more unsolved problems with dynamo theory. The "mean-field" model is simplified because it assumes a dynamical separation between the small and the large scales. The large-scale field is assumed to be smoothed by turbulent diffusion, which requires fast and efficient field reconnection. One of the main future tasks is to compute the "mean" quantities α and η from the small-scale properties of the interstellar medium, which is only possible with numerical MHD modeling. Such models must include the whole rotating galaxy, including the halo, consider all relevant effects, and need a spatial resolution of smaller than the turbulence scale ($\approx$ 10 pc). The multiphase interstellar medium has also to be taken into account [55]. Rapid progress in modeling galactic fields can be expected in near future. The predictions of the dynamo model have been generally confirmed by present-day observations the primordial model of field amplification is less developed than the dynamo model and is not supported by the data. A wound-up large-scale seed field can generate only the even bisymmetric mode (S1) or the odd dipolar mode (A0), both of which were not observed so far. On the other hand, the number of galaxies

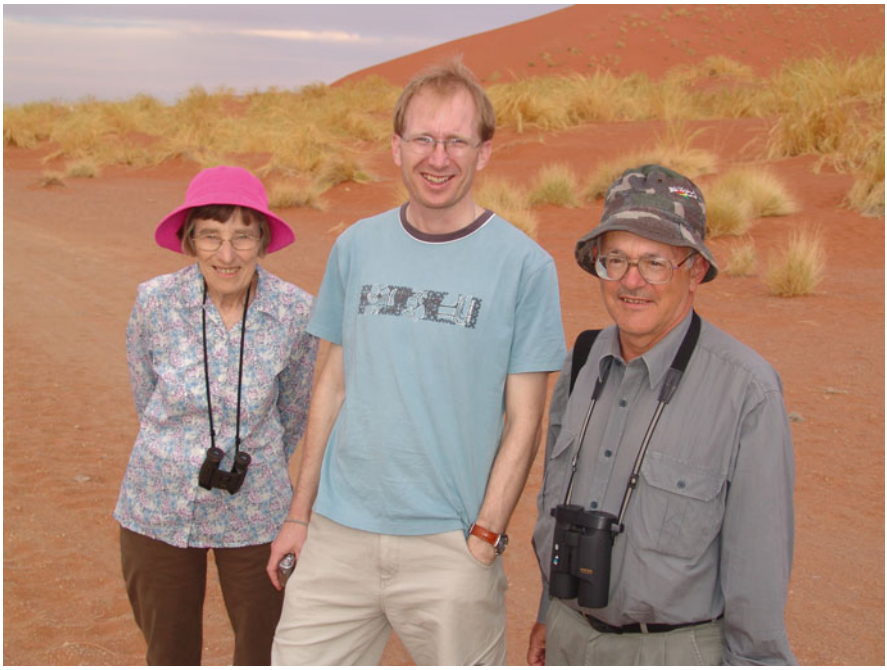
with a well-determined field structure is still limited. Future radio telescopes will be able to decide whether the dynamo or the primordial model is valid or whether a new model has to be developed.

Acknowledgments We would like to thank many of our colleagues who have pursued the studies of magnetic fields in the Milky Way and in galaxies. At the MPIfR we thank Wolfgang Reich, Marita Krause, and Patricia Reich. We had many excellent cooperation projects in this field with groups in Kraków (Poland), DRAO Penticton (Canada), NAOC Beijing (China), ATNF (Australia), Moscow and Perm (Russia), Newcastle (UK), Potsdam and Bochum (Germany).

References

1. Appenzeller, I. 1967, PASP, 79, 600
2. Athanassoula, E. 1992, MNRAS, 259, 345
3. Battaner, E., Florido, E. 2000, Fund Cosmic Phys, 21, 1
4. Beck, R. 2005, In: R. Wielebinski, R. Beck, eds. Cosmic Magnetic Fields, Lecture Notes in Physics, Springer, Heidelberg, vol. 664, p. 41
5. Beck, R. 2007, A&A, 470, 539
6. Beck, R., Wielebinski, R. 2010, In: G.G. Gilmore, ed. Star and Stellar Systems, Springer, Heidelberg, in press
7. Beck, R., Berkhuijsen, E.M., Wielebinski, R. 1980, Nature, 283, 272
8. Beck, R., Brandenburg, A., Moss, D., Shukurov, A., Sokoloff, D. 1996, ARA&A, 34, 155
9. Bernet, M.L., Miniati, F., Lilly, S.J., et al. 2008, Nature, 454, 302
10. Brandenburg, A., Subramanian, K. 2005, Phys Rep, 417, 1
11. Brandenburg, A., Donner, K.J., Moss, D., et al. 1993, A&A, 271, 36
12. Brentjes, M.A., de Bruyn, A.G. 2005, A&A, 441, 1217
13. Chyży, K.T., Beck, R., Kohle, S., Klein, U., Urbanik, M. 2000, A&A, 356, 757
14. Crutcher, R.M., Kazès, I. 1983, A&A, 125, L23
15. Crutcher, R.M., Kazès, I., Troland, T.H. 1987, A&A, 181, 119
16. Davis, L.J., Greenstein, J.L. 1951, ApJ, 114, 206
17. Deiss, B.M., Reich, W., Lesch, H., Wielebinski, R. 1997, A&A, 321, 55
18. Dumke, M., Krause, M., Wielebinski, R., Klein, U. 1995, A&A, 302, 691
19. Fabbiano, G., Klein, U., Trinchieri, G., Wielebinski, R. 1987, ApJ, 312, 111
20. Ferrière, K. 2009, A&A, 505, 1183
21. Fiebig, D., Gsten, R. 1989, A&A, 214, 333
22. Fletcher, A., Berkhuijsen, E.M., Beck, R., Shukurov, A. 2004, A&A, 414, 53
23. Fletcher, A., Beck, R., Shukurov, A., et al. 2010, MNRAS, in press
24. Gao, X.Y., Reich, W., Han, J.L., et al. 2010, A&A, 515, A64
25. Han, J.L. 2007, In: J.M. Chapman, W.A. Baan, eds. Astrophysical Masers and their Environments, Proc. IAU Symp. 242, CUP, Cambridge, pp. 55–63
26. Han, J.L., Beck, R., Ehle, M., Haynes, R.F., Wielebinski, R. 1999, A&A, 348, 405
27. Han, J.L., Demorest, P.B., van Straten, W., Lyne, A.G. 2009, ApJS, 181, 557
28. Heesen, V., Beck, R., Krause, M., Dettmar, R.-J. 2009, A&A, 494, 563
29. Heiles, C., Crutcher, R. 2005, In: R. Wielebinski, R. Beck, eds. Cosmic Magnetic Fields, Lecture Notes in Physics, Springer, Heidelberg, vol. 664, p. 137
30. Hiltner, W.A. 1958, ApJ, 128, 9
31. Hinshaw, G., Weiland, J.L., Hill, R.S., et al. 2009, ApJS, 180, 225
32. Kim, K-T., Kronberg, P.P., Dewdey, P.E., Landecker, T.L. 1990, ApJ, 355, 29
33. Klein, U., Beck, R., Buczilowski, U.R., Wielebinski, R. 1982, A&A, 108, 176
34. Klein, U., Wielebinski, R., Morsi, H. 1988, A&A, 190, 41

35. Klein, U., Haynes, R.F., Wielebinski, R., Meinert, D. 1993, A&A, 271, 402
36. Kogut, A., Dunkley, J., Bennett, C.L., et al. 2007, ApJ, 665, 355
37. Kronberg, P.P., Bernet, M.L., Miniati, F., et al. 2008, ApJ, 676, 70
38. Landecker, T.L., Reich, W., Reid, R.I., et al. 2010, A&A, in press
39. Lazar, M., Schlickeiser, R., Wielebinski, R., Poedts, S. 2009, ApJ, 693, 1133
40. von Linden, S., Otmianowska-Mazur, K., Lesch, H., Skupniewicz, G. 1998, A&A, 333, 79
41. Mathewson, D.S., Ford, V.L.1970, Mem RAS, 74, 139
42. Mitra, D., Wielebinski, R., Kramer, M., Jessner, A. 2003, A&A, 398, 993
43. Nota, T., Katgert, P. 2010, A&A, 513, A65
44. Noutsos, A., Johnston, S., Kramer, M., Karastergiou, A. 2008, MNRAS, 386, 1881
45. Otmianowska-Mazur, K., Elstner, D., Soida, M., Urbanik, M. 2002, A&A, 384, 48
46. Parker, E.N. 1979, Cosmical Magnetic Fields, Clarendon, Oxford
47. Rees, M.J. 2005, In: R. Wielebinski, R. Beck, eds. Cosmic Magnetic Fields, Lecture Notes in Physics, Springer, Heidelberg , vol. 664, p. 1
48. Reich, W., Reich, P. 2009, In: K.G. Strassmeier et al., eds. Cosmic Magnetic Fields: From Planets, to Stars and Galaxies. Proceedings of the International Astronomical Union, IAU Symposium, vol. 259, pp. 603–612
49. Reich, W., Sofue, Y., Wielebinski, R., Seiradakis, J.H. 1988, A&A, 191, 303
50. Rüdiger, G., Hollerbach, R. 2004, The Magnetic Universe, Wiley, Weinheim
51. Scarrott, S.M., Ward-Thompson, D., Warren-Smith, R.F. 1987, MNRAS, 224, 299
52. Schlickeiser, R. 2005, Plasma Phys Contr F, 47, 205
53. Schmidt, Th. 1976, A&AS, 24, 357
54. Seiradakis, J.H., Lasenby, A.N., Yusef-Zadeh, F., et al. 1985, Nature, 317, 697
55. Shukurov, A. 2005, In: R. Wielebinski, R. Beck, eds. Cosmic Magnetic Fields, Lecture Notes in Physics, Springer, Heidelberg, vol. 664, p. 113
56. Sun, X.H., Han, J.L., Reich, W., et al. 2007, A&A, 463, 993
57. Sun, X.H., Han, J.L., Reich, W., et al. 2010, in preparation
58. Testori, J.C., Reich, P., Reich, W. 2008, A&A, 484, 733
59. Thum, C., Morris, D. 1999, A&A, 344, 923
60. Verschuur, G.L. 1968, Phys Rev Lett, 21, 775
61. Widrow, L.M. 2002, Rev Mod Phys, 74, 775
62. R. Wielebinski, R. Beck, eds. 2005, Cosmic Magnetic Fields, Lecture Notes in Physics, Springer, Heidelberg, vol. 664
63. Wielebinski, R., Krause, F. 1993, Rev A&AR, 4, 449
64. Wolleben, M. (private communication)
65. Wolleben, M., Landecker, T.L., Reich, W., Wielebinski, R. 2006, A&A, 448, 411
66. Yusef-Zadeh, F., Morris, M., Chance, D. 1984, Nature, 310, 557



The Gaseous Halo Mask

Mary E. Putman, M. Ryan Joung, Jana Grcevich, and Fabian Heitsch

Abstract Support for spiral galaxies harboring an extended cosmologically significant gaseous halo is presented through evidence of the stripping of gas from dwarf galaxies in both observations and local and cosmological simulations. Dwarf galaxies within 270 kpc of the Milky Way and Andromeda are largely devoid of gas, and those with significant amounts of gas are beyond this radius. This is most easily explained by ram pressure stripping as the dwarfs move through perigalacticon. The star formation histories of the dwarf galaxies can also be partially understood in the context of their movement through this halo medium. Since the dwarfs are stripped at distances generally greater than 20 kpc, the gas is unlikely to make it directly into the spiral galaxy disk as cold star formation fuel. Some of the dwarfs' gas may be integrated into the spiral galaxy's hot halo and warm clumps may undergo a recooling process close to the disk.

1 Introduction

Gaseous halos are a baryonic component of galaxies that remain elusive in terms of their role in the evolution of galaxies and their total mass contribution. There is indirect evidence that the hot gaseous halos of spiral galaxies extend out to large radii [6, 10, 18], but no direct detection beyond a few kpc from a galaxy's disk [16, 24], and we do not know the exact location of our own Galaxy's hot gas [3, 23]. The detection of a large covering fraction of warm gas at large radii from galaxies with absorption line studies provides another clue that there exist some form of multi-phase medium around galaxies [4]. If the hot gas does indeed extend out to the virial radius it would provide support for cosmological models of galaxy formation that predict the collapse of baryons into the dark matter potential wells of large galaxies heated the gas to high temperatures [12, 25]. If it has enough density at these large radii, it could also host a significant fraction of a galaxy's baryons [5, 21].

M.E. Putman (✉)
Columbia University, New York, NY 10027, USA
e-mail: mputman@astro.columbia.edu

D.L. Block et al. (eds.), *Galaxies and their Masks*,

DOI 10.1007/978-1-4419-7317-7_6,

Dwarf galaxies surrounding large spirals show a clear deficit of baryons compared to their dark matter content, in particular the newly discovered ultra-low luminosity dwarfs in the Local Group have mass to light ratios approaching 1,000 [19]. The explanation for the deficit of both stellar and gaseous baryons remains unclear. The dwarf galaxies most likely lost their gaseous baryons at early times and were therefore limited in the stellar population they could form. One of the possible gas-loss mechanisms is the passage of the dwarf through the extended gaseous halo of a larger primary galaxy. Therefore, the study of gas loss in dwarf galaxies can provide information on both the baryon content of dwarf galaxies and the presence of an extended, gaseous component to the large spiral galaxies they are found around. In this chapter we present observational results, local simulations of dwarf galaxy stripping, and a cosmological simulation of a Milky Way mass galaxy and its companions to study the presence and properties of the extended gaseous halos of spiral galaxies and the baryonic content and star formation history of dwarf galaxies.

2 Gas Stripping from Dwarf Galaxies

The distribution of dwarfs with and without gas in the Local Group suggests there is a distant-dependent mechanism stripping them of their fuel. Figure 1 shows the HI mass of the Local Group dwarf galaxies (or limit on their HI mass) versus their distance from the Milky Way or Andromeda (whichever is closer) [6]. Dwarfs within 270 kpc of either galaxy are devoid of gas to low detection limits with the exception of the small amount of gas in the two dwarf elliptical galaxies. The presence of HI in these dwarf ellipticals galaxies and the ongoing stripping of the LMC/SMC system may be partially due to the larger total masses of these galaxies in comparison to the other dwarfs. Indeed the LMC/SMC system is not included in Fig. 1 due to their combined total mass putting them above our 10^{10} $M_{\odot}$ dwarf galaxy limit. Figure 1 is an update from Grcevich and Putman [6] as it includes five new Andromeda dwarfs or all of the dwarfs discovered through the beginning of 2010. The HI mass limits on many of the Andromeda dwarfs are close to where the lowest HI mass dwarf galaxies have been detected, so these limits should be taken deeper in the future.

Both tidal and ram pressure stripping are distance-dependent mass-loss mechanisms that could explain Fig. 1. However, most ultra-faint dwarfs show no velocity gradients and thus no evidence for tidal disruption in their cores [19, 22]. Ram pressure stripping can recreate the distance-dependent gas content trend while leaving the stars unaffected. According to the Gunn and Gott [7] ram pressure stripping criterion, the dwarfs will lose their gas if the pressure exerted on the dwarf's gas by the hot halo gas exceeds the gravitational binding force. This criterion depends on the square of the dwarf's orbital velocity and the halo gas density and is therefore most likely to be satisfied when the dwarf galaxies are near perigalacticon in their orbit.

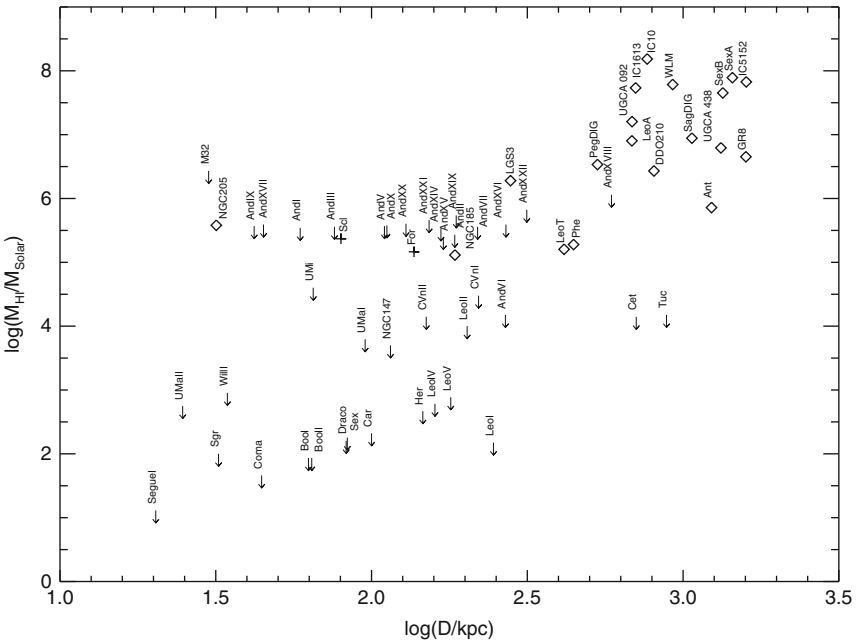


Fig. 1 The relationship between dwarf galaxies HI mass (*open diamond* is a detection, *downward arrow* is the upper limit) and distance to either the Milky Way or Andromeda (whichever is closer). Fornax and Sculptor are noted as crosses because they have nearby HI clouds that may or may not be associated (see Grcevich and Putman [6]). This plot includes all dwarfs discovered through the beginning of 2010

The halo density required to strip the dwarfs was calculated in Grcevich and Putman and reasonable values were obtained compared to the predictions of cosmological simulations.

To test the Gunn and Gott stripping criterion in the case of low mass dwarf galaxies, we conducted 2D hydrodynamic “wind tunnel" simulations of the stripping process using the Proteus code [8, 9, 20]. Simulations were run that include the range of realistic dwarf central gas densities, hot halo densities, orbital velocities, and dark matter halo characteristics. Figure 2 shows an example of the disruption of a Leo T-like dwarf galaxy [17] at two different timesteps. The spatial resolution of the simulations is $\sim$ 1 pc, high enough to resolve instabilities along the interface of the dwarf and halo gas, and therefore include the contribution of these instabilities to the mass loss rate.

If we examine the strength of the ram pressure acting on the simulated dwarfs versus their initial central gas density, we find that they follow the Gunn and Gott stripping criterion in the sense that they completely lose their gas within 500 Myr. Those dwarfs that never meet the stripping criterion simply experience continual low-level stripping from their edges. Extrapolating this slower dynamical stripping

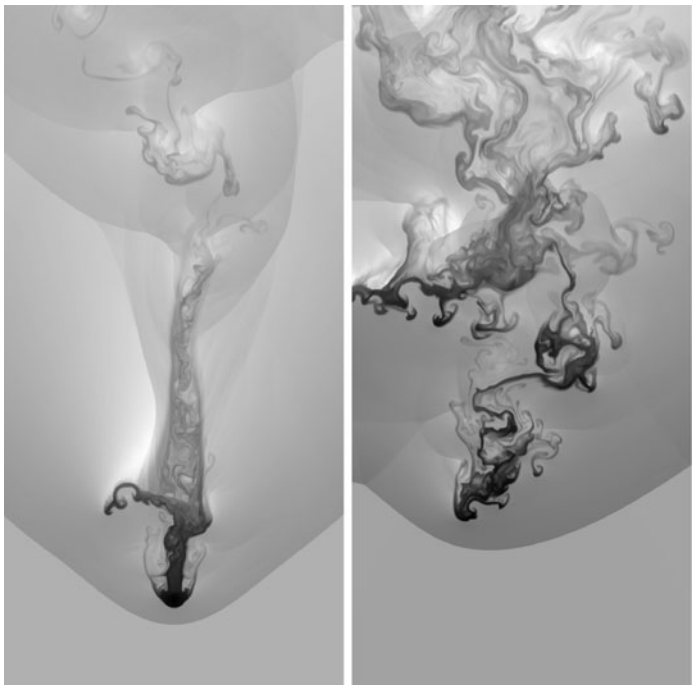


Fig. 2 Gas density at 200 Myr (*left*) and 500 Myr (*right*) for a simulated dwarf galaxy having $M(< 300\ pc)_{tot} = 3.5 \times 10^7\ M_\odot$. The dwarf has an initial central gas density of 0.35 cm^{-3} and is moving through a 1.5×10^{-4} cm^{-3} hot halo medium at a constant velocity of 300 km s^{-1}. This dwarf meets the criterion for ram pressure stripping and the galaxy is completely stripped of gas in this simulation. The dynamical instabilities acting on the gas are evident

into the future, it would require several Gigayears to leave the dwarf devoid of gas with only these mechanisms. In future work we plan to expand the simulations to the third dimension, include changing orbital velocities and directions, and extend the runs further in time to see the ultimate fate of the low-level dynamical stripping.

3 Tracking the Evolution of Dwarf Galaxies in Cosmological Simulations

To further explore the gas content and star formation history of Local Group dwarf galaxies, we examined a high-resolution cosmological adaptive mesh refinement (AMR) simulation of a Milky Way mass ($\sim 2 \times 10^{12}\ M_\odot$) galaxy using the Enzo code and tracked the evolution of satellites that are well resolved in the simulation. The simulation includes the radiative cooling of gas, meta-galactic UV background, star formation, and stellar feedback self-consistently and has a dark matter mass resolution of $2.8 \times 10^6\ M_\odot$ and spatial resolution of better than ~ 350 pc at all times.

We track the evolution of nine chemical species including HI, HII, and H_2 by solving a chemical reaction network simultaneously. The entire cosmological simulation having a volume of 64 (comoving Mpc/h)^3 was run at lower resolution first. The dark matter halo containing the disk galaxy, ($\sim$ 6.5 Mpc/h)^3 in size was selected from this volume and resimulated at the higher (mass and spatial) resolution.

Figure 3 shows a snapshot of the HI content of the simulation at z=0 with example dwarf galaxies circled (see caption). The majority of the dwarf galaxies come in along the cosmic filaments shown in low-density gas visible in this figure. The two

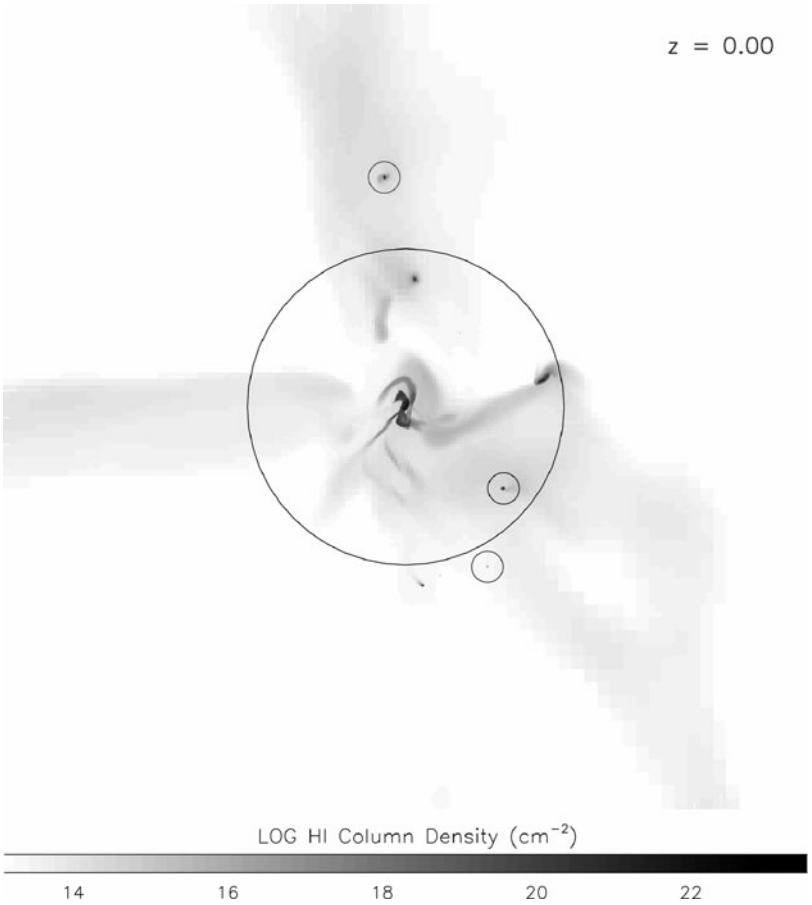


Fig. 3 HI column density map of a cosmological simulation of a Milky Way mass disk galaxy at z=0. The field of view is 1.4 Mpc^2 and all of the gas along 1 Mpc centered on the parent galaxy is included. The *large circle* represents the virial radius of the parent galaxy ($\sim$ 250 kpc), while the *three small circles* are example dwarf satellites. The one at bottom has already had a close passage with the parent galaxy (see Fig. 5) and has had most of its gas stripped. The next one from the bottom has just entered the virial radius of the parent galaxy and the one at the top has not yet reached the parent galaxy. Low column density gas representing the cold flow of gas into the galaxy is also evident in this figure

upper dwarf galaxies have not yet entered the virial radius of the simulated spiral and have largely retained their gas. The bottom dwarf was at its perigalacticon of 50 kpc at z=0.5 and subsequently lost a significant fraction of its gas mass. This loss in gas mass did not have a corresponding loss in stellar mass, indicating a stripping method that only affects the gas was active. It is also evident in the images of the dwarf's gas at various timesteps that it appears to be undergoing ram pressure stripping from the medium surrounding the spiral.

The HI mass of the satellite galaxies versus distance from the simulated spiral is shown in Fig. 4. The observational relationship of Fig. 1 is roughly reproduced, as all the dwarf galaxies with significant detectable neutral hydrogen gas are beyond at least 200 kpc in the simulation. This simulation represents dwarf galaxies with masses between $\sim 10^{8-10}$ $M_{\odot}$. A higher resolution simulation is currently being completed to examine those dwarfs with masses in the 10^7 $M_{\odot}$ range, as this is thought to be the mass of many of the newly discovered Local Group dwarf galaxies [19]. There is a higher percentage of dwarf galaxies at large distances without HI in the simulation than in the observational data. Additional gas-less dwarfs at large radii may exist in the Local Group and just not yet be discovered due to the sensitivity limitations of the SDSS survey and the area covered by the deep surveys of Andromeda [11, 13].

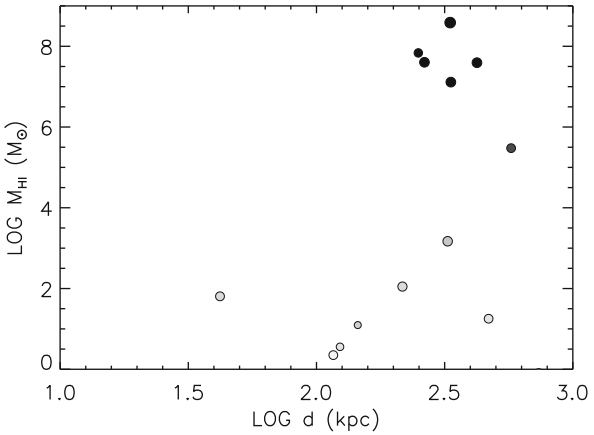


Fig. 4 The HI masses of the dark matter halos with total masses $< 10^{10}$ $M_{\odot}$ and star formation plotted against distance to the center of the simulated Milky Way (see Fig. 3). Most of the simulated dwarf galaxies with significant gas are found beyond $\sim$ 200 kpc. The shade of the symbol represents the ratio of M_{HI}/M_{tot} (darker is higher values) and the size of the symbol is scaled by the M_{tot} value

With the simulation results we can track the evolution of the mass and star formation of the dwarf galaxies with time and distance from the parent galaxy. Figure 5 shows this for the dwarf galaxy at the bottom of Fig. 3 that has orbited within the virial radius of the simulated Milky Way. The stellar mass of the galaxy initially increases as it forms stars and then remains flat after $z \sim 0.5$. The HI (and total gas)

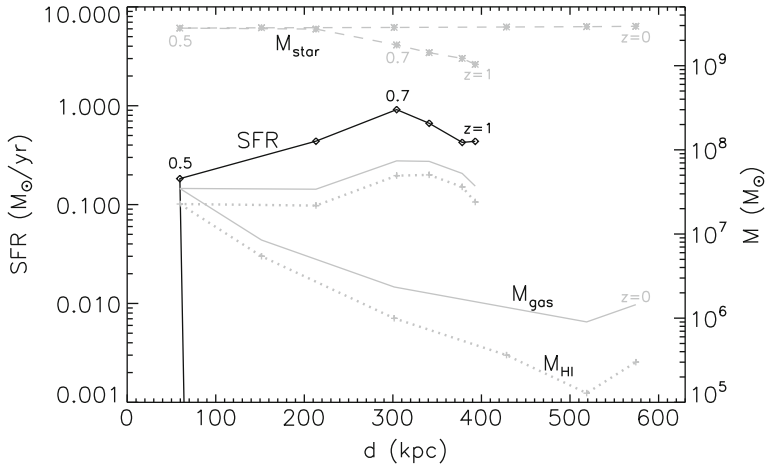


Fig. 5 This plot describes the gas content and star formation history of the dwarf at the bottom of Fig. 3 that has had a close passage with the simulated Milky Way. The *gray dashed line* shows the stellar content of the dwarf that increases initially as it forms stars and then remains approximately flat from z=0.5 to 0. The *dotted gray line* shows the HI mass of the galaxy and the *solid gray line* the total gas mass

mass of the galaxy initially remains relatively flat and then begins to decrease due to ram pressure stripping as the dwarf has a close passage with the primary galaxy at a distance of ∼50 kpc. The efficiency of the stripping of gas from the dwarf may be aided by the star formation rate for the dwarf slightly increasing at $z \sim 0.7$. The increase in star formation for this particular dwarf is due to a merger between two satellites, but in cases without a merger there are also increases in star formation that may correspond to the dwarf passing through a diffuse medium and having its gas compressed. As shown in Fig. 5, with the loss of the majority of the dwarf's gas, the star formation in the dwarf drops dramatically. The lack of evolution of the stellar mass of the dwarf indicates that the dwarf is not undergoing significant tidal effects that could potentially also play a role in stripping the dwarf. The mass and star formation evolution of other dwarfs in the simulation will be presented in a future paper with the higher resolution simulation. Comparisons to the star formation histories of the Local Group dwarf galaxies will provide significant insight into understanding their stellar populations.

4 What happens to the Stripped Gas?

The fate of the gas stripped from the dwarf galaxies in the context of fueling spiral galaxies is another interesting aspect to explore. Galaxies like the Milky Way require infalling low metallicity star formation fuel to explain their metallicity distribution

and star formation rate [15, chapter "Chemical Enrichment in Galaxies: Constraints on Nucleogenesis and Galaxy Evoluation" by Matteucci, this volume], and gas from dwarf galaxies could help to satisfy this requirement. The gas from the dwarf galaxies will most likely be stripped at radii > 20 kpc given the results from observations and simulations. This stripped gas is therefore unlikely to make it to the disk directly as cold star formation fuel, as simulations are finding that gas clouds moving through the halo are rapidly destroyed and do not travel further than ~ 10 kpc [2, 8]. The gas stripped from dwarf galaxies will therefore have to recool to successfully play a role in fueling star formation in the disk. The effect of the warm remnants of halo clouds recooling was seen in the simulations of Heitsch and Putman [8]. The simulations were set up to have a cloud moving through a evolving halo density, and as the remnant warm cloud moved through a denser and denser medium (similar to a cloud approaching the disk) it slowed down and was compressed, causing the cloud to recool. There is a large population of clouds that are being detected with new high resolution and sensitivity HI surveys that successfully observe the velocity range between the Galactic disk and the halo [1, 14]. The distances to all of these clouds are uncertain, but a large percentage of the population may be at the disk–halo interface, and potentially represent the cooling of star formation fuel near the disk. Figure 6 shows an example of these clouds at the disk–halo interface as detected by the Galactic Arecibo L-band Feed Array (GALFA)-HI Survey. This survey is covering 13,000 deg^2 at 180 m/s and 3.4′ resolution and is revealing a large population of small, cold features that were missed by previous surveys.

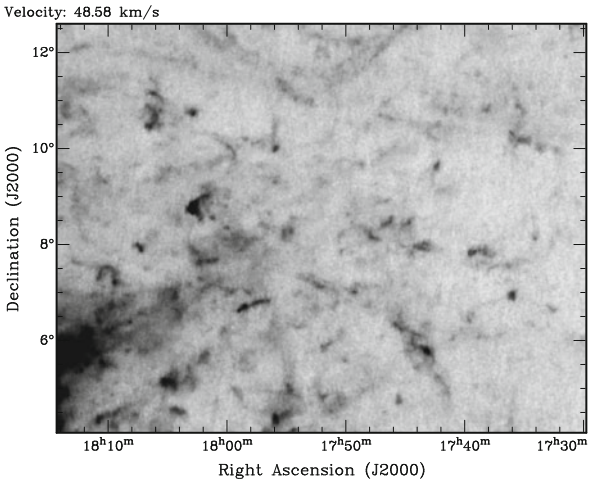


Fig. 6 Small, cold clouds at the disk–halo interface over a $\sim$3 km s^{-1} channel of GALFA-HI data at $V_{LSR} = 48.6$ km s^{-1} . This region is centered at $l, b \sim 18°, 32°$ and the resolution is 3.4′

5 Summary

The observations and simulations shown here are supportive of spiral galaxies harboring an extended cosmologically significant gaseous halo through evidence of the stripping of the gas from dwarf galaxies. Further work will continue to constrain the detailed properties of the gaseous halo in terms of its temperature, full extent, and clumpiness. The temperature is largely dependent on the galaxy's total mass in the simulations and ranges between 1×10^6 and 3×10^6 K. The star formation histories of the dwarf galaxies can also be potentially understood in the context of their movement through this halo medium initially triggering the dwarf to form more stars and then stripping the gas from the dwarf and preventing further star formation. The stripped gas will rain onto the disk of the galaxy, but since the dwarfs are stripped at distances generally greater than 20 kpc the fuel is unlikely to make directly into the galaxy in cold form. Some of the gas may be integrated into the hot halo, and warm clumps stripped from the dwarfs may undergo a recooling process as they fall through the denser gaseous medium close to the disk. Future observations that probe the multiple phases of the halo gas (e.g., the GALFA-HI Survey and the Cosmic Origins Spectrograph) and local and cosmological simulations (to achieve the resolution required and examine the evolution of an entire galactic system over cosmic time) will cast further light on the gaseous halo mask and its role in fueling galaxies.

Acknowledgments Ken Freeman of course! Ken is a thesis advisor extraordinaire and an important role model. M.E.P. acknowledges funding from NSF grants AST-0904059, AST-0917810, the Luce Foundation, and the Research Corporation, and thanks David Block for his hard work organizing the conference. Computations were performed at the NASA Advanced Supercomputing (NAS) Division and at the National Center for Supercomputing Applications supported by the NSF.

References

1. Begum, A. et al. 2010, ApJ, in press
2. Bland-Hawthorn, J., Sutherland, R., Agertz, O., Moore, B. 2007, ApJ, 670, L109
3. Bregman, J. et al. 2009, ApJ, 699, 1765
4. Chen, H.-W., Helsby, J.E., Gauthier, J.-R., Schectman, S.A., Thompson, I., Tinker, J. 2010, ApJ, 714, 1521
5. Fukugita, M., Peebles, P.J.E. 2006, ApJ, 639, 590
6. Grcevich, J., Putman, M.E. 2009, ApJ, 696, 385
7. Gunn, J., Gott, J.R. 1972, ApJ, 176, 1
8. Heitsch, F., Putman, M.E. 2009, ApJ, 698, 1485
9. Heitsch, F. et al. 2004, ApJ, 674, 316
10. Hsu, W., Putman, M.E., Heitsch, F., Stanimirovic, S., Peek, J.E.G., Clark, S.E., 2010, AJ, in press
11. Ibata, R. et al. 2007, ApJ, 671, 1591
12. Keres, D., Katz, N., Fardal, M., Dave, R., Weinberg, D.H. 2009, MNRAS, 395, 160
13. Koposov, S. et al. 2007, ApJ, 669, 337
14. Lockman, F. J. 2002, ApJ, 580, L47
15. Peek, J.E.G. 2009, ApJ, 698, 1429

16. Rassmussen, J. et al. 2009, ApJ, 697, 79
17. Ryan-Weber, E.V., Begum, A., Oosterloo, T., Pal, S., Irwin, M.J., Belokurov, V., Evans, N., Zucker, D. 2008, MNRAS, 384, 535
18. Sembach, K.R. et al. 2003, ApJS, 146, 165
19. Simon, J., Geha, M. 2007, ApJ, 670, 313
20. Slyz, A., Prendergast 1999, A&AS, 139, 199
21. Sommer-Larsen, J. 2006, ApJ, 644, L1
22. Strigari, L.E., Bullock, J.S., Kaplinghat, M., Simon, J.D., Geha, M., Willman, B., Walker, M.G. 2008, Nature, 454, 1096
23. Wang, D., McCray, R. 1993, ApJ, 409, L37
24. Wang, Q.D., Chaves, T., Irwin, J.A. 2003, ApJ, 598, 969
25. White, S.D.M., Rees, M.J. 1978, MNRAS, 183, 341

Molecular Gas Properties of Galaxies: The SMA CO(2-1) B0DEGA Legacy Project

Daniel Espada, S. Martin, P.-Y. Hsieh, P.T.P. Ho, S. Matsushita, Lourdes Verdes-Montenegro, J. Sabater, Simon Verley, M. Krips, and V. Espigares

Abstract In the last two decades high-resolution ($< 5''$) CO observations for ~ 150 galaxies have provided a wealth of information about the molecular gas morphologies in the circumnuclear regions. While in samples of "normal" galaxies the molecular gas does not seem to peak toward the nuclear regions for about 50% of the galaxies, barred galaxies and mergers show larger concentrations. However, we do not exactly know from an observational point of view how the molecular gas properties of a galaxy evolve as a result of an interaction. Here we present the SMA CO(2–1) B0DEGA (Below 0 DEgree GAlaxies) legacy project (http://b0dega.iaa.es) in which we are imaging the CO(2–1) line of the circumnuclear regions ($1'$) of a large (~ 70) sample of nearby IR-bright spiral galaxies, likely interacting, and that still remained unexplored due to its location in the southern hemisphere. We find different molecular gas morphologies, such as rings, nuclear arms, nuclear bars, and asymmetries. We find a centrally peaked concentration in about 85% of the galaxies with typical size scales of about 0.5–1 kpc. This might be related to perturbations produced by recent interactions.

1 Introduction

The molecular gas is one of the dominant components of the interstellar medium (ISM) in the inner few kiloparsecs of spiral galaxies. It is not only an essential ingredient to understand the structure and kinematics of the ISM there but also constitutes the fuel for future star formation (SF). Molecular gas, due to its dissipative nature, efficiently loses angular momentum via gravitational torques and may fall into the central hundreds of parsec. This gaseous component is eventually one of the main drivers for the existence of a central starburst and/or the fueling an active galactic nuclei (AGN) [8]. As a result, the concentration index in barred galaxies

D. Espada (✉)
Instituto de Astrofísica de Andalucía – CSIC, Glorieta de la Astronomia s/n, 18080 Granada, Spain
e-mail: daniel@iaa.es

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_7,

is seen to be larger than in non-barred galaxies [6, 7]. Galaxy–galaxy interactions are expected to play a major role in triggering the formation of bars, resulting in larger central gas concentrations in each individual galaxy. Then, the merging of two galaxies can lead to a violent starburst episode [1, 9].

We aim to study the circumnuclear morphologies of the bulk of the molecular gas and its relation with both intrinsic (i.e., stellar morphology) and external variables (i.e., environment). High-resolution molecular gas studies are needed. However, the limited number of observed galaxies to date (∼150) and large number of variables at play seem insufficient to understand how they affect the general properties of the bulk of the molecular gas. The most numerous high-resolution CO studies (best tracer at our disposal to trace the bulk of the molecular gas) of nearby spiral galaxies performed to date are NRT-OVRO (N = 20 galaxies [6]), BIMA-SONG ($N = 44$ [4]), and the NUGA projects (N = 28 [3] and references therein). These images reveal a wide variety of morphologies in the circumnuclear molecular disks. BIMA-SONG is one of the largest compilations of CO interferometric data to date for "normal" galaxies and shows that molecular gas is not centrally peaked in 55% of the galaxies [4]. However, this result is not likely valid for other samples of galaxies with other environmental properties. We aim to enlarge the number of observed galaxies that are likely suffering recent interactions (although not mergers), which in addition to previously studied samples will allow us to study how interactions modify the molecular gas concentration.

2 The B0DEGA Sample

We focus on a sample of: (i) nearby galaxies, with recession velocities $V < 7,000\,\mathrm{km\,s^{-1}}$, being the bulk of galaxies at about 2,000 km s^{-1}; (ii) IR-bright, following the criterion: $2.58 \times S_{60\,\mu m} + S_{100\,\mu m} > 31.5$ Jy ($S_{60\,\mu m}$ and $S_{100\,\mu m}$ are the IRAS flux densities in the IRAS Point Source Catalog [5]) so that CO(2–1) fluxes are high enough to be easily detected with the Submillimeter Array (SMA) in a reasonable amount of time of about 2–3 h and with resolutions better than $\sim 5''$; (iii) located in the southern hemisphere and observable from SMA ($-45° < \delta < 0°$). We use the latter criterion because the lack of millimeter/submillimeter interferometers in the southern hemisphere has prevented to study detailed molecular gas properties of many southern galaxies. The SMA can observe sources down to $\delta = -45°$, and it is revealing the molecular gas properties of many unique southern sources (for example, Centaurus A [2]).

With these criteria it yields a total of $N = 134$ galaxies, out of which only 14 had high-resolution CO(2–1) maps in the literature. Here we present data for a subsample of the IR brightest $N = 30$ B0DEGA galaxies. Further observations for other galaxies in the sample are being carried out at the SMA as a legacy project.

The galaxies reside in mid-density regions in the local Universe, usually members of pairs, triplets, or groups, thus representing the galaxies that have suffered a recent interaction. In most cases they have peculiarities in the optical, such as

asymmetries and dust lanes, and physical companions are found quite nearby in some cases. Most of the galaxies are in the morphological type range Sa–Sc, although the distribution is mainly dominated (70%) by Sb–Sc galaxies (see optical DSS images in Fig. 1). A total of 72% of the galaxies of the here presented subsample possess a bar feature.

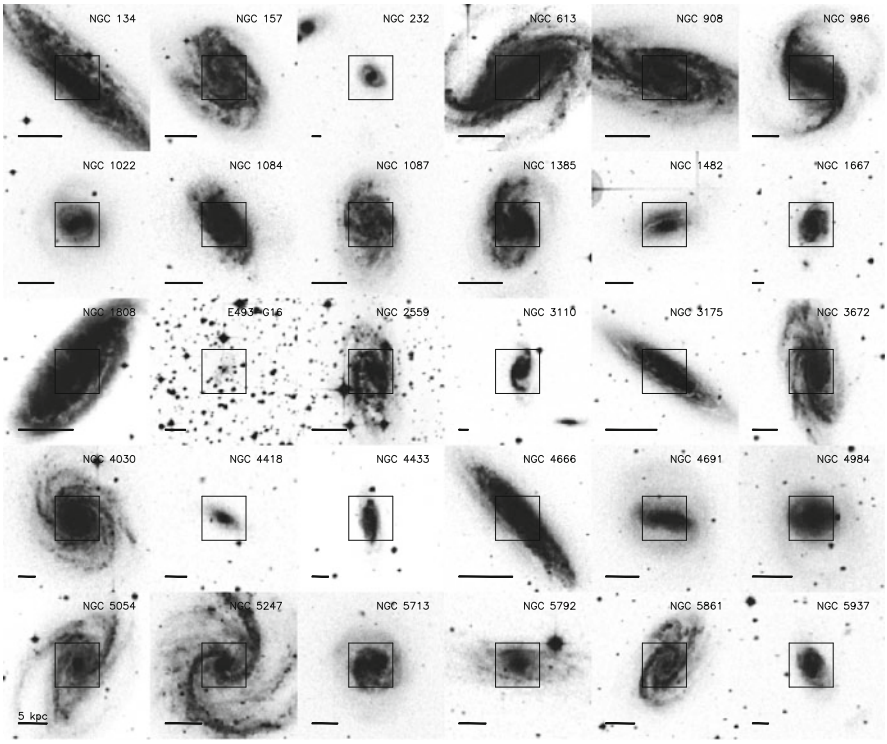


Fig. 1 DSS optical maps of the 30 galaxies in the B0DEGA subsample. The F.O.V. of the images is 220″. The *boxes* indicate the 66″ F.O.V. of the CO(2–1) maps. The horizontal segment on the *bottom left side* of each plot represents a 5 kpc linear scale

3 Preliminary Results: CO(2–1) Morphologies and Concentration

We have produced channel maps, integrated maps (moment 0), velocity fields (moment 1), CO(2–1) spectra, and azimuthally averaged radial distribution profiles for each of the galaxies. All these data products will be publicly available at `http://b0dega.iaa.es`.

We show the CO(2–1) distribution maps for our initial subsample of 30 galaxies in Fig. 2. It is remarkable that we detect molecular gas in the central regions in all

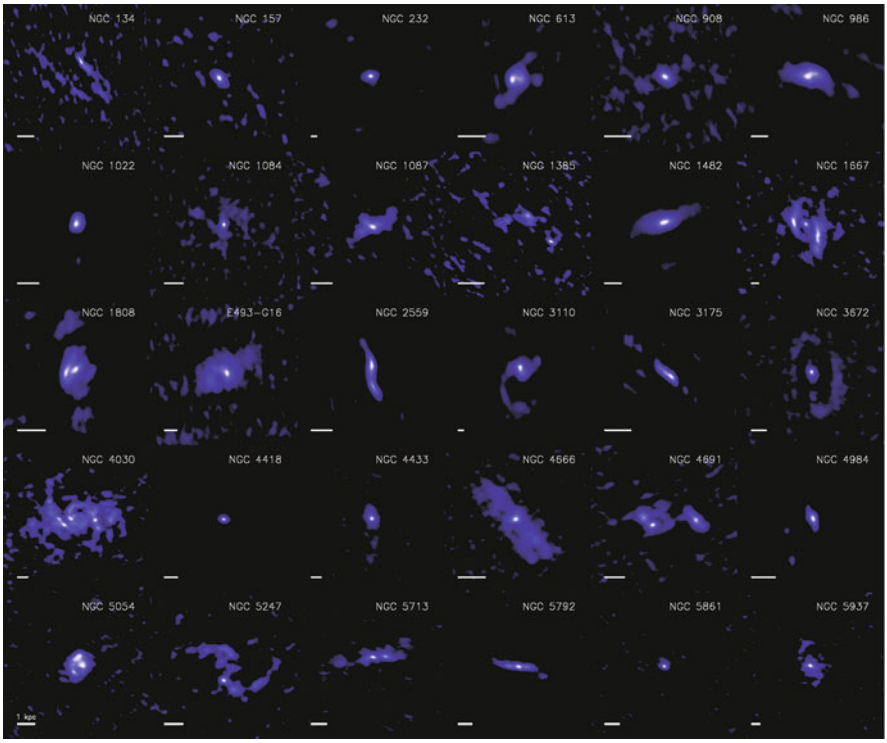


Fig. 2 CO(2–1) distribution of the 30 galaxies in the B0DEGA subsample. The F.O.V. is 66″ , slightly larger than the 52″ HPBW of SMA at 230 GHz. The horizontal segment on the *bottom left side* of each plot represents a 1 kpc linear scale

of the sources. These maps reveal a wide variety of molecular gas morphologies: extended disks (i.e., NGC 1482, NGC 4666), ring-like structures (i.e., NGC 134, NGC 157, NGC 1084), large-scale arms (i.e., NGC 3110, NGC 4030), nuclear arms (i.e., NGC 613, NGC 986, NGC 1087, NGC 2559, NGC 5247), as well as elongated nuclear barred structures (i.e., NGC 1385, NGC 4691 and NGC 5247).

It is remarkable that the molecular gas distribution peaks at the center for 86% of the galaxies in this subsample. When we restrict to galaxies for which the linear scale is better than 300 pc/″, in total $N = 24$ galaxies, we find that these galaxies have a characteristic Gaussian width for its circumnuclear components of < 1 kpc, and in average of about 0.5 kpc. An exception is NGC 4030, which has a width of about 2 kpc. The galaxies that do not peak at the center show ring structures or nuclear bars with nuclear spiral arms emerging from it. These galaxies are NGC 134, NGC 1385, NGC 4691, and NGC 5247. NGC 134 shows a wide ring feature at 1.5 kpc. NGC 1385, NGC 4691, and NGC 5247 have quite asymmetric and elongated structures, probably nuclear bars, and with spiral arms emerging from them (NGC 4691 and NGC 5247).

We plot in Fig. 3a the global star formation (SF) law represented as the 8–1,000 μm FIR luminosity versus the total gas mass M_{gas} calculated as $1.36 \times M_{H_2}$, where the 1.36 factor is to take into account species other than hydrogen. The scatter is not large if we take into account that we might have missed extended CO(2–1) emission in our interferometric observations and suggests that most of the SF is taking place in the inner 1′, where most of the molecular gas is.

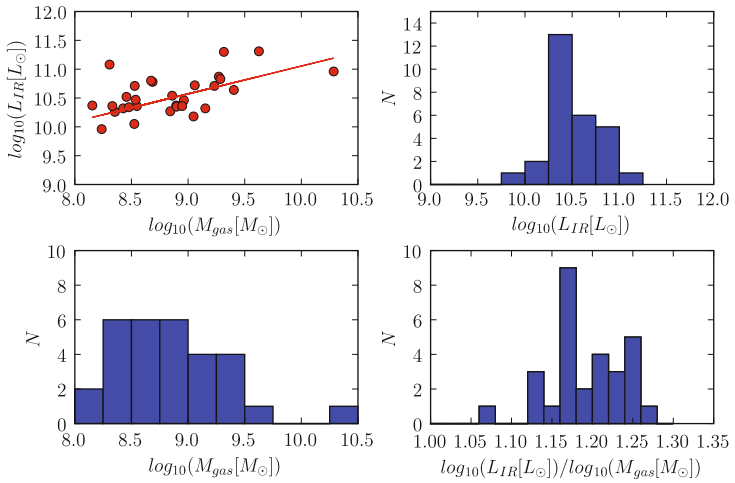


Fig. 3 **a** M_{gas} vs L_{IR} in decimal logarithmic scale, as well as a *line* showing the fit to the data points. The slope is 1.21 ± 0.05; **b** L_{IR} distribution; **c** M_{gas} distribution; **d** M_{gas} / L_{IR} distribution

In this preliminary study, we do not find any correlation between gas concentration and existence of AGN. Actually only ∼ 10% of the galaxies in this subsample are confirmed AGNs (NGC 1667, NGC 4418, and probably NGC 613), although note that the classification may not be complete. Large starbursts are apparent, especially for barred galaxies such as NGC 1808 and NGC 613 as well as in galaxies with the largest optical perturbations, such as NGC 1022. Most of our galaxies (∼ 70%) have been classified as starburst and/or nuclear HII.

4 Discussion

Note that most of the galaxies with CO interferometric maps in the literature are IR-bright galaxies too. In particular, although the selection criteria of the BIMA-SONG, NUGA, and NRT-OVRO samples are different, ∼ 90% of the galaxies there fulfill our IR-flux criterion. The distribution of IR luminosity in the B0DEGA subsample spans a range from 10.0 to 11.3 $L_\odot$, with an average of 10.6 $L_\odot$ (see Fig. 4). The L_{IR} distribution in our sample presents in average larger values with respect to the previously mentioned samples of nearby galaxies. This effect is specially important with the BIMA-SONG sample. However, it shows low values when compared with

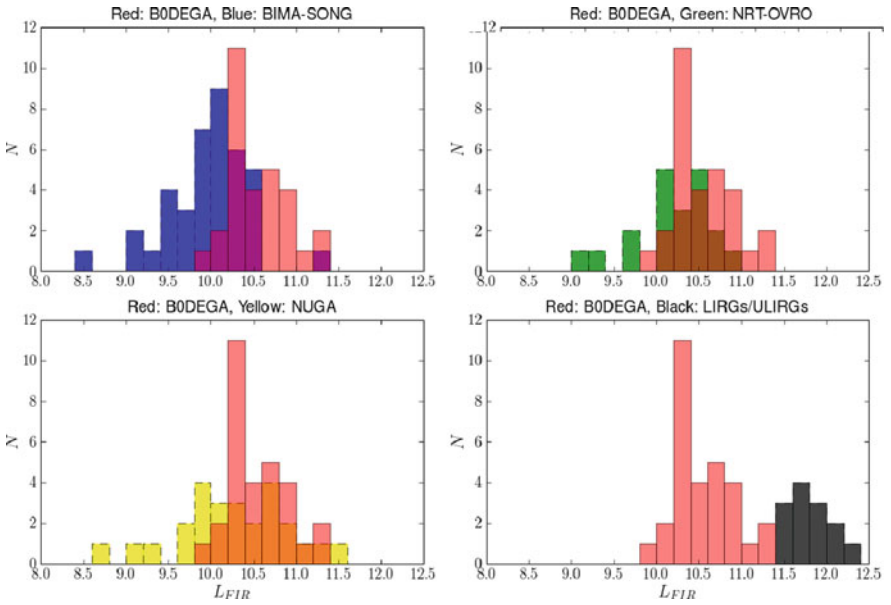


Fig. 4 Comparison of the L_{IR} distribution of our sample with that of other studies in the literature: BIMA-SONG (*blue histogram*), NUGA (*yellow*), and LIRG/ULIRGs (*black*)

ultra/luminous IR objects (LIRGs/ULIRGs) with $L_{IR} > 11$ L$\odot$, which are usually strongly interacting or merger systems located at larger distances. This indicates that our sample represents galaxies whose properties are between normal galaxies and mergers. On the other hand, the stellar content distribution as traced by L_B for our galaxies is not significantly different to that of other samples, suggesting that our objects have an enhancement of SF rate when compared with similarly massive and nearby galaxies.

The rate of galaxies with molecular gas peaking at the center seems to be larger (by $\sim 45\%$) in our subsample than in BIMA-SONG galaxies, which in principle is not likely to be due just by the use of different transitions of CO (J=2–1 and J=1–0, respectively). Thus, the larger concentration rate of molecular gas in our sample may mean that it is partly caused by recent interaction events, which probably trigger bar formation and other mechanisms that drive a large amount of gas to the center. Note that the rate of bars in BIMA-SONG and our subsample is similar, 67% versus 73%. However, in our sample the central regions might be in an evolutionary stage where the gas driven toward the center has not been fully consumed. Despite the central concentrations, most of the galaxies have not been classified to host an AGN. On the other hand, starburst and nuclear HII are apparent in many of the galaxies.

Acknowledgments We thank the SMA staff members who made the observations reported here possible. D.E. was supported by a Marie Curie International Fellowship within the 6th European Community Framework Programme (MOIF-CT-2006-40298).

References

1. Barnes, J.E., Hernquist, L.E. 1991, ApJ, 370, L65–L68
2. Espada, D., Matsushita, S., Peck, A., Henkel, C., Iono, D., Israel, F.P., Muller, S., Petitpas, G., Pihlström, Y., Taylor, G.B., Dinh-V-Trung 2009, ApJ, 695, 116–134
3. García-Burillo et al. 2003, A&A, 407, 485–502
4. Helfer, T.T., Thornley, M.D., Regan, M.W., Wong, T., Sheth, K., Vogel, S.N., Blitz, L., Bock, D.C.J. 2003, ApJS, 145, 259–327
5. Helou, G., Walker, D.W. eds. 1988, NASA, Washington, DC
6. Sakamoto, K. et al. 1999, ApJ, 525, 691–701
7. Sheth, K., Vogel, S.N., Regan, M.W., Thornley, M.D., Teuben, P.J. 2005, ApJ, 632, 217–226
8. Shlosman, I., Frank, J., Begelman, M.C. 1989, Nature, 338, 45–47
9. Wilson, C.D., Petitpas, G.R., Iono, D., Baker, A.J., Peck, A.B., Krips, M., Warren, B., Golding, J., Atkinson, A., Armus, L., Cox, T.J., Ho, P., Juvela, M., Matsushita, S., Mihos, J.C., Pihlstrom, Y., Yun, M.S. 2008, ApJS, 178, 189–224

The DiVA's Mask: Iconifying Galaxies and Revealing HI Anomalies

Jayanne English, Jason Fiege, Theresa Wiegert, Baerbel Koribalski, Wolfgang Kerzendorf, and Kenneth C. Freeman

Abstract We introduce GalAPAGOS (Galaxy Astrophysical Parameter Acquisition by Genetic Optimization Software), a software package that models the density distribution and velocity behaviour of HI emission using the 3-D information in data cubes. We explore creating 3-D masks from model galaxy cubes generated using parameters determined by GalAPAGOS. We call our technique of applying these masks to the original HI data cubes, and subsequent analysis of the faint outer regions of galaxies, "DiVA: DIstribution of Velocity Anomalies." In this preliminary investigation we study the extraplanar features of the nearby, edge-on, disk galaxy NGC 2188.

1 Introduction

There are many possible origins for neutral hydrogen (HI) gas structure in the halos of disk galaxies. An excellent review of the literature exploring the possibilities is given by [4]. Briefly, processes internal to the disk predict the existence of blowout features such as extended, vertical spurs associated with chimneys or walls of superbubbles and fast moving halo clouds. Alternatively structures observed away from the plane could be associated with external processes such as tidal debris or inflow from a cooling halo or the "cosmic web." The interaction of galaxies could thicken disks and generate warps. Discerning the most appropriate scenario can be facilitated by examining the velocity behaviour as well as the spatial distribution of HI features beyond the disk.

We explore applying preliminary models of the disk of edge-on galaxy NGC 2188 as masks to reveal the distribution of anomalies away from the plane. Culturally, masks iconify particular roles and delineate expectations. For example, a Diva or Divo in makeup and costume personifies a particular type of beauty, such

J. English (✉)
University of Manitoba, Winnipeg, MB, Canada R3T 2N2
e-mail: jayanne_english@umanitoba.ca

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_8,

as symmetry of common features (eyes, mouth). In analogy our models delineate the common features of a disk, including the warp. What is noticeable beyond the mask in the performing Diva/Divo are their particular deviations from the norm, their characteristics that make them uncommon. In the DiVA aspect of our study of galaxies we also use the mask to highlight the uncommon, in this case the distribution of spatial features and their potential velocity anomalies.

To create our masks we use the output from our GalAPAGOS modelling software which generates parametric solutions for the spatial distribution of HI gas and rotation curves. This package models a three-dimensional rotating HI disk, simulating the emission, and minimizing the χ^2 goodness of fit by comparing model solutions with a real data cube. The computational challenge lies in finding the global χ^2 minimum over a parameter space that is necessarily large for a realistic model. GalAPAGOS performs this search by means of a state-of-the-art genetic algorithm called Ferret (§ 2). An important advantage of GalAPAGOS is that only broadly defined ranges are required for each parameter. The code samples the parameter space thoroughly and does not need to be started close to the final best-fitting solution. Thus, the burden on the user is decreased, while also ensuring greater reliability than other methods. As GalAPAGOS explores its parameter space, it automatically maps out the distribution of χ^2 values near the best solution, thus providing confidence intervals on all model parameters. The modelling equations can be defined by the user and the parametric solutions resulting from GalAPAGOS can be computed at arbitrary resolution, which is useful for visualizing the spatial and kinematic structure of galaxies. Moreover, the high-quality rotation curves can be used for mass modelling.

Here we apply GalAPAGOS to NGC 2188 ($\alpha = 06\,\mathrm{h}, 10\,\mathrm{m}, 09.5\,\mathrm{s}$, $\delta = -34°06'22''$) which is a SB(s)m galaxy at a distance of 7.9 Mpc [3]. In VLA observations, by Domgorgen, Dahlem & Dettmar (1996) with $12''$ ($\sim$ 460 pc per beam) resolution, this galaxy appears to have a worm-like filament emanating from the southwest as well as 2 spurs extending from the central region of the galaxy towards the east. They interpret these spurs as the walls of an expanding bubble, noting, however that the diffuse ionized gas (DIG) emission does not correlate with these features. Interestingly, they propose that the banana shape of NGC 2188's disk could be due to interaction with an intergalactic HI cloud.

We present Local Volume HI Survey (LVHIS) observations with an average resolution of 2 kpc per beam along with a Siding Spring Observatory R-band image.

2 Observations, Analysis, and Modelling

LVHIS consists of deep HI line and 20 cm radio continuum Australia Telescope Compact Array observations for all nearby, gas-rich galaxies closer than 10 Mpc and DEC $< -30°$. The sample is catalogued at
http://www.atnf.csiro.au/people/bkoribal/LVHIS/ along with a list of papers. For the following analysis we use naturally weighted 21 cm spectral line data of NGC 2188

collected 15 January 2006. Processed in the usual way using Miriad, the cube covers the heliocentric velocity range of 646–842 km s^{-1} with a Hanning smoothed channel width of 4 km s^{-1}. The beam size is 76.8 × 46.8 arcsec. The r.m.s.(= 1 σ) is 1.7 mJy/beam.

To remove the noise in the data cube for analysis purposes we first created a smoothed data cube. This was applied as a mask such that only intensity values in the data cube that were above 3 sigma in the smoothed cube were retained, generating a "noise-reduced" data cube for easier interpretation.

R-band images were acquired, for a total 1 h integration time, using the Imager on 2.3 m telescope at Siding Spring Observatory 27 and 28 February 2009. These were reduced in the usual manner using Pyraf.

GalAPAGOS (Galaxy Astrophysical Parameter Acquisition by Genetic Optimization Software) uses the Ferret Genetic Algorithm as its graphical user interface and optimization engine. Ferret is the principle optimizer in the Qubist Global Optimization Toolbox for MATLAB, which was developed by J. Fiege (www.nqube.ca). The functional form of the empirical models used for our preliminary runs [2] is as follows:

The rotation curve

$$v_\phi = v_0 \tanh\left(\frac{r}{r_{0,v}}\right)\left[1 + a_v\left(\frac{r}{r_{\text{out}}}\right)\right] \tag{1}$$

where $r_{0,v}$ determines the radius where the tanh function turns over, r_{out} is the outer radius of the galaxy where the density falls off, and a_v is a coefficient that controls the asymptotic slope of the rotation curve past $r_{0,v}$.

The HI distribution:

$$N_H(r) = \frac{N_{H,0}}{2}\left[\tanh\left(\frac{r_{\text{out}} - r}{\Delta r_{\text{out}}}\right) + 1\right] \cdot SDM(r; \mathbf{a}), \tag{2}$$

where the density falls off to zero over a width approximately Δr_{out} and SDM is a cubic spline function defined by N coefficients that modulate density enhancements and depletions.

We also model the scale height and assume that the axis symmetry is broken by a warp. Our model differs from tilted ring models in that our rings are linked together by a smoothly varying function of radius to model the extent of the warp and the azimuthal variation of its direction.

For input into GalAPAGOS, the 21 cm data cube was binned by 3 pixels in each spatial dimension and 2 pixels along the velocity axis and the noise level of the brightness temperature (T_B) in the cube was determined. During this preliminary run 2,811 solutions were found where the T_B in the model was within 1 σ of the T_B of the data. To examine particular solutions we scale the parameters to the resolution of the original data cube and use these to create a full resolution model galaxy data cube in fits format.

3 Preliminary Results

The integrated intensity map contours are overlaid on DSS red image in Fig. 1. In comparison to the optical disk, the HI extends more towards the south than to the north. We have not determined that there is a starless HI cloud in the field of view of our data cube as suggested by previous authors (§ 1).

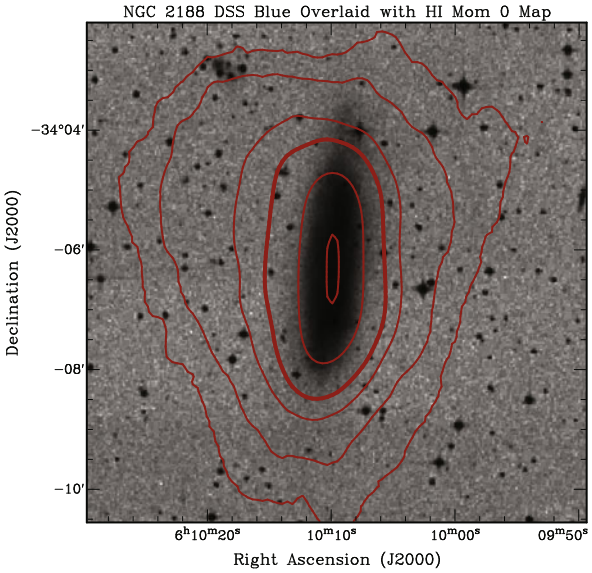


Fig. 1 Integrated intensity HI map overlaid on DSS *blue image*. The noise-reduced cube was used to generate the zeroth moment map. The contours, in Jy beam^{-1} × km s^{-1}, are 0.1, 0.3, 1.0, 2.0, 5.0, 9.0

3.1 Iconifying Galaxies

GalAPAGOS generates a family of solutions within 1 σ (§ 2). Figure 2 plots the deprojected rotation curves of the set while Fig. 3 plots the HI distribution. None of the rotation curves are declining. However, density distributions with both depletions and non-depletions of gas in the centre are valid.

3.2 Revealing HI Anomalies

Visual inspection shows that the model disk of the best fit solution has a regular shape at a contour level of 9 mJy/beam (5 σ of the noise in the data cube). We use the model cube to mask all values above this level in the "noise-reduced" data

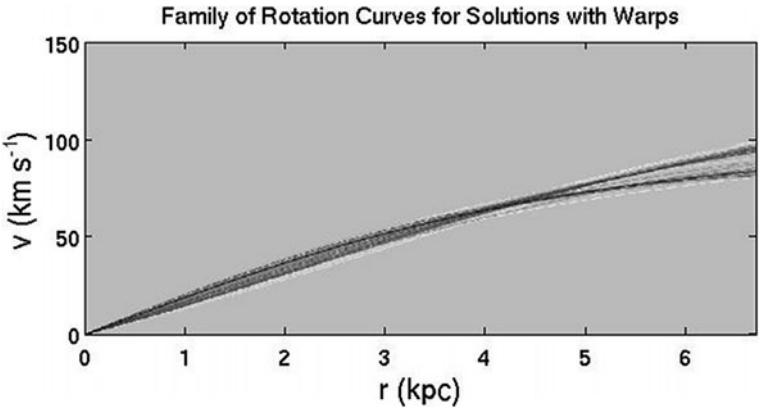


Fig. 2 Family of solutions to the rotation curve model. The solutions with lower χ^2 values have *darker lines* in the plot

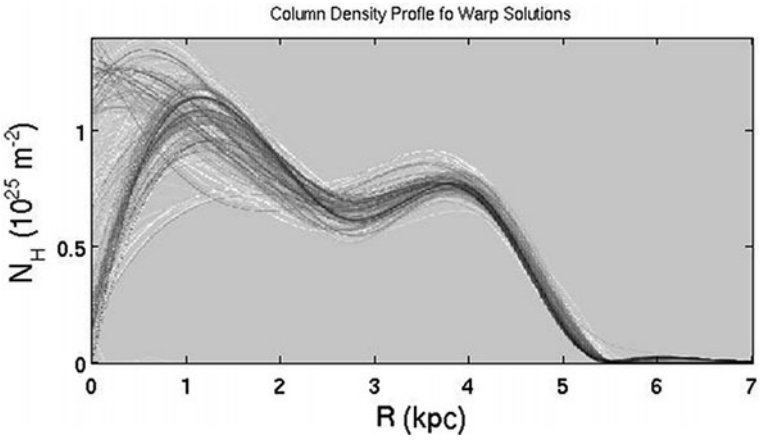


Fig. 3 Family of solutions to the HI density model. The solutions with lower χ^2 values have *darker lines* in the plot. As expected when modelling inclined galaxies, the centre can be filled or have a hole

cube. This removes the disk in each channel and reveals faint structures that may be extraplanar. We call the result of this masking the DiVA cube. The various extended structures on the east side of NGC 2188 include an extension towards ESO 364-29; see Fig. 4. At the resolution of our HI data ESO 364-29 does not display HI. However, ESO 364-29 has a systemic velocity of 787 km s^{-1} while NGC 2188 has a systemic velocity of 747 km s^{-1} so these galaxies could be interacting and this extension could be a tidal feature. Yet the interpretation is not clear since an apparent background 2MASS galaxy, J06101414-3404429, lies close to this extension; Fig. 4. An optical spectrum for its systemic velocity is required to determine if it contributes HI emission to this feature.

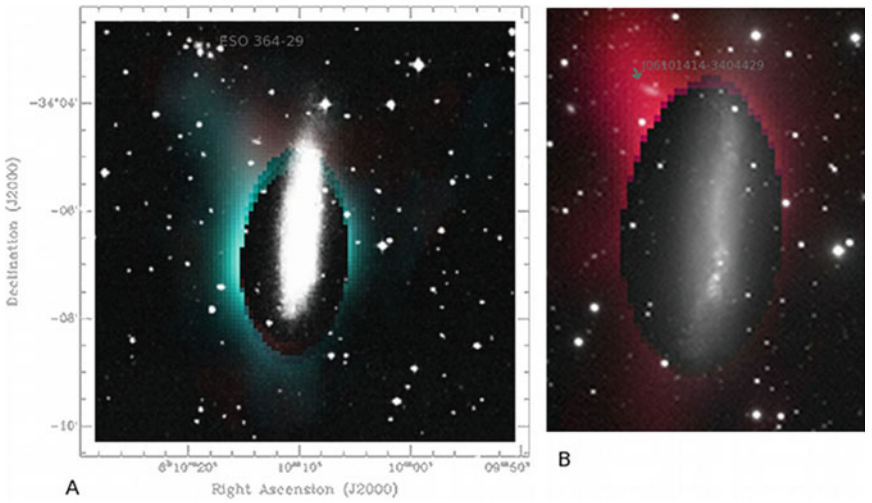


Fig. 4 **a** Northeast HI extension towards ESO 364-29. Three colourized channels from the DiVA (*masked*) cube covering velocities from 730 km s^{-1} (*turquoise*) to 746 km s^{-1} (*red*) overlaid on a *blue* DSS image. **b** J06101414-3404429 and HI feature. HI with velocities 750 and 754 km s^{-1} in the DiVA cube overlaid on an R-band image from Siding Spring Observatory

Like Domgorgen et al. [1] we find extensions perpendicular to the middle of the disk; Fig. 5. However, their apparent "worms" are centred at 750 km s^{-1} while our features are predominate at $\sim$ 710 km s^{-1}. At the resolution of our data we do not detect an extension of their southwest "worm."

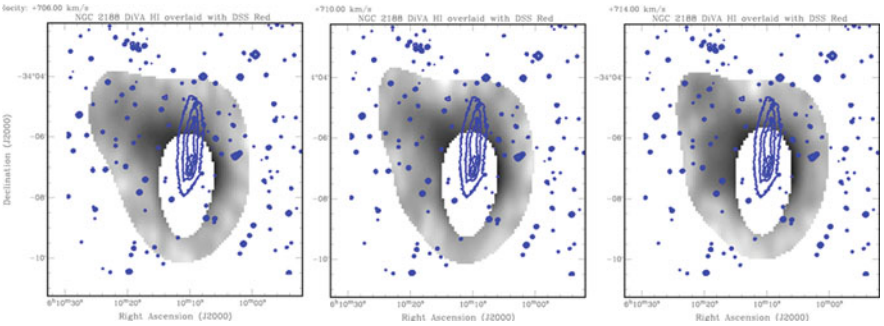


Fig. 5 HI Extensions over central and star forming regions. Three channels from the DiVA (*masked*) cube with the DSS *red* image in contours from 20 to 90 % in steps of 20%

4 Summary and Future Plans

Our preliminary DiVA exploration shows a number of HI extensions perpendicular to the east side of the plane of NGC 2188. One is towards ESO 364-29 indicating that NGC 2188's distorted, warped appearance may be due more to a galaxy–galaxy interaction rather than due to star formation blowouts from the disk interior. We plan other GalAPAGOS runs on NGC 2188 as well as a continued exploration of masking and other visualization techniques. Using the SSO data we will model the mass in NGC 2188.

Future plans for GalAPAGOS include adding greater asymmetry in addition to our asymmetric warp and exploring other mathematical models for the disk. GalAPAGOS will be a useful analysis tool for the HI galaxy surveys by ASKAP and MeerKAT in preparation for the SKA.

Acknowledgments We acknowledge the inspiration from the Diva/Divo Kingston project.

References

1. Domgorgen, H. Dahlem, M, Dettmar, R.J. 1996, Astron Astrophys, 313, 96
2. Fiege, J., van Vliet Wiegert, T., English, J., in progress
3. Miller, S.T., Veilleux, S. 2003, ApJ Sup, 148, 383
4. Rand, R.J., Benjamin, R.A. 2008, ApJ, 676, 991

Enigmatic Masks of Cosmic Dust: Lessons from Nearby Galaxies Through the Eyes of the Spitzer Space Telescope

Robert Groess, David L. Block, and Giovanni G. Fazio

Abstract Evident from early observations of barred galaxies, such as NGC1300, NGC1530, NGC3351 and NGC5921 by Curtis (1918 Descriptions of 762 nebulae and clusters photographed with the Crossley reflector. Publications of the Lick Observatory, XIII, 11–42), was the prevalence of characteristic dark bands or streaks across the disk of the galaxies. These dark bands were interpreted to be lanes of light obscuring "occulting material". While being nothing more than a nuisance to observers at the turn of the twentieth century, we now turn to the lanes of dust themselves, seen glowing in emission at 8.0 μm, to understand more of the internal dynamics of *barred* galaxies. We find a separation of 3.6 μm images of nearby barred galaxies into two distinct classes or form families. The first class comprises galaxies whose 8.0 μm morphology tightly traces the underlying barred stellar backbone at 3.6 μm, while the second class does not reveal a bar at all at 8.0 μm. Spectacular dust lanes are, however, evident in this second class and are understood to be signatures of shock loci at the outer edges of the bar resulting in dust cascading down the gravitational potential of the galaxy, towards the nucleus.

1 Introduction

It was already in 1918 that Heber D. Curtis [9] made the observation that

> there is one fairly common type of spiral [galaxy] . . . Its main characteristic is a band of matter extending diametrically across the nucleus and inner parts of the spiral. Frequently the whorls in this type form a nearly complete ring; in other examples the whorls appear to begin at the ends of this cross-arm. The general appearance is that of the Greek letter ϕ.

Curtis [10] goes on to say, in a masterful paper entitled, "A Study of the Occulting Matter in the Spiral Nebulae",

R. Groess (✉)
School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa
e-mail: robert.groess@wits.ac.za

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_9,

> the occurrence of such dark bands in the spiral nebulae is a relatively common phenomenon. This fact can scarcely fail to be of *great importance* in the study of this highly interesting and *still imperfectly understood* class of celestial objects. [Italics, Author]

That barred galaxies would present an important niche in the study of extragalactic nebulae was always anticipated. Less so was the rôle that dust would come to play in the next century, and possibly beyond, in what is still today, in the words of H.D. Curtis, an "imperfectly understood class of celestial objects." Barred galaxies have continued to hold many of their secrets shrouded in mystery. In the Hubble Atlas, Allan Sandage [19] remarks of NGC1300 and NGC5383,

> [One can see] dust lanes leaving the nucleus, one on each side of the bar, and extending into the spiral arms.

The mysterious "dust lanes" have indeed vexed astronomers over the ages. Only recently, with the advent of large format near-infrared detectors, such as those on the Spitzer Space Telescope, have we been able to peal away the layers which constitute the mask.

2 Two Form Families of Stellar-Barred Spirals at 8.0 μm

The Infrared Array Camera (IRAC) onboard Spitzer brings us a dramatic new view on the morphology and structure of warm dust glowing in emission at 8.0 μm. The Rayleigh–Jeans tail of most population I and II stars is already strongly attenuated at 8.0 μm, where the contribution from warm dust in emission dominates. We find evidence for at least two distinct classes or form families of 8.0 μm morphology of nearby spiral galaxies which are barred at 3.6 μm (see Fig. 1). These two classes are

(1) Barred galaxies where the 8.0 μm warm dust morphology tightly traces the older population I stars at 3.6 μm, i.e. the galaxy is unmistakably identifiable at 8.0 μm, compared with its 3.6 μm morphology.

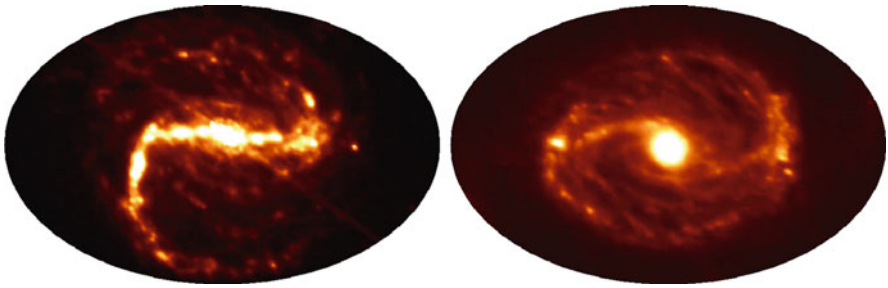


Fig. 1 Type 1 barred spiral galaxy NGC7479 seen at 8.0 μm (on the *left*), tightly traces the underlying stellar morphology when compared with its 3.6 μm morphology. NGC3368 (at *right*) is an archetype of the Type 2 form family, with no obvious trace of a bar at 8.0 μm. Two prominent curved spiral arms are evident however. These strikingly reveal loci of shocks in the presence of the bar potential seen as warm dust glowing in emission

(2) S-shaped dust spirals at 8.0 μm of varying curvature for different galaxies, with no obvious trace of a bar-like feature.

We present these two cases in further detail.

2.1 Type 1: Tightly Coupled Morphologies at 3.6 and 8.0 μm

Strongly barred galaxies such as NGC7479, NGC4314 and NGC1097 (see Fig. 2) present qualitatively and quantitatively similar or near-identical morphologies when observed at 3.6 and 8.0 μm. The similarities are so striking that subjecting these

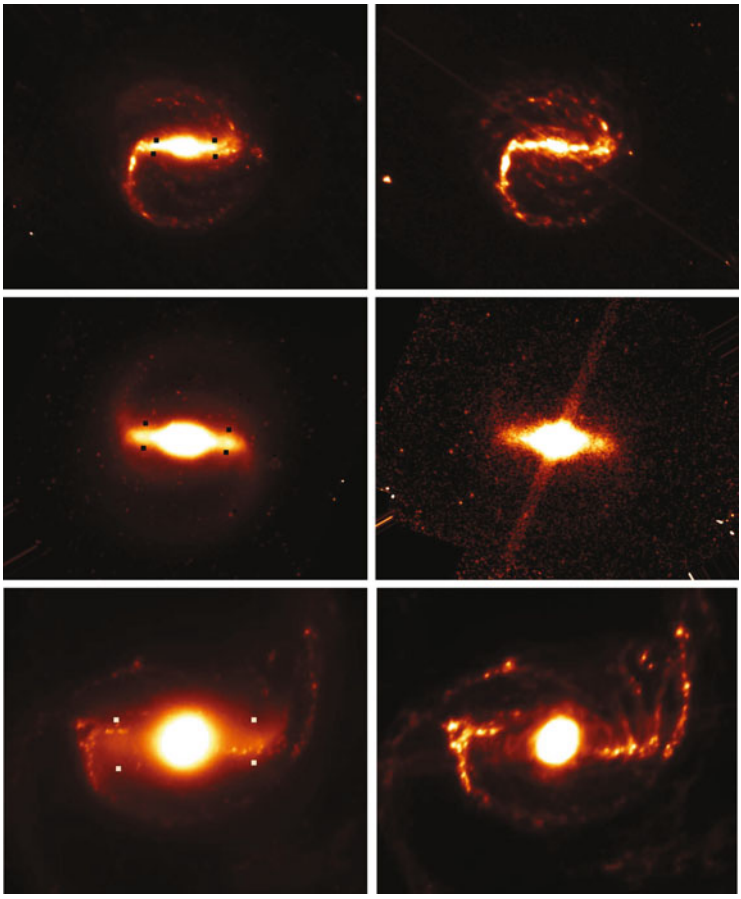


Fig. 2 Type 1 barred spiral galaxies NGC7479, NGC4314 and NGC1097 with a tightly coupled morphology between 3.6 μm (*left column*) and 8.0 μm (*right column*). These *bars* contain both stars and dust, with the dust distribution being a very good tracer of the underlying stellar morphology. The *four squares* delineate the positions of the torque maxima for each bar at 3.6 μm

galaxies to a Gravitational Torque Analysis as developed by Combes and Sanders [7] and commonly referred to as the Gravitational Torque Method (GTM) by Buta and Block [6] reveals relative gravitational torque strengths (Q_g) which are generally very similar when compared between the two channels. The best case in our sample is NGC7479 with a $Q_{g(3.6)}$=0.65±0.09 and $Q_{g(8.0)}$=0.65±0.17

The GTM calculates the ratio of the maximum tangential force with respect to the mean axis-symmetric radial background force and selects the maximum value over all radii. This single figure is the relative gravitational torque strength Q_g:

$$Q_{\mathrm{g}}(R) = \frac{F_{\mathrm{T}}^{\max}(R)}{F_0(R)} = \frac{\left(\frac{\partial \Phi(R,\varphi)}{\partial \varphi}\right)_{\max}}{R\frac{\partial \Phi_0(R)}{\partial R}} \tag{1}$$

Although the GTM was developed for determining relative gravitational torques at near-infrared wavelengths, where M/L is found to be nearly constant within the luminous disks of spiral galaxies [12, 15] applying this procedure to the "Type 1" 8.0 μm images results in very similar "Q_g" values. While these $Q_{g(8.0)}$ values cannot be interpreted as gravitational torques per se, as that would require knowledge of the three-dimensional mass distribution of a galaxy embedded in a gravitational potential, we propose these values provide a way of measuring the similarity between the older stellar population I backbone and warm dust morphologies of the same galaxy at different wavelengths, respectively.

2.2 *Type 2: Straight and Curved Dust Lanes in Emission, Without Trace of a Bar*

In NGC3368, NGC4725 and NGC4303 (see Fig. 3), very pronounced curved lanes of dust are evident as well as the complete lack of any indication of a bar. For example, NGC3368 shows tightly wrapped ring-like arms [11] at optical wavelengths with a moderate bar strength of $Q_g = 0.24 \pm 0.01$ [13] at 3.6 μm, but no bar is evident at all at 8.0 μm.

The shape and spatial location of dust lanes in spiral galaxies have been modelled in a number of theoretical studies by focusing on the hydrodynamic response of gas and dust in disks, in close proximity to rotating bars. Prendergast [17] pioneered hydrodynamic models including the phenomena of gas shocks in the neighbourhood of a bar. Early dust velocity field predictions are detailed in Huntley [14], which have subsequently been summarized by Prendergast [18] for SBb and SBc galaxies. These were further investigated by Athanassoula [3] who predicts two distinct dust lane morphologies: (i) straight lanes and (ii) concave curved lanes, where the concave side is towards the bar major axis. In these models, the dust lanes appear on the leading edges of bars as defined by a trailing set of spiral arms.

Athanassoula [4] presents detailed and comprehensive simulations showing the degree of curvature of dust lanes coincides with different optical bar strengths. Ann and Lee [1] performed smoothed particle hydrodynamics (SPH) simulations of the

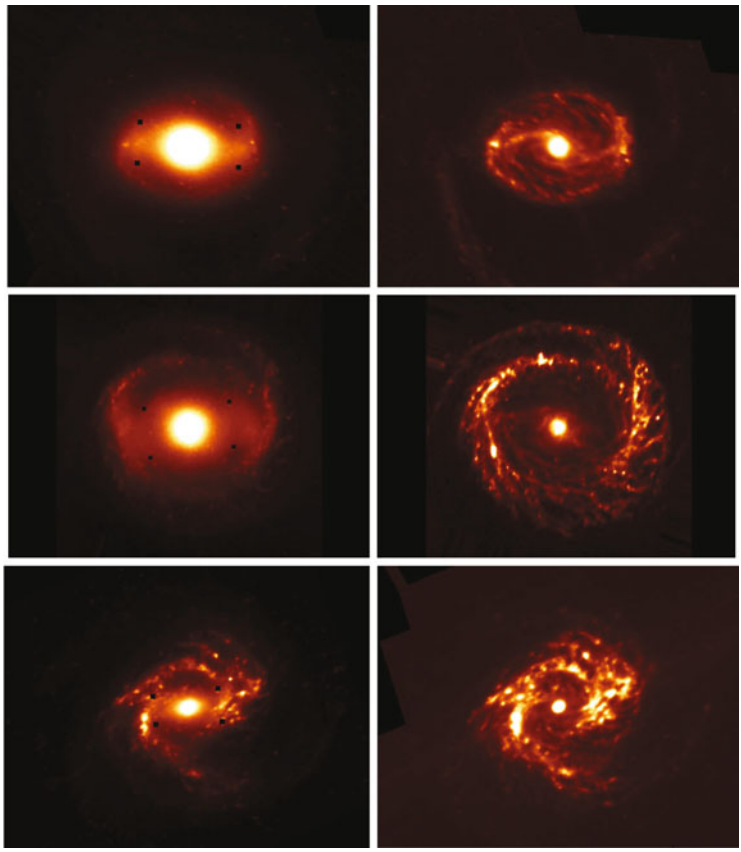


Fig. 3 Spectacular spirals or filaments of dust in emission are observed to connect the spiral arms with the nuclear regions of NGC3368, NGC4725 and NGC4303. The galaxies have a definite stellar bar at 3.6 μm (*left column*) but are totally devoid of any bar-like features at 8.0 μm (*right column*)

evolution of gaseous disks and present gas morphologies of a range of artificial barred galaxies. Comerón et al. [8], using a sample of 55 barred spiral galaxies, observationally verify in a rigorous dust lane curvature treatment that strong bars (high Q_g) imply dust lanes with low curvature, whereas weak bars could harbour dust lanes with a wide range of curvature.

Dust lanes are understood to be loci of shocks which occur where the velocity fields of gas and dust due to both families of x_1 orbits (with long axis parallel to the bar axis) and x_2 orbits (perpendicular) intersect and experience an abrupt velocity change in magnitude and direction. A shear field is set up in these regions and the velocity vectors of the gas and dust abruptly turn inwards towards the galactic nucleus. These velocity fields are likened to a spectacular waterfall where shock fronts induce vast streams of galactic dust cascading down the gravitational potential towards the centres of galaxies. A well-known prototype for this pattern is

NGC1300, which is classified by Hubble as SBb(s). Optical images show there are two straight dust lanes which are on symmetrically opposite sides of the bar, on the leading edges as defined by the trailing sense of the spiral arms.

The velocity vectors also reveal the curvature around the orbit apo-centres. Here a higher curvature at the x_1 orbits apo-centre results in dust lanes that are less curved. Thus very straight dust lanes are expected where there is a rapid change in direction for the gas and dust velocity field at the outermost orbits where they cross the bar long axis. Athanassoula investigates various apo-centre curvatures using second-order flux-splitting models [20] and predicts that strong bars harbour relatively straight dust lanes. The gas and dust density, along with the velocity field vectors, are seen in Athanassoula [4] Figs. 2a,b. Here bright areas in Fig. 2a indicate regions of higher gas density and the presence of flocculent spurs emanating from the gas and dust lanes at the shock positions reveal a well ordered but intricate velocity field along the shock loci. The velocity field vectors in Fig. 2b have a length which corresponds to the magnitude of the gas and dust velocity at points that are equidistant along a grid.

As regards gas and dust transport, in cases where no shocks are evident there is no net inflow towards the nucleus. However, in cases with shocks, the gas flow is considerably more complicated [16]. The trailing sides of the bar exhibit gas outflow where the gas density is very low. This outflow continues until it reaches the shock front, where it abruptly turns inwards and results in the formation of dust lanes. Ann [2] further investigates the formation of nuclear gas and dust rings by making extensive use of SPH simulations.

Since dust lanes are understood to be regions of higher gas density and dust concentrations, this raises the question of whether this should influence the local star formation rate. The reason star formation is not prevalent at all is attributed to the high shear encountered in straight dust lanes. Molecular clouds will thus tend to shear out before they have time to collapse, preventing any star formation in that region. Dust lanes with a higher degree of curvature, however, may well harbour more benign environments for incipient star formation.

3 Conclusion

We know that the optical component of disk galaxies may represent an extremely poor tracer of the underlying stellar backbone. Block and Wainscoat [5] present a strong case in favour of a true duality of spiral structure at optical versus infrared wavelengths. We have now found that there is also a decoupling in the distribution of warm dust grains in the bars of a significant number of nearby galaxies. The degree of decoupling is proportional to the uncertainty that arises in predicting the morphology of warm dust in emission compared with underlying stellar component. Of the nearby galaxies in our sample that present a well-defined bar at 3.6 μm, there are two distinct classes or form families evident. Type 1 are tightly coupled to the 3.6 μm morphology at 8.0 μm, while Type 2 do not reveal a bar at all at

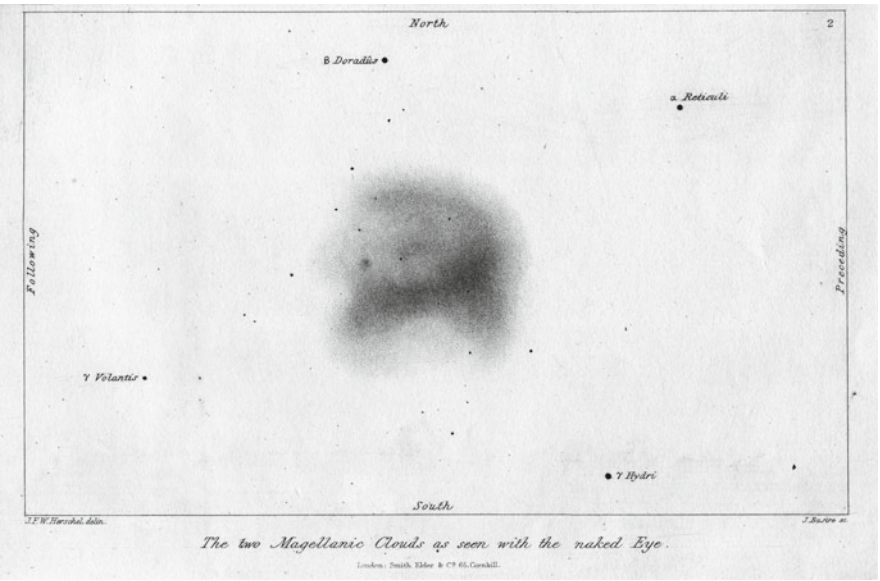
8.0 μm. Instead, very distinct dust lanes are evident whose curvature is limited by the strength of the stellar bar. A strong bar at 3.6 μm implies straight dust lanes, while a weak bar could have a range in dust lane curvatures.

It would seem that just as the dynamical mass distribution of a galaxy cannot, a priori, be predicted or inferred from its transient Population I morphology (Hubble Type) so too the morphology of dust emission in a barred spiral galaxy cannot be definitively predicted from optical photographs or even images at 3.6 μm.

Acknowledgments R.G. would like to thank the University of the Witwatersrand for the generous FRC Startup Grant enabling attendance of this conference as well as the School of Computational and Applied Mathematics for additional funding. A great word of appreciation also goes to the members of the IRAC team at the Harvard-Smithsonian Center for Astrophysics, in particular (alphabetically) Matt Ashby, Michael Pahre and Steve Willner. In conclusion, we wish Professor Ken Freeman (FRS) many happy returns on this occasion of his 70th birthday.

References

1. Ann, H.B., Lee, H.M. 2000, JKAS, 33, 1
2. Ann, H.B. 2001, JKAS, 34, 261
3. Athanassoula, E. 1984, Phys Rep, 114, 319
4. Athanassoula, E. 1992, MNRAS, 259, 345
5. Block, D.L., Wainscoat, R.J. 1991, Nature, 353, 48–50
6. Buta, R., Block, D.L. 2001, ApJ, 550, 243
7. Combes, F., Sanders, R.H. 1981, A&A, 96, 164
8. Comerón, S., Martinez-Valpuesta, I., Knapen, J.H., Beckman, J.E. 2009, ApJ, 706, L256
9. Curtis, H.D. 1918a, Descriptions of 762 nebulae and clusters photographed with the crossley reflector. Publications of the Lick Observatory, California, XIII, 11–42
10. Curtis, H.D. 1918b, A study of occulting matter in the spiral nebulae. Publications of the Lick Observatory, California, XIII, 45–54
11. Elmegreen, D.M., Elmegreen, B.G. 1987, ApJ, 314, 3
12. Freeman, K.C. 1992, Physics of Nearby Galaxies: Nature or Nurture? eds. T.X. Thuan, C. Balkowski, J. Tran Thanh, Editions Frontiere Gif-sur-Yvette, p. 201
13. Groess, R. 2007, Gravitational Torques of the Dust Penetrated Stellar Backbone of Extragalactic Spiral Disks, PhD Thesis, University of the Witwatersrand
14. Huntley, J.M. 1978, ApJ, 225, 101
15. Mathewson, D.S., Ford, V.L., Buchhorn, M. 1992, ApJS, 81, 413
16. Piner, B.G., Stone, J.M., Teuben, P.J. 1995, ApJ, 449, 508
17. Prendergast, K.H. 1962, IAUS, 15, 126
18. Prendergast, K.H. 1983, IAUS, 100, 215
19. Sandage, A. 1961, The Hubble Atlas of Galaxies, Carnegie Institution of Washington, Washington, DC
20. van Albada, G.D., Roberts, W.W. Jr. 1981, ApJ, 246, 740


North
2
β Doradûs
α Reticuli
Following
Preceding
γ Volantis
γ Hydri
South
J.F.W. Herschel. delin.
The two Magellanic Clouds as seen with the naked Eye.
London: Smith, Elder & Cº 65, Cornhill.

The Large Magellanic Cloud: A Power Spectral Analysis of Spitzer Images

Ivânio Puerari, David L. Block, Bruce G. Elmegreen, and Frédéric Bournaud

Abstract We present a power spectral analysis of Spitzer images of the Large Magellanic Cloud. The power spectra of the FIR emission show two different power laws. At larger scales (kpc) the slope is ~ -1.6, while at smaller ones (tens to few hundreds of parsecs) the slope is steeper, with a value ~ -2.9. The break occurs at a scale $\sim 100 - 200$ pc. We interpret this break as the scale height of the dust disk of the LMC. We perform high-resolution simulations with and without stellar feedback. Our AMR hydrodynamic simulations of model galaxies using the LMC mass and rotation curve confirm that they have similar two-component power laws for projected density – and that the break does indeed occur at the disk thickness. Power spectral analysis of velocities betrays a single power law for in-plane components. The vertical component of the velocity shows a flat behavior for large structures and a power law similar to the in-plane velocities at small scales. The motions are highly anisotropic at large scales, with in-plane velocities being much more important than vertical ones. In contrast, at small scales, the motions become more isotropic.

1 Introduction

The LMC belongs to the de Vaucouleurs classification bin SB(s)m and is ~ 50 kpc away (based on a distance modulus $m - M = 18.50 \pm 0.10$, following Freedman et al. [9]). The LMC is thus an ideal laboratory for studies of both morphology and turbulence, due to the high-spatial resolution possible due to its close proximity. Furthermore, the LMC has relatively little shear and has no strong spiral arms: the structure of the interstellar medium (ISM) is less affected by density wave shocks.

I. Puerari (✉)
INAOE, Santa María Tonantzintla, Mexico, School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa
e-mail: puerari@inaoep.mx

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_10,

The study of the morphology of galaxies facilitates our understanding of the physical processes of formation and of secular evolution which forge the distribution of stars, gas, dust, and the like.

Particularly interesting is the study of the spatial structure of the ISM. Atomic hydrogen HI gas is a sensitive tracer of the ISM. Due to the fact that radio observations of nearby galaxies can attain a good spatial and velocity resolution, HI gas distributions may be analyzed in terms of power spectral analysis [3–6, 8, 10, 15].

Another tracer of the structure of the ISM is cold dust, in the form of carbonaceous and silicate grains, at temperatures of 10–20 K. These grains are the dominant emitters in the FIR. Li and Draine [11] have extensively modeled COBE/FIR observations of the diffuse ISM in the Milky Way. Li and Draine [12] focus their attention on the SMC and show that carbonaceous and silicate grains of different sizes dominate the spectra at wavelengths in the range 50–400 μm. Large ($a > 250$ Å) grains become more important for wavelengths larger than 50 Å, while small carbonaceous grains prevail at shorter wavelengths, where one observes PAH's lines.

In this study, we present a power spectral study conducted on images of the LMC secured with the Spitzer Space Telescope and MIPS. We analyze images at 24, 70, and 160 μm. The longest of these wavelengths traces the dominant component of interstellar matter: cold dust grains. At 70 μm the emission arises from both warm and cold dust grains, while at 24 μm the emission is from warm dust grains and PAH's. Cold dust grains are responsible for the extinction in optical images, the scattering of starlight, as well as polarization. It is reasonable to state that in studying emission longward of 60 μm, one is investigating the structure of the optical mask itself.

2 Data

The coverage in the MIPS images of the LMC is $\sim 8 \times 8°$, with a total integration time of 217 h. The point source sensitivity estimates for the SAGE survey improve previous surveys of the LMC (such as those with the Infrared Astronomy Satellite IRAS) by three orders of magnitude [14].

At 24 and 70 μm, the images sizes are 8192×8192 pixels, at a scale of $4.98''$ and $4.8''$ per pixel, respectively. At 160 μm, the image size is 2048×2048 pixels, at a scale of $15.6''$ per pixel. As far as deprojecting images of the LMC are concerned, there is no unique "center" – the dynamical center of the HI is offset by almost a full degree from the photometric center of the LMC bar [19]. We choose to conduct our deprojections about the dynamical center of the LMC, which has right ascension $\alpha = 5^h27.6^m$ and declination $\delta = -69°52'$ (J2000.0), following Section 7 in van der Marel et al. [18]. In order to test the robustness of our analysis, we also deprojected the MIPS images about the bar center as opposed to the dynamical center of the LMC; the resulting power spectra are only marginally affected.

3 Analysis

A power spectral calculation is straightforward. In two dimensions, the prescription is as follows:

$$\mathcal{F}(k_x, k_y) = \int_x \int_y I(x, y) e^{-i(k_x x + k_y y)} dxdy \tag{1}$$

where $I(x, y)$ is the intensity of each pixel, and k_x and k_y are the conjugate variables to x and y, respectively. The full 2D power spectra are given by

$$Power(k_x, k_y) = (Re[\mathcal{F}])^2 + (Im[\mathcal{F}])^2 \tag{2}$$

A 1D PS $Power(k)$ can be calculated using $k^2 = k_x^2 + k_y^2$. Mathematically, this corresponds to take circles in the Fourier space $\mathcal{F}$.

4 Results

In Fig. 1, the left-hand panel shows emission from LMC dust at 160 μm. At top right is seen the Fourier space $\mathcal{F}$ while the power spectrum itself $Power(k)$ is shown at the lower right of Fig. 1.

The 2D power spectrum of the LMC at 160 μm demonstrates two distinct power laws (see [1] for further details). The steeper least-squares slope at higher spatial frequencies has a value of −3.08, while the shallower least-squares slope is −2.16. A distinct "knee" or "break" in the power spectrum is seen at approximately 200 parsecs. Figure 2 shows a very similar behavior – this time for the 70 μm image. At all three MIPS wavelengths, the knee is found to lie in the interval spanning the $100 - 200$ pc. The least-square slopes at the three different MIPS wavelengths are similar, but with a slight progression, as follows: at high-spatial frequency, the values are −3.08, −2.97, and −2.60 at 160 μm, 70, and 24 μm, respectively, while at low-spatial frequency, they are computed to be −2.16, −1.84, and −0.80 in these three passbands. The implication may mean that cool dust is more diffuse or dispersed than hot dust.

The spatial resolution of the MIPS images allows one to probe structures of dust emission at both small and large physical scales. At small scales (smaller than the disk scale height), the behavior of the power spectra is a 3D one, while for large scales, the PS is more akin to that in 2D. The LMC disk is relatively thin, with a height/size ratio of $\sim 1/15 - 1/20$; nevertheless, motions do not occur in strictly 2D, as in an infinitely thin 2D sheet.

We have no information on the actual velocities seen in the LMC images, but we believe that two different physical process are acting in the distribution of the ISM at large and small scales. Conceivably, more localized energy sources (from stars,

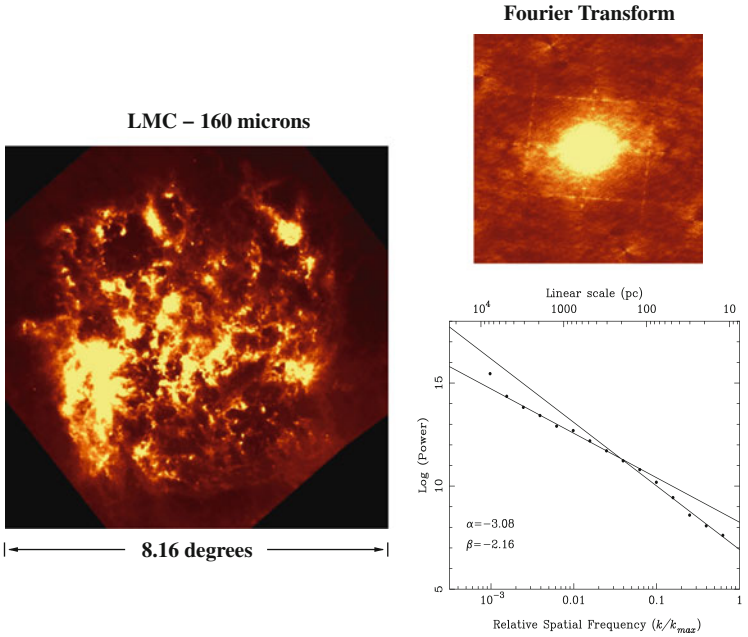


Fig. 1 *Left Panel*: The Large Magellanic Cloud imaged at 160 μm, as part of the SAGE survey. North is up, and East to the left. Emission from cold dust grains is seen in striking contrast. The bright emission cloud in the lower left of the MIPS image corresponds to the giant molecular and atomic clouds south of the 30 Doradus star-forming region. *Upper Right Panel*: The two-dimensional Fourier transform of the deprojected ("face-on") MIPS image at 160 μm. The rectangle betrays the presence of low-level striping in the SAGE MIPS image. *Lower Right Panel*: Two-dimensional power spectrum of the Large Magellanic Cloud. The power spectrum was divided into 15 intervals of relative spatial frequency, and each dot represents the mean of the power spectrum in that interval. Least-squares regression lines are used to determine the slopes of the two distinct power laws. The power laws are $P(k) \propto k^{\alpha}$ and $P(k) \propto k^{\beta}$ for high- and low-spatial frequencies, respectively. The point of intersection of the two regression lines at ~ 200 pc demarcates the "knee," interpreted as the line-of-sight depth of the cold dust disk of the Large Magellanic Cloud

OB associations, and SN explosions) should drive 3D motions and contribute to the small-scale sector of the power spectrum. In contrast, density-wave and bar-driven streaming motions are often 5–10 times faster than the perpendicular motions which produce the disk thickness and which are thus responsible for the 2D behavior of the ISM distribution.

5 Simulations

We have performed high spatial and mass resolution hydrodynamic simulations (up to 0.8 pc and 5×10^3 $M_{\odot}$), to study the properties of the ISM substructure and turbulence of LMC models. We use an AMR (Adaptive Mesh Refinement) software

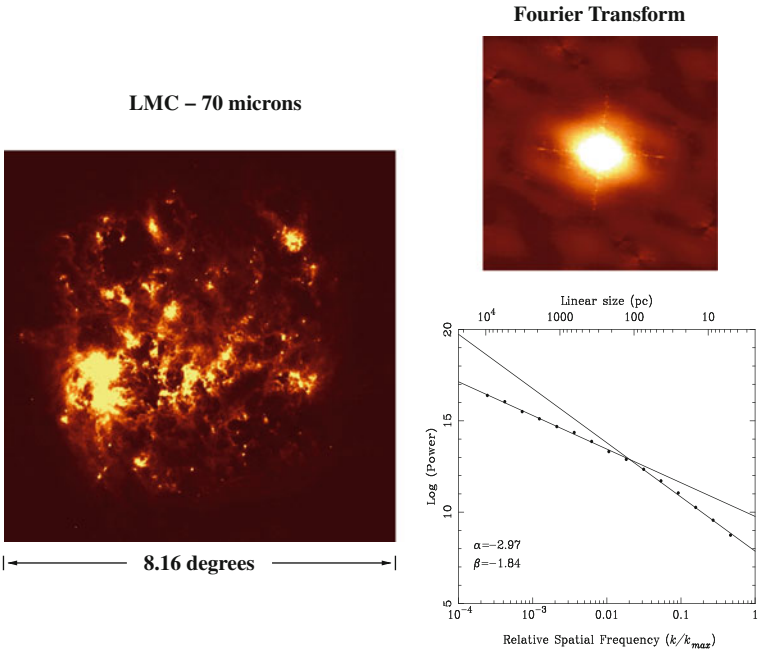


Fig. 2 Two-dimensional power spectral analysis of the Large Magellanic Cloud at 70 μm. The original 70 μm image appears in the *left panel*. Details as in Fig. 1. Shown in the *lower right-hand panel* are two power laws, and an associated "knee" which occurs between 100 and 200 pc. Particularly striking is the smallest scale sampled in the 70 μm power spectrum of only 2.32 pc, at the Nyquist limit. The power spectrum spans almost four orders of magnitude. The value of the steeper (more negative) slope of −2.97 in the power spectrum is close to the −3 slope representative of 2D incompressible Kolmogorov turbulence

written by Romain Teyssier (see [16]). A complete discussion appears in Bournaud et al. [2]; in this chapter, we highlight some details. We initialize an exponential stellar disk containing 3×10^9 $M_\odot$, a scalelength of 1.5 kpc, and a truncation radius of 3 kpc. A non-rotating bulge (3×10^8 $M_\odot$) and a halo (5×10^9 $M_\odot$) are both included. The initial gas disk is exponential, with a scalelength of 3.0 kpc and a scaleheight of 50 pc (truncated at 500 pc). The total gas mass is 6×10^8 $M_\odot$. Initially, the gas is purely rotating, with no macroscopic velocity dispersions. With time, gravity generates turbulence at several scales. For purposes of comparison, we generated two models: with and without the feedback from star formation.

In Fig. 3 we show the power spectra of the gas distribution for the "with star formation feedback" run at times T=254 Myr and at three subsequent times, incremented by 7 Myr. The two slopes are clearly seen, and their values are −3.12 for large scales and −1.89 for small scales. These values agree with those calculated observationally, using the MIPS FIR images above. We have also computed power spectra of the velocities (in-plane V_R and V_θ, and vertical V_Z) (seen in Fig. 4). At small scales, all of the PS of velocities have the same slope with a marginal

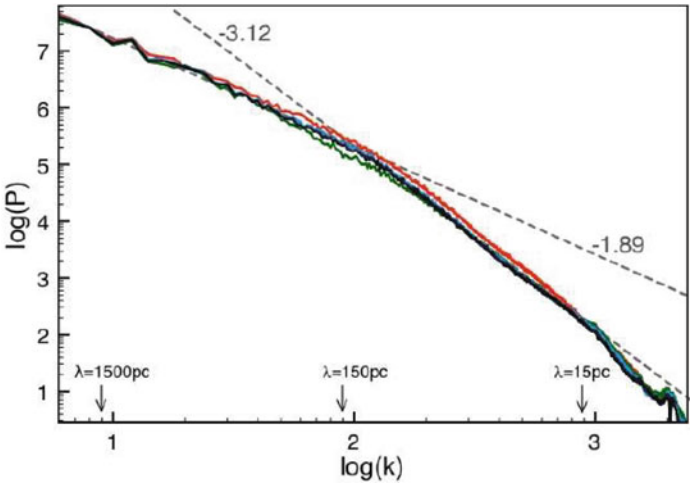


Fig. 3 Density power spectrum for four different times in the simulation with feedback. T = 254 Myr (*dark*), 261 Myr (*green*), 268 Myr (*blue*) and 275 Myr (*red*)

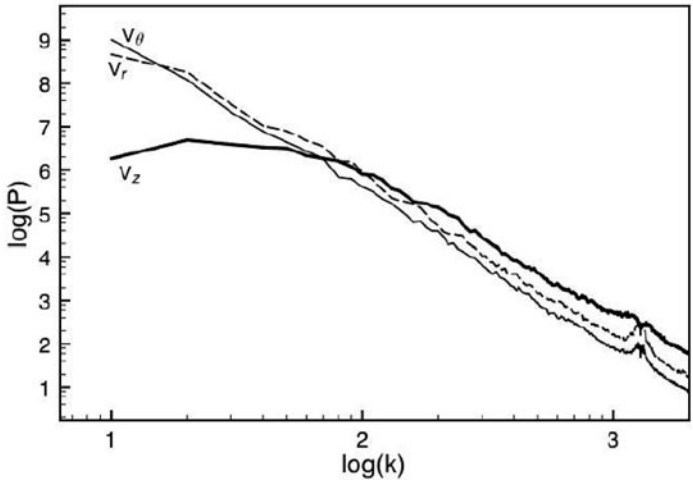


Fig. 4 Velocity power spectrum for the three separate components V_R, V_θ, and V_Z, at T = 254 Myr (LMC model with stellar feedback)

difference in power. At large scales, the motions are strongly non-isotropic, with the V_Z power spectra being flat, with less power, compared to in-plane velocities. The latter ones present the same slope at high and low frequencies. This result is significant, because it demonstrates that the vertical component of the velocity behaves very different for high- and low-spatial frequencies. At high frequencies, the turbulence is more 3D (almost the same power for V_R, V_θ, and V_Z). At low frequencies (large scales), however, the motions are 2D (much less power for V_Z than for the in-plane velocities). These 2D streams, generated primarily by disk

self-gravity – but possibly also by tidal forces from companion galaxies, as in Mastropietro et al. [13] and from intergalactic ram pressure, as discussed by Tonnesen and Bryan [17] and by Dutta et al. [7] – contribute to the low-frequency portion of the power spectrum.

6 Conclusions

For the first time, power spectra are presented for emission for dust grains in the LMC. This is facilitated by analyzing FIR images at wavelengths of 24, 70, and 160 μm from MIPS, on board the Spitzer Space Telescope. The power spectra reveal two distinct power laws in the high- and low-frequency domain, representing small and large physical scales, respectively. The "knee" or "break" in the power spectra occurs at $\sim 100 - 200$ pc. We interpret this break as the scaleheight of the dust disk of our closest magellanic neighbor, the LMC, following Elmegreen, et al. [8]. High-resolution AMR simulations of LMC models affirm that the same behavior follows in the power spectra generated from our simulated ISM: simulations confirm the existence of two different power laws, with the 'break' occurring at the scale height of the disk. Observations and simulations are in excellent accord. As far as velocity components are concerned, we find that for high frequencies (small scales), the motions are quite isotropic, while at low frequencies, in-plane motions are much more important than vertical ones. These results affirm that the power spectra of the ISM does indeed depend upon the geometry of the system being simulated. At large scales, perturbations create flows akin to those in two dimensions, while at small scales, simulations betray a behavior which is more three dimensional in nature.

Acknowledgments I.P. extends his thanks to the Mexican Foundation Conacyt. He also expresses profound gratitude to the School of Computational and Applied Mathematics at the University of the Witwatersrand, Johannesburg, for their hospitality during his many visits to the University, spanning more than 12 years.

References

1. Block, D.L., Puerari, I., Elmegreen, B.G., Bournaud, F. 2010, ApJ, 718, L1
2. Bournaud, F., Elmegreen, B.G., Teyssier, R., Block, D.L., Puerari, I. 2010, MNRAS, arXiv1007.2566
3. Crovisier, J., Dickey, J.M. 1983, A&A, 122, 282
4. Dutta, P., Begum, A., Bharadwaj, S., Chengalur, J.N. 2008, MNRAS, 384, L34
5. Dutta, P., Begum, A., Bharadwaj, S., Chengalur, J.N. 2009a, MNRAS, 397, L60
6. Dutta, P., Begum, A., Bharadwaj, S., Chengalur, J.N. 2009b, MNRAS, 398, 887
7. Dutta, P., Begum, A., Bharadwaj, S., Chengalur, J.N. 2010, MNRAS, 405, L102
8. Elmegreen, B.G., Kim, S., Staveley-Smith, L. 2001, ApJ, 548, 749
9. Freedman, W.L. et al. 2001, ApJ, 553, 47
10. Green, D.A. 1993, MNRAS, 262, 327
11. Li, A., Draine, B.T. 2001, ApJ, 554, 778
12. Li, A., Draine, B.T. 2002, ApJ, 576, 762

13. Mastropietro, C., Burkert, A., Moore, B. 2009, MNRAS, 399, 2004
14. Meixner, M. et al. 2006, AJ, 132, 2268
15. Stanimirovic, S. et al. 1999, MNRAS, 302, 417
16. Teyssier, R. 2002, A&A, 385, 337
17. Tonnesen, S., Bryan, G.L. 2009, ApJ, 694, 789
18. van der Marel, R.P., Alves, D.R., Hardy, E., Suntzeff, N.B. 2002, AJ, 124, 2639
19. Westerlund, B.E. 1997, The Magellanic Clouds, Cambridge Astrophysics Series, Cambridge University Press, Cambridge

Light Cores Behind Dark Masks

Ruben J. Díaz, Damián Mast, Germán Gimeno, Horacio Dottori, Irapuan Rodrigues, María Paz Agüero, and Peter Pessev

Abstract The dusty circumnuclear environments of many spiral galaxies hide important clues for understanding the tight correlation between the supermassive black hole and the host galaxy bulge mass. Two of the nearest spiral galaxies with nuclear starburst host off-centered compact cores associated with the peak of the infrared emission. These compact objects are lightweight, up to a few million solar masses. Their locations imply that they are triggering the present star formation within the circumnuclear disk. Numerical modeling implies that the giant star clusters formed in the process evaporate in a few system revolutions (due to the strong asymmetric gravitational potential). Therefore the off-center compact object scenario provides a model in which the dynamically cold gaseous component is rapidly converted into a dynamically hot stellar component. Meanwhile only a small fraction of the gas contributes to the growth of the off-centered compact object.

1 Introduction

One century after the discovery of strong emission lines in the center of NGC 1068, galactic nuclei are still mysterious places. Several fundamental questions on the Super Massive Black Holes (SMBH) formation and its relation with the structure of the host galaxies remain unsolved. In particular, the tight correlation between the SMBH mass and the velocity dispersion/mass of the galactic bulge within which it resides [10, 11] is compelling evidence for a close connection between the formation of the black hole and the host galaxy [15]. It is yet not clear how the many different mechanisms for angular momentum removal that may affect the evolution of a spiral galaxy (interactions, minor mergers, bars, etc.) will end up with a bulge harboring a compact dark core of about one thousandth of its mass. A key to understand the relationship is probably hidden in the domain of the smallest spatial scales and

R.J. Díaz (✉)
Gemini Observatory, AURA, Colina El Pino, La Serena, Chile
e-mail: rdiaz@gemini.edu

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_11,

SMBH masses, behind the dusty nuclear environment of spiral galaxies. Moreover, the presence of nuclear starbursts in some of the nearest spiral galaxies is indicative of nuclear fueling processes building up the central regions. This scenario is usually masked by the byproducts of the star formation (mainly obscuring material) and kinematic feedback. The observational approach includes the use of near-infrared spectroscopic data cubes at moderate to high spectral and high-spatial resolution. A theoretical approach involves the use of N-body models with smoothed particle hydrodynamics.

We report that two of the nearest spiral galaxies with nuclear starburst in the southern hemisphere, NGC 5236 and NGC 253, host off-center compact cores associated to the peak of infrared emission at the central region. These compact objects are up to a few million solar masses, and their locations imply that they are triggering the present star formation within the circumnuclear disk. We have already thoroughly studied the case of NGC 5236, for which numerical modeling implies that the formed giant star clusters evaporate in a few system revolutions due to the strong asymmetric gravitational potential. After presenting the last observations for both galaxies, we discuss the implications of the off-center compact object scenario for the SMBH−Bulge co-evolution at the nuclei of spirals.

2 Observations of NGC 5236

NGC 5236 (Fig. 1) is one of the two-dominant galaxies at the Hydra-Centaurus group. This object has more than 20 satellites, a very large disc of neutral hydrogen and shows evidence of tidal perturbations. Two large-scale stellar streams were mentioned at this meeting [3]. The starburst nucleus harbors about 20 clusters as massive as 30 Dor and shows asymmetric appearance [23].

We used the Cambridge Infrared Panoramic Survey Spectrograph [19], during its visit to the Gemini South 8.1m telescope. The observations were taken in March 2003 with an Integral Field Unit (IFU) sampling of 490 spaxels of 0.36″ (7.2 pc) in an elliptical arrangement with a size of 13″ × 5″. The IFU was centered in a point midway between the optical nucleus position and the possible position of the hidden nucleus previously determined from our optical 2D kinematics [18]. Three fields were observed in order to construct a mosaic. The total set of almost 1,500 spectra covers the spectral range $1.2 - 1.4\,\mu$m, at spectral resolution 3,200. The achieved image quality was FWHM$\approx$ 0.5″ at the continuum maps. The data reduction techniques used were described in a previous publication [6]. We fitted gaussians to the Paβ and [Fe II] $1.26\,\mu$m emissions of each spectrum, as well as to the most prominent sky emission lines near Paβ for using them as wavelength and profiles references. S/N ratio higher than 10 was achieved over most of the field. For the wavelength calibration we used the sky emission lines. The mean dispersion between the values obtained (the radial velocity determined from each reference pair of sky lines) never surpassed $2\,\mathrm{km\,s^{-1}}$, what we consider the mean error of the linearization. In order to determine the total error of the radial velocity, we sum in quadrature the above error and the one obtained from the empirical relation for the gaussian fit [16]. The average radial velocity uncertainty is about $6\,\mathrm{km\,s^{-1}}$.

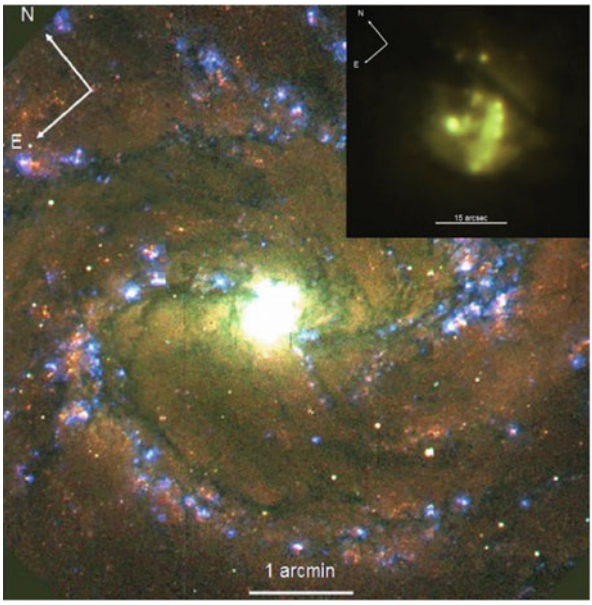


Fig. 1 False color image composed from continuum subtracted Hα, [OIII]λ5007 and [SII]λ6730 narrow band images obtained with GMOS-S. The inset shows a z band image of the nucleus where the giant star formation arc is evident

In Fig. 2 we present a new large mosaic velocity field, the [FeII]/Paβ ratio map, W(Paβ) and the corresponding age map. The equivalent widths of hydrogen recombination lines are powerful age indicators [7]. We used the W(Paβ) map making a polynomial fit to the age output of the corresponding SB99 model [17]. Using instantaneous star formation, Z=2Z$_\odot$, Salpeter IMF with limits 1 and 100 M$_\odot$ as input parameters, we obtained an equation that allowed us to calibrate the W(Paβ) map into age. The results regarding the star formation history will be discussed elsewhere. We confirm that the dark nucleus is located at the youngest end of the star forming arc and find no evidence of shocks associated with the position of the hidden nucleus (which could eventually mimic a rotation pattern). The spider kinematic diagram corresponding to the dark compact core is indeed located away from the bulge center. We have shown [6] that this kinematic center is located exactly at the location of the peak mid-infrared emission and also the youngest end of the starburst arc. It was shown that the age gradient of the giant star clusters studied by HST multiband optical and NIR imaging [12] is similar to the dynamical crossing time calculated for the dark core. The maximum unresolved mass for a compact object that can be fitted to the observed velocity field is of 3×10^6 M$_\odot$. We processed Chandra archival data of the nucleus and found that most of the diffuse X-rays in the spectral range 0.5–8 kev come from a cone pointing to the hidden nucleus position, indicating that a strong ionizing source is present under very high extinctions in the line of sight (Fig. 2).

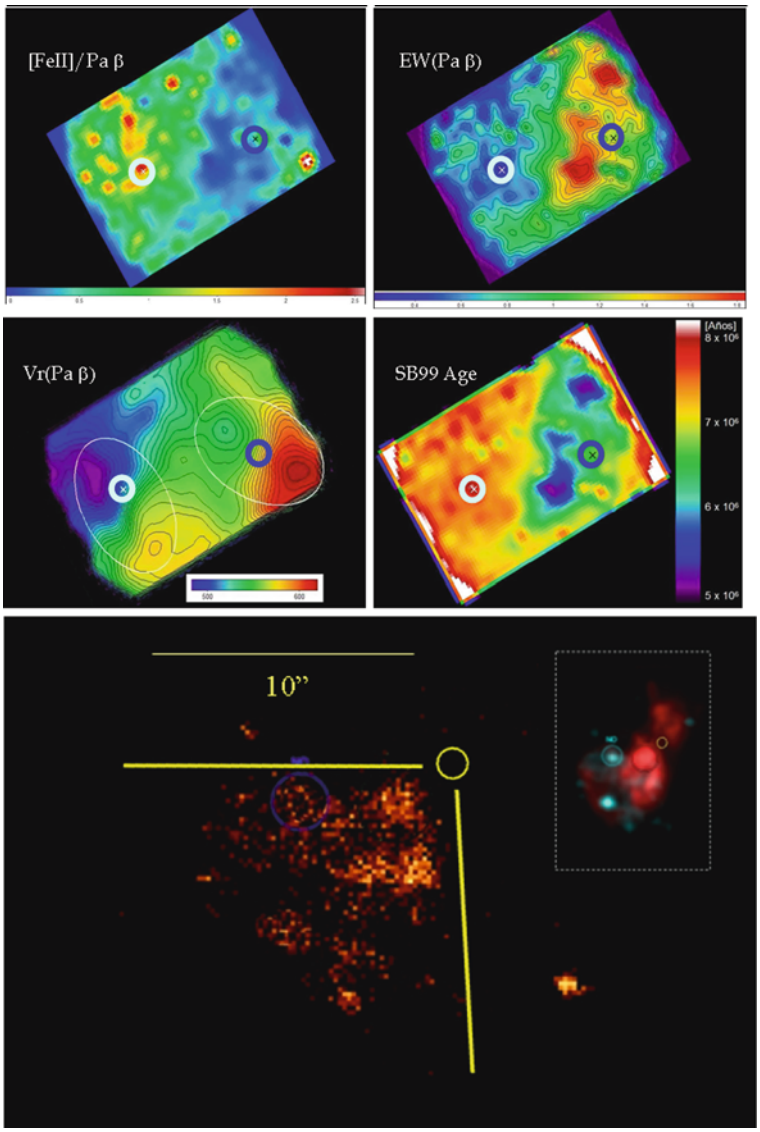


Fig. 2 *Top*: Some relevant maps ($13'' \times 10''$) derived from merged data cube, which is centered on the bulge symmetry center. The maps depict the distribution of [FeII]/Paβ ratio, equivalent width W(Pβ), smoothed radial velocity field, age distribution map from a polynomial fit to the SB99 simulation. The optical nucleus (*white circle*) and dark nucleus (*blue circle*) are indicated. *Bottom*: Distribution of the integrated X-ray emission in the energy range 0.5–8 kev, using Chandra archival data. After removal of the compact sources, a scattering cone is apparent, with vertex at the dark nucleus position (*yellow circle*). The optical nucleus position is marked with a *faint blue circle*. The *inset* shows a composition of optical narrow band emission and X-ray emission, unveiling a biconical structure with the same vertex

3 Modeling of NGC 5236 Nucleus

In order to understand the dynamical evolution of the double nucleus we performed N-body simulations of a system composed of masses corresponding to the bulge and global potential, the hidden nucleus, the optical nucleus, and four condensations representing the largest clusters in the star formation arc. Hernquist's models [13] were used with stellar and gaseous components constrained by the bodies rotation curves. The clusters were represented by Plummer's models, with 1,000 particles each, with position and masses corresponding to previous references (regions 2–5 in Fig. 8 of [8]). The model details, as masses and orbital parameters, are discussed elsewhere [22]. A large set of simulations were run in order to refine the models and orbital parameters used in the final simulation. The simulations were performed with Gadget2 [24]. A total 91,552 particles were used, and the simulations were run for 300 Myr, starting at the present time configuration. The simulation (Fig. 3) shows that the optical nucleus and the dark nucleus would form a single massive core in as short timescale as 16 Myr. All Plummer cluster cores fall into the nucleus in 130 Myr. Considering the range of uncertainties in the orbit determination, we can state that this massive core would finally settle as the new nucleus of NGC 5236 in a few tens of Myr, implying a net growth of the galactic core. The simulations also show that the tidal stripping of the condensations boosts the velocity field of the external shell to escape velocity, which, if mistaken for the dispersion velocity

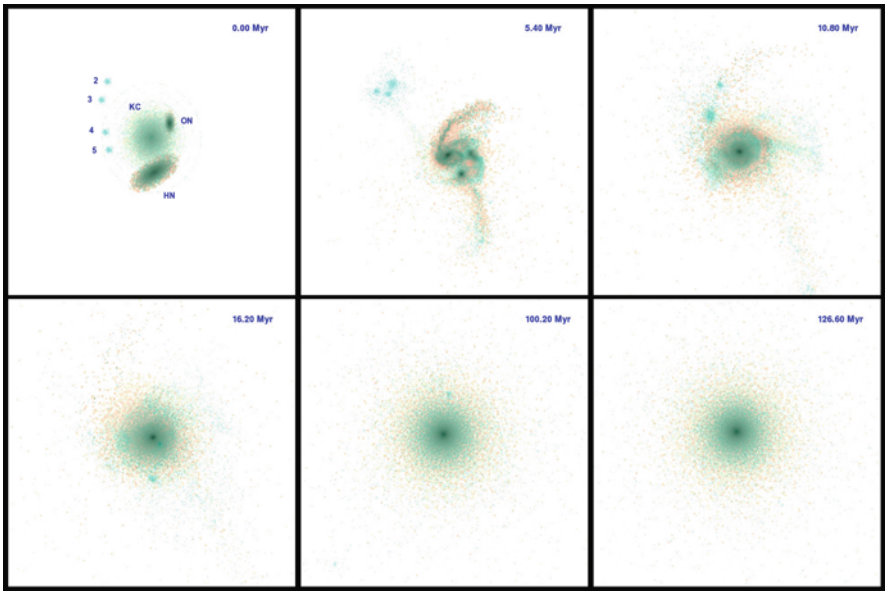


Fig. 3 Evolution of the NGC 5236 Gadget2 model, with ∼ 1,00,000 particles including gaseous components. The main masses merge in less than 20 Myr, and the last Plummer sphere representing a giant stellar cluster coalesces with the central body in 130 Myr

of systems in equilibrium, might lead to overestimation of the cluster masses. This seems to be case of disagreement between the photometric and kinematic masses derived by previous stellar population studies of the optical nucleus [25]. The circumnuclear HII regions arc is far from being a stable system. In fact, it will spread out in an orbital time and be swallowed by the refurbished nucleus in a period slightly larger than the merging of the nuclei. Therefore the claimed existence of two ILRs and a regular system of X2 bar orbits [14] is difficult to sustain in this scenario. It is worth to remark that the dynamically cold gaseous component which yields to the giant star clusters is rapidly converted into a dynamically hot stellar component, joining the bulge mass that could be measured, e.g., through the central stellar velocity dispersion, while a small fraction of it may contribute to the growth of the compact object as it reaches the sphere of influence of the resulting compact object.

4 Observations of NGC 253

NGC 253 (Fig. 4) is the brightest object in the Sculptor group and, as the previous case, the nucleus appears to be heavily obscured by dust and its position does not coincide with the peak of the near- and mid-IR emission [21]. Nevertheless, apparently there are at least four compact star clusters in the nuclear region – apart from the optical nucleus itself [26]. One of them (the so-called bright blob) coincides with the peak IR emission [4, 20] and a bipolar outflow cone is observed in X-ray emission (see [9] for references and a study of the starburst). It has already been shown [9] that there is no evidence of a compact source associated to the symmetry center of the circumnuclear disk, at optical, near-infrared, and radio wavelengths. The infrared spectral energy distribution of the clusters measured by these authors

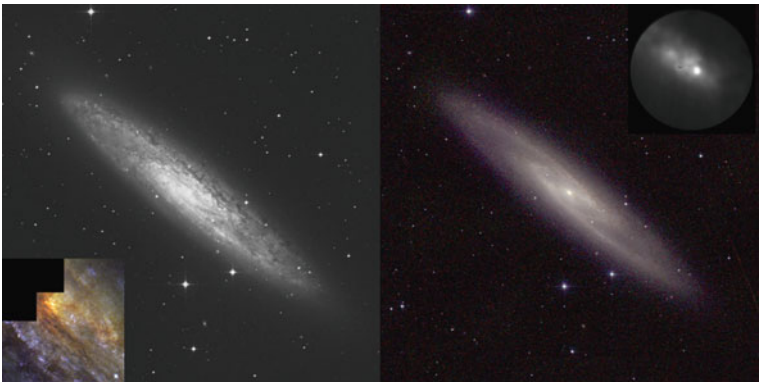


Fig. 4 *Left*: Red Digitized Sky Survey image of NGC 253, $26' \times 26'$ field. The lower inset shows the color composite image (F814W, F555W, F336W filters) of the nucleus from the *HST* archive. Right: Two Micron All Sky Survey color composite (J, H, K_S bands). The *upper inset* shows the Phoenix acquisition image at 2.2 μm ($10''$ field of view)

implies up to 10^5 young stellar objects ($<M>\sim 5\ M_\odot$) in the largest knots, which in turn means that near-infrared observations at spectral resolutions larger than 10,000 are needed to pinpoint the gravitational potential well center and define a perturbation pattern geometry from the kinematical data.

We are scanning the circumnuclear Brγ emission with the Phoenix high-resolution spectrograph, which has a slit length of $14''$ and a sampling of $1.5\,\mathrm{km\,s^{-1}}$ pixel^{-1}. We are also scanning the MgI absorption line at $2.108\,\mu\mathrm{m}$, aiming to unveil the nature of the true nucleus and the compact clusters and measure the velocity dispersions. We are building a datacube from the Phoenix long slit scanning, which will cover a region $5'' \times 14''$, at resolution of about $0.4''$, and a spectral resolution of 25,000 (after binning) covering the kinematics of the HII emission in the K-band. This will achieve unprecedented level of detail of the central kinematics of a nearby nucleus from NIR ground-based observations. These data will allow to constrain the models of SMBH evolution inside a nuclear star forming disk in detail. Some other physical processes, like the SNe feedback would be better understood. Figure 5 shows two spectral frames. One of them goes through the true kinematical center (located $\sim 4''$ SW from the global geometry center) characterized by a larger local velocity gradient. The slope is uniform around the area of the bulge center, while it becomes steeper at the position of the Object 4, following the previous nomenclature [9]. The largest compact object that can be fitted in the rotation pattern is in the range 0.8–1.2 $\times\ 10^6\ M_\odot$, which is consistent with the values inferred from

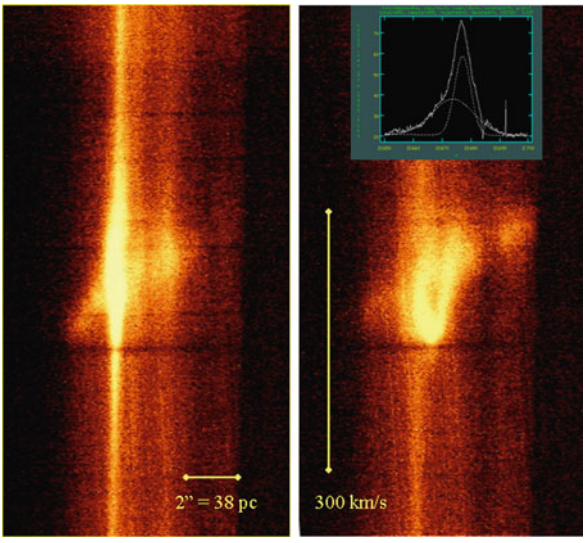


Fig. 5 Two preliminary Phoenix spectral frames at different positions on the nucleus of NGC 253, showing the Brγ emission at spectral resolution 50,000. *Left*: The largest local velocity gradient is located $4''$ SW of the bulge center, through the brightest mid-IR source. *Right*: A different slit position depicts a bubble with a size of 60 pc and expansion velocity of $100\,\mathrm{km\,s^{-1}}$. Note the richness of the data cube being elaborated

the large emission line broadening also detected at these position. As in the case of NGC 5236, the off-center kinematic core is located on the spot of the maximum IR emission, at the youngest side of the circumnuclear star formation region, and this dynamical configuration indicates that strong evolution is happening at the nucleus.

5 Concluding Remarks

We are witnessing dramatic dynamical changes in the central regions of the nearest spiral galaxies with starburst nucleus, with the largest local mass concentration not coincident with the global kinematic center in NGC 5236 and NGC 253. Considering the case of Milky Way and Andromeda galaxies in the local group, and other known cases in nearby spiral galaxies [1, 2, 5] this starts to seem a fairly common feature. In the case of a presence of a circumnuclear gaseous disk the off-centering seems to be triggering the star formation. Our simulations show that the dynamically cold gaseous component would be rapidly converted into a dynamically hot stellar component, while a small fraction of it may contribute to the growth of the compact object. We are planning to perform numerical modeling with several million particles in order to better understand of the fate of this kind of configuration. In case that the gaseous mass transformed into stars evaporated to the bulge would be about three orders of magnitude the mass that would reach the sphere of influence of the collapsed object, the off-center core mechanism could be one of the preferred scenarios yielding to the SMBH–Bulge mass relationship at some evolutionary stages of the spiral galaxies.

Acknowledgments We acknowledge the support of MINCYT (Argentina) and CAPES (Brazil), a grant by CONICET (PIP 5097), and the Gemini time allocations GS-2003A-Q-17, GS-2007A-DD-17, and GS-2009B-Q-39. The Gemini Observatory is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership, consisting of the NSF (USA), STFC (United Kingdom), NRC (Canada), ARC (Australia), MINCYT (Argentina), CNPq (Brazil), and CONICYT (Chile). CIRPASS was provided and supported by the Instrumentation Group at the Institute of Astronomy, Cambridge. We have also used Chandra and HST archival data.

References

1. Arribas, S., Mediavilla, E. 1994, Astrophys. J. 437, 149
2. Arribas, S., Mediavilla, E., Garcia-Lorenzo, B. 1996, Astrophys. J. 463, 509
3. Block, D., Freeman, K. 2008, Shrouds of the Night, Springer, New York, NY, p. 304
4. Das, M., Anantharamaiah, K., Yun, M. 2001, Astrophys. J. 549, 896
5. Díaz, R.J., Carranza, G., Dottori, H., Goldes, G. 1999, Astrophys. J. 512, 623
6. Díaz, R.J., Dottori, H., Agüero, M.P., Mediavilla, E., Rodrigues, I., Mast, D. 2006, Astrophys. J. 652, 1122
7. Dottori, H. 1981, Astrophys. Space Sci. 80, 267
8. Elmegreen, D., Chromey, F., Warren, A. 1981, Astronom. J. 116, 2834

9. Fernández-Ontiveros, J.A., Prieto, M.A., Acosta-Pulido, J.A. 2009, Mon. Not. R. Astron. Soc. 392, L16
10. Ferrarese, L., Merrit, D. 2000, Astrophys. J. Lett., 539, L9
11. Gebhardt, K., et al. 2000, Astrophys. J. Lett. 539, L13
12. Harris, J., Calzetti, D., Gallagher, J.S., III, Conselice, C.J., Smith, D.A. 2001, Astron. J. 122, 3046
13. Hernquist, L. 1993, Astrophys. J. Suppl. 86, 389
14. Houghton, R.C., Thatte, N. 2008, Mon. Not. R. Astron. Soc. 385, 1110
15. Kauffmann, G., Haehnelt, M. 2000, Mon. Not. R. Astron. Soc. 311, 576
16. Keel, B. 1996, Astrophys. J. Suppl. 106, 27
17. Leitherer, C. 1999, Astrophys. J. Suppl. 123, 3
18. Mast, D., Díaz, R.J., Agüero, M.P. 2006, Astron. J. 131, 139
19. Parry, I. 2000, SPIE 4008, 1193
20. Prada, F., Gutierrez, C., McKeith, C. 1998, Astrophys. J. 495, 765
21. Sams, B., Genzel, A., Tacconi-Garman, L., Hofmann, R. 1994, Astrophys. J. 430, L33
22. Rodrigues, I., Dottori, H., Díaz, R.J., Agüero, M.P., Mast, D. 2009, Astron. J. 137, 4083
23. Sérsic, J.L., Pastoriza, M. 1967, Publ. Astron. Soc. Pacific 79, 152
24. Springel, V. 2005, Mon. Not. R. Astron. Soc. 364, 1105
25. Thatte, N., Tecza, M., Genzel, R. 2000, Astron. Astrophys. 364, L47
26. Watson, A., Gallagher, J., Holtzman., J, et al. 1996, Astron. J. 112, 534

Globalization, Open Access Publishing, and the Disappearance of Print: Threat or Opportunity?

J.J. Blom

Abstract Academic publishing is changing rapidly as a result of mutually dependent developments. Electronic publishing was born through the emergence of the Internet. The Internet has accelerated globalization, which in turn has enhanced the consolidation of commercial publishing houses and the collaboration between libraries. Old business models are being replaced by new business models, such as open access publishing. The familiar print format is disappearing for journals and perhaps for books. This chapter presents an overview of these developments and addresses potential threats and opportunities to the parties involved in publishing.

1 Introduction

Scientific, Technical and Medical (STM) publishing is going through rapid changes. Electronic publishing took off around 1995 and less than two decades later publishers are considering to stop printing journals, while this was the preferred method of communicating scientific discoveries for more than three centuries. Indeed, one of the earliest science journals still published, The Philosophical Transactions of the Royal Society, was started in 1665. The first electronic journal established by Richard H. Zander, Flora Online, started in 1987 (details in Davis and Schmidt [3]).

The main driver of change in academic publishing is the ongoing development of Internet functionalities, which now goes hand-in-hand with commercial globalization. The Internet was first a tool to connect academic institutions for the rapid exchange of information, but after the development of user-friendly browser software (Mosaic in 1993) and efficient search engines (Google in 1997) it was fully embraced by the private sector as a means to sell products and services. The electronic commerce, or "e-commerce", became ubiquitous and successful.

J.J. Blom (✉)
Physics and Astronomy, Springer, New York, USA
e-mail: harry.blom@springer.com

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_12,

Commercial academic publishing took off in 1960s when many new journals were founded or existing journals expanded. This coincided with a rapid expansion of scientific output, itself a result of sustained economic growth in countries with a strong tradition in science and education. For decades, the business model of selling subscriptions to libraries worked without problems, but toward the end of the 1980s the growth of subscription prices no longer matched the growth of library budgets and serials cancelations started to erode the library collections. Diminishing circulation numbers against ever increasing prices announced the end of the print subscription business model. It is right in the middle of this often-quoted "serials crisis" that electronic publishing took off and came to the rescue.

Electronic subscriptions can be shared by many subscribers simultaneously, which means that available budgets for partial subscriptions can be combined to buy access to a package of journals. Combining subscriptions is done by a group of libraries in a so-called consortium. Large groups of libraries gain negotiation power, but also the stability offered by multi-year deals in combination with a smaller number of important accounts to do business with, enhances the overall efficiency of the sales efforts, which lowers transaction costs.

2 Globalization

Science, and certainly astronomy, has been an international effort for centuries. The greatest minds, and their ideas, traveled across boundaries and cultures since the birth of modern science around the sixteenth Century. To give just one of many possible examples, the book *Cosmotheoros* by Christiaan Huygens was published in Latin in 1698. Less than 20 years later, this book was also available in English, Dutch, French, German, and Russian and Huygens was a member of scientific academies in several countries. But for a long time, most of the scientific research was funded and carried out in Europe (with western Russia) and North-America. Indeed, in 1980 still about 85% of all journal articles documented in the Web of Science database are published by researchers working on these continents [1]. Only during the last decade, we see a fast rise of the scientific output from Asia and to a lesser extent Latin-America and the middle-East. Based on these numbers one could argue that the twenty first Century marks the true start of science as a global activity. Perhaps not surprisingly, this coincides with a more global spread in Gross Domestic Product (GDP) and budgets for research and education (see [5]). The global scientific output of journal articles grows each year by about 2.5% and the Web of Science database grew from about 0.5 million articles in 1980 to 1 million articles in 2009. Everybody with an interest in science, whether a researcher, librarian, or publisher, will likely see these trends almost as opportunities.

US economists define a fifth wave of mergers and acquisitions (M&As) that started in the early 1990s and still continues [4]. This is the wave characterized by cross-border M&As driven by globalization. The publishing industry has largely followed and contributed to the overall trend of mergers and acquisitions. Besides

common motives such as better economies of scale, benefits of larger market shares and synergies, it seemed natural for academic publishers to have a network of offices close to international research centers of excellence. So when Kluwer Academic Publishers and Springer-Verlag merged in 2004, the legacy of many imprints from different countries came under one roof. Among others these imprints are D. Reidel, Martinus Nijhoff, Chapman & Hall, Plenum, Birkhäuser, Vieweg, Bohn Stafleu van Loghum, Apress, Steinkopff, Valdemars, Urban & Vogel. The four largest STM publishers Elsevier, Springer, Wiley–Blackwell, and Taylor & Francis are each the result of many M&As and have become multi-national companies, each publishing thousands of books and journals per year.

These large STM publishers consider their increased global presence and customer base as a great opportunity and certainly their costs of production per article or book have gone down steeply. In fact, the wide global infrastructure of outsourced activities initiated by commercial publishers is now adopted or followed by smaller society and university presses, either directly or indirectly by having their journals and books distributed by these large global publishers. But critical voices from academia and libraries keep warning against monopolies and the power of these companies to dictate prices. For example, in response to the news of the Kluwer–Springer merger, the Scholarly Publishing and Academic Resources Coalition (SPARC) sent a letter in the summer of 2003 to the Commission of the European Communities (EC). They were worried about a reduction in access to scientific information and were perhaps hoping the EC would apply anti-trust legislation to prevent the merger or at least impose measures against its potential negative effects [9]. The EC did not act on these worries immediately, but such sentiments are one of the drivers of the Open Access movement discussed in Section 3.

The trend to move to larger scales didn't skip academic institutional libraries. As was mentioned in the introduction, a group of libraries, or consortium, can save costs by sharing subscription licenses and other resources. Indeed, with the development of electronic publishing in the early 1990s, nationwide (purchasing) library consortia were founded in many countries, such as the UK, Israel, China, Italy, Spain, Brazil, and Australia [8]. In the USA, library collaborations have a longer tradition [2] but also in this country large statewide consortia in Georgia, Louisiana, Ohio, Texas, and Virginia were started in the late 1980s to mid-1990s as well [7]. The pan-national consortia appear in the late 1990s, such as the International Coalition of Library Consortia (ICOLC) in 1997. Relevant for astronomy is the Physics-Astronomy-Mathematics (PAM) Division of the (international) Special Library Association. The daily e-mailed opinions and experiences, and the quarterly bulletin of this group, provide informative peeks in the kitchen of everything that is happening in PAM publishing.

The result of this concentration is that large publishers now do business with roughly 400 large consortia, instead of close to 35,000 individual academic-level libraries that exist in the world. Subscriptions (and more recently eBooks) are mainly sold in large packages, often encompassing the entire portfolio of a publisher. Seizing the opportunity of a smaller pool of decision-makers, the

definition and implementation of industry standards is accelerating. For example, most academic publishers support the not-for-profit organization CrossRef (www.crossref.org), which registers Digital Object Identifiers (DOIs), now widely used as references. DOIs are set up as (eternal) URLs to the Internet source of journal articles, book chapters, figures, or any other element that one would like to identify and locate separately. The most recent development of this organization is CrossCheck. This tool compares newly submitted articles with databases of published literature, other submitted articles, and the Internet. The level of text duplication detected subsequently provides an early indication of plagiarism, so CrossCheck should ultimately be capable of preventing the publication of serious forms of plagiarism in peer-reviewed journals.

3 Open Access Publishing

As soon as the Internet reached critical mass for academic institutions, i.e., a sufficient number of researchers got connected, people started to share electronic versions of scientific articles. Physicists and astronomers were leading the pack. As of 1991, they started to deposit their mostly not-yet peer-reviewed articles in Paul Ginsparg's now famous pre-print repository called arXiv, originally hosted at Los Alamos National Laboratory, now at Cornell University. In fact, this started before destination journals distributed official electronic editions and most users were unaware of legal implications (i.e., breach of copyright). The host institution of arXiv (so far) absorbed the costs of moderating the content and maintaining the server, allowing immediate public access to all new discoveries drafted in the deposited articles. The wide use of the arXiv means that the astronomy, physics, mathematics, computer science, and quantitative biology communities have de facto created an "open access" situation in relative harmony with subscribed-for peer-reviewed destination journals.

In the biomedical research community, the push for public access had a slower start, but developed more aggressively. Limitations to access of discoveries in medicine was perhaps felt to be more unethical. Harold Varmus, Nobel laureate and director of the U.S. National Institutes of Health (NIH) from 1993 to 1999, has been an important driver behind what became PubMed Central in 2000, a public database of full-text articles. Since April 2008, the NIH requires that all findings resulting from the research they have sponsored should appear in the public domain within 1 year after the official date of publication. This 12 months embargo mostly preserves the business model of peer-reviewed "toll access" journals.

But the central idea that tax payer sponsored research results should become freely available to the tax payer is gaining further momentum. Indeed, science funding agencies in several countries now require, or are considering to require, publishing exclusively with open access (see [6]). Many publishers started open or hybrid access journals in response. In this Open Access model, the author keeps the copyrights and pays an article fee to cover all costs of publishing, e.g., peer review,

journal and database management, type setting, marketing, rights management, Web development, accounting, uninterrupted perpetual hosting on a Web site, and a margin for the publisher. Since such costs still need to be covered, the science funding agencies and other sponsors are indeed factoring in publishing expenses when they are defining the value of their research grants.

Is this new model a threat or opportunity? From the point of view of the publisher it is both. In astronomy, the average open access fee is roughly 1,100 Euro/ 1,500 USD. One can debate how high this fee needs to be, but at this level it is possible to maintain a healthy journal if a sufficient stream of good quality articles is submitted. No more dependence on shrinking library budgets, no need for a sales force to go out and sell. But almost all of the well-known astronomy journals are currently receiving more per article than this figure, providing a surplus to either learned society or profit to the commercial publisher. Future demands for open access by funding agencies may have to be met with a transition to fully open access journals. Financial success and stability will have to come from superior journal (transition) management, online innovation, and cost savings.

From the point of view of a science institution it is also both. Open access guarantees a wider distribution and library budgets may go down as subscription charges disappear. There is a serious opportunity to save costs. But is public access worth losing the freedom to choose one's journal? Isn't it likely that an institutional library budget will simply be replaced by an institutional open access budget and an associated infrastructure? For large science producing universities, the new system may not be cheaper and cost savings may not flow back to the institute's overall budget. It is perhaps in this light that one can understand the move of Harvard, MIT, and Columbia to set up their own article repositories.

4 The Disappearance of Print

About 10 years ago, the final and official article publication was the version that got printed in a journal and the print date was equal to the published date. How this has changed today! The electronic article edition has become the final publication, upon its online appearance (although the accepted date still has the most important value). The print edition appears (much) later and is often incomplete. It lacks all clickable links to citations, figures, catalogs, data, etc. and possibly multimedia elements. It is no surprise that publishers are now preparing to cancel the printed journal editions in response to a diminishing functionality, usage, and demand.

For Springer, the year 2010 is likely to mark the tipping point in journal sales: less revenue in print subscriptions than in electronic subscriptions. In other words, libraries are swapping the print for electronic-only subscriptions. The same is happening for books, although with some delay. Springer's electronic book sales revenues are about 25% of the total book sales revenues.

The benefits of electronic publishing need no further explanation, although there is still a strong tendency to mimic the traditional printed article. The use of multi-

media is limited, the article structure and length remains the same, many scientists prefer to cite volume and page numbers, there is no inter-activity and no updates within the published article are allowed. But it is likely that the electronic article will evolve and move away from the traditional print format in several different ways.

Meanwhile, what are the remaining benefits of print? Are we losing something with its disappearance? Most concerns are centered around the fact that paper lasts for centuries, while we don't know the longevity or accessibility of electronic formats. Publishers still have library customers that buy almost everything in print, because they are trying to maintain strategic archives. In recognition of this need, organizations like Portico (www.portico.org), which was founded in 2002, are entirely focusing on a sustainable digital preservation. It seems there is no choice, the information boom simply outpaces our ability to capture all that's new in print.

5 The Future

How will the publishing landscape look like when we are celebrating Ken Freeman's 80th birthday? Will science results appear exclusively on handheld devices, next generations of Amazon Kindles and Apple iPads, enhanced by dozens of apps for better display and interactivity? Will the publishers have changed into mere IT companies, or perhaps been cut out completely as middle men in the information handling chain? Has the librarian become a broker for lowest open access fees? Whatever the future brings us, the safe bet seems that it will be very different from today.

But the familiar does not always disappear so fast. There is a striking irony emerging in this electronic era. The Internet is saving many printed and forgotten texts! Books and journal articles that were out of print for decades are re-digitized and put online. Google has become (in)famous for scanning more than 10 million books and publishers who own the copyrights are selling these again in print. It is the rapid development of high-quality digital printing techniques which will soon allow the printing "around the corner" of your favorite book in any format you desire, corrected and updated, colored and annotated, etc.

References

1. Archambault, E. 2010, Science-Metrix report "30 Years in Science", http://www.science-metrix.com/30years-Paper.pdf
2. Bostick, S.L. 2001, Liber Q, 11, 6–13
3. Davis, E.B., Schmidt, D. 1996, In: Guide to Information Sources in the Botanical Sciences, ISBN 1-56308-075-3, 2nd edn, Libraries Unlimited, Inc, Englewood, CO
4. Gaughan, P.A. 2002, In: Mergers, Acquisitions, and Corporate Restructurings, ISBN 978-0471121961, 3rd edn, Wiley, New York, NY

5. NSF report online, Globalization of Science and Engineering Research, http://www.nsf.gov/statistics/nsb1003/
6. Open Access List of all mandates and policies, http://www.eprints.org/openaccess/policysignup/
7. Potter, W.G. 1997, Libr Trends, 45(3), 416–434
8. Shachaf, P. 2003, Libri, 53, 94–102
9. SPARC Letter online www.sparceurope.org/news/EU%20Letter.doc

Super Star Clusters and Supernovae in Interacting LIRGs Unmasked by NIR Adaptive Optics

Petri Väisänen, Zara Randriamanakoto, Erkki Kankare, Seppo Mattila, and Stuart Ryder

Abstract We report on an ongoing near-IR adaptive optics survey targeting interacting luminous IR galaxies. High-spatial resolution NIR data are crucial to enable interpretation of kinematic, dynamical and star formation (SF) properties of these very dusty objects. Whole progenitor nuclei in the interactions can be missed if only optical HST imaging is used. Here we specifically present the latest results regarding core-collapse supernovae found within the highly extincted nuclear regions of these galaxies. Direct detection and study of such highly obscured CCSNe are crucial for revising the optically derived SN rates used for providing an independent measurement of the SF history of the Universe. We also present thus far the first NIR luminosity functions of super star cluster (SSC) candidates. The LFs can then be used to constrain the formation and evolution of SSCs via constraints based on initial mass functions and cluster disruption models.

1 Introduction

Interacting and merging galaxies are sites of violent star formation (SF) in extreme environments, as well as sites of the most massive known clustered star formation. Many of these objects are classified as luminous or ultra-luminous IR galaxies (LIRGS have $\log(L_{IR}/L_{\odot}) = 11-12$ and ULIRGs $\log(L_{IR}/L_{\odot}) > 12$). Although it is still debated exactly how much of the high-redshift SF is interaction triggered, and how local ULIRGs relate to their higher-z cousins, it nevertheless is clear that a significant fraction, if not most, of SF in the high-z universe is happening in these extreme environments. It is thus important to understand the physical processes in local examples, where detailed studies are possible, over a wide range of L_{IR} and SF rate (SFR) output, and a range in environment and interaction stage.

P. Väisänen (✉)
South African Astronomical Observatory, Cape Town, South Africa
e-mail: petri@saao.ac.za

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_13,

Moreover, much of the most violent SF is hidden behind dust. The total SF output can be estimated from far-IR luminosities, but the spatial resolution is usually too low to see more details of, e.g., the spatial distribution of SF. We have used NIR adaptive optics (AO) to image a sample of < 200 Mpc LIRGs to bridge the gap between high-resolution optical imaging and the mid/far-IR (U)LIRG studies. See [12, 20, 22] for more information of the observations and data. We have looked for the elusive population of core-collapse SNe and for populations of massive super star clusters (SSCs), in addition to studying the galaxies themselves.

2 Super Star Cluster Candidates

Most stars are born in clusters but most stars within galaxies do not live in clusters. This basic observation has led to a lively discussion on the physical processes involved in SSC formation and evolution, their disruption and survival and how they can be used to trace SF and dynamical histories of their host galaxies. SSCs with masses even up to 10^7 $M_\odot$ have been found in large numbers especially in interacting galaxies (e.g. [2] for a recent review). However, it still remains unclear whether the most massive clusters can only form in starburst environments or are SSCs seen in those locations just because of larger number statistics. SSC studies have thus far been driven by optical HST observations, and we now add the first significant NIR sample to be studied further. We detect from a dozen to hundreds of SSC candidates per galaxy in our sample of LIRGs (Fig. 1).

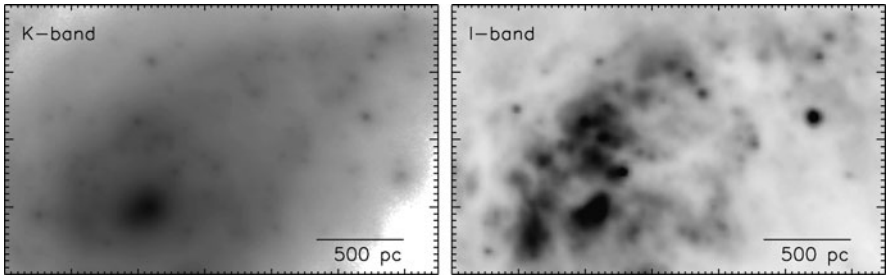


Fig. 1 A region of $6'' \times 4.5''$ northwest of the nucleus (*at bottom left*) of IRAS 18293-3413 showing dozens of SSC candidates. Our VLT/NACO data are on the *left* and HST/ACS I-data on the *right*: SSCs are only partially detected simultaneously in the two wavelengths because of a combination of extinction and age effects

2.1 The SSC NIR Luminosity Function

The current debate over basic SSC system characteristics involves questions such as: is the mass function (MF) of SSCs universal, or mass and galaxy dependent, are the MFs/LFs described with power laws, Schecter functions or Gaussians,

and is SSC disruption internally driven and mass independent or mass dependent and externally driven (see e.g. [3–5, 9] and references therein). There clearly is need for more data, preferably of new and varied targets and at different wavelengths.

While still waiting to complete our Gemini survey, we have already started putting together epoch-stacked deep images to study the LIRGs themselves. As a preliminary result, Fig. 2 shows the first ever K-band luminosity function of SSCs in external galaxies from six of our targets. The LF slope alone is a useful diagnostic to probe SC formation/evolution since different disruption models result in different LFs and would indicate different initial LFs and MFs (e.g. [6]). Intriguingly, the fitted LF slopes are fairly consistently around $\alpha \sim -1.6$, i.e. significantly shallower than in optical studies which typically find values of $\alpha \sim -2$. At face value, this would indicate support for models advocating non-universal, mass-dependent, LFs; in case of a universal power law index of $\alpha \sim -2$ with mass-independent disruption, even though the observed LF is a sum of SSCs of different ages and masses, the final LF should show the same $\alpha \sim -2$ slope.

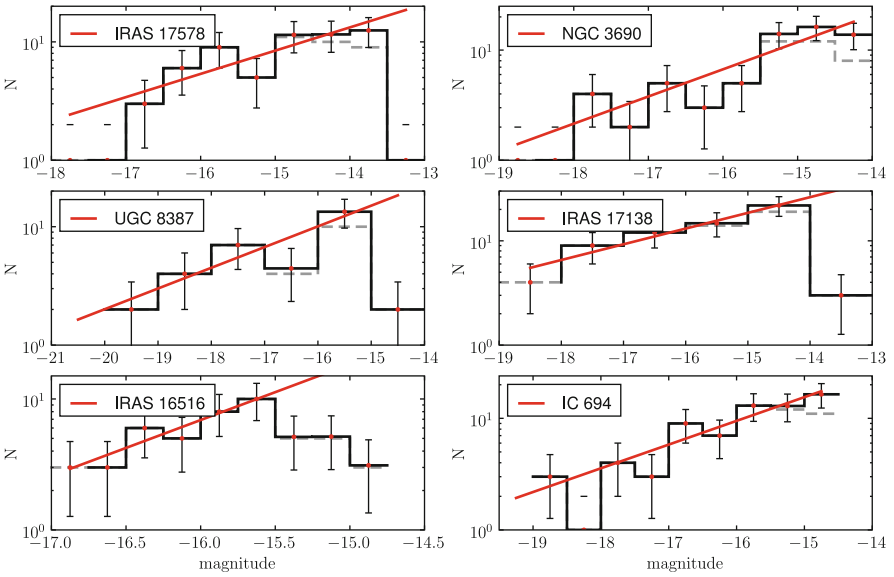


Fig. 2 A preliminary K-band LF of SC candidates in our sample of Gemini images. The *black histogram* and *red data points* with error bars show completeness corrected SSC number counts against K-band absolute magnitude, and the *straight lines* the weighted least-squares fit. The *dashed histogram* is the raw SSC count. Intriguingly, the power law LF index in these cases shows quite consistently $\alpha \sim -1.6$, significantly shallower than those seen in the optical

Detecting SSCs in the K-band might be expected to result in different slopes given both extinction and age effects, however – in the NIR we are seeing, at least partially, a more obscured population, and also likely a somewhat older population

compared to the optically detected SSCs. The age difference might not necessarily be widely different, however: our Starburst99 modelling indicates that the K-band is likely to pick out SSCs close to age $\sim$ 10 Myr, while the optical SSCs peak only slightly earlier (excluding extinction). Nevertheless, more careful analysis of completeness and selection effects are needed to securely reach implications from our preliminary SSC NIR-LF results presented here.

3 Searching for Extincted Core-Collapse SNe

SF-dominated (U)LIRGs are expected to hide in their central regions large numbers of undetected core-collapse supernovae (CCSNe), i.e. stars more massive than $\sim$8 $M_\odot$ exploding at the end of their (short) lives. LIRGs with SFRs of tens to hundreds of $M_\odot$/year can be expected to host $\sim 0.2 - 1$ CCSNe/year, a couple of orders of magnitude higher rate than in ordinary field galaxies. Radio observations have revealed hidden SN factories in the nuclear regions of (U)LIRGs, e.g. [16, 19]. Such SNe cannot be detected at optical wavelengths, even in the local universe, because of severe ($A_V > 10$ mag) dust extinction. However, they *must* exist in the nuclear (central kpc) regions of such galaxies, if the high SFRs derived from their IR luminosities and spectra are to be believed.

Previous NIR searches [10] have found an SN-rate lower by factors of several than expected. We have argued [11] that this rate misses the SNe exploding close to the centres of LIRGs in very extinguished regions and have an ongoing programme to search for these cases making use of AO-assisted NIR imaging.

SN2004ip was detected in a pilot study using VLT/NACO [12]. It was at a projected distance of 1.4″, or 500 pc, from the nucleus of IRAS 18293-3413, among the closest SNe detected (in IR) to a LIRG nucleus. Our subsequent radio observations confirmed its CCSN nature [15]. Our current Gemini/ALTAIR/NIRI search has so far produced 3 new SNe. The first one was SN2008cs [7], located at 4.2″, or 1.5 kpc, projected distance from the nucleus or IRAS 17138-1017. Follow-up observations in both radio and NIR bands were again consistent with a core-collapse event and showed the SN to suffer from a very high host galaxy extinction of $A_V \approx 16$ mag, the highest definite measurement yet for any SN. In addition, a "historical" SN2004iq was detected in the same galaxy from HST images and is located 660 pc from the nucleus. Earlier this year, three more SNe were detected in our Gemini programme, though two of them were actually first seen in non-AO optical and NIR searches [13, 14, 17] thanks to their relatively large galactocentric distances. The third one [8] in IC 883, however, turned out to be the closest SN to a LIRG nucleus yet discovered, at just 180 pc (or 0.4″) projected distance from the core (Fig. 3).

The intrinsic number of CCSNe based on the FIR luminosities for our sample of seven galaxies over the Gemini programme is calculated to be about 10 SNe. Although the number statistics are arguably still low, and though we have not finished the survey, with five SN detections in the sample it already appears clear that

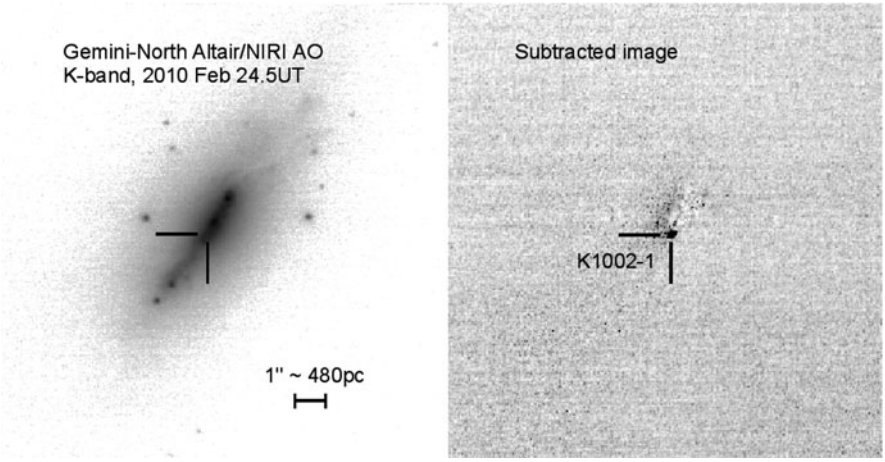


Fig. 3 The central regions of IC 883 observed with Gemini. The location of the recent SN candidate PSN K1002-1 is marked, though it is visible only in the subtracted image from two different epochs in the *right panel*. The projected distance to the core of the galaxy is a mere 180 pc

AO-assisted NIR observations provide an excellent window to detect SNe in the obscured nuclear regions of nearby LIRGs.

4 SF and Dynamics in Interacting Galaxies

We have followed-up in detail two of our early high-quality deep NACO AO images. In addition to combining the AO data with archive optical HST imaging we obtained spectroscopy using SALT and AAT. Both cases revealed surprises. We found an unexpected third component in the supposed pair interaction of IRAS 19115-2124, dubbed the Bird [20]. Moreover, this third, least massive and irregular component dominates the current star formation output of the whole system. This is in contrast to the widely held picture that tidal interactions are expected to drive large quantities of gas into the *central regions* of the interaction resulting in starbursts. While there is strong SF in the central regions as well, it is clearly the smaller "extra" component that elevates the Bird into almost a ULIRG-class object. We are in the process of studying this system further with integral field spectroscopy and mid-IR imaging to piece together what has happened in this intriguing system.

Follow-up of IRAS18293-3134, on the other hand, showed evidence of an extremely rare leading arm spiral, i.e. a galaxy where the spiral arms open up in the same direction as the disk is rotating [21]. Simulations have shown that some retrograde encounters should produce these kind of galaxies [18], though they have not been studied much and only a couple of candidates exist in the literature [1]. Their very existence, however, would have implications for, e.g., the dark matter halo mass of spirals.

Both of these systems highlight the fact that NIR AO imaging is crucial in interpreting the dynamical state of dusty LIRGs. Relying only on optical data, even HST data in the I band, will often not reveal all the components of the interactions and will not resolve the locations of major stellar mass distribution. Without the NIR AO data it would have been virtually impossible to disentangle the true velocity dispersions of the nuclei from other kinematic components such as tidal tails and gas outflows. Similarly, the leading arms of IRAS 18293-3134 would not have been possible to detect without NIR AO data, raising the tantalizing possibility that more such cases might be hidden in the chaotic interactions of dusty LIRGs.

5 Summary

We have presented results of our ongoing survey to unmask CCSNe in the highly extincted nuclear regions of interacting luminous IR galaxies, and presented, for the first time, luminosity functions of SSCs found in these galaxies in the NIR.

References

1. Byrd G.G., Freeman T., Howard S. 1993, AJ, 105, 447
2. de Grijs, R. 2010, RSPTA, 368, 693 (arXiv:0911.0778)
3. Elmegreen B.G., Hunter D.A. 2010, ApJ, 712, 604
4. Fall S.M., et al. 2009, ApJ, 704, 453
5. Gieles M. 2009, MNRAS, 394, 2113
6. Gieles, M. 2010, Galaxy Wars, In: eds B. Smith et al., ASPC, 423, 123 (arXiv:0908.2974)
7. Kankare E., et al. 2008, ApJ, 689, L97
8. Kankare E., et al. 2010, CBET, 2213, 1
9. Larsen S.S. 2009, A&A, 494, 539
10. Mannucci F., et al. 2003, A&A, 401, 519
11. Mattila S., Meikle W.P.S., Greimel R. 2004, New AR, 48, 595
12. Mattila S., et al. 2007, ApJ, 659, L9
13. Mattila S., et al. 2010, CBET, 2145, 1
14. Newton J., Puckett T., Orff, T. 2010, CBET, 2144, 1
15. Pérez-Torres, M.A., et al. 2007, ApJ, 671, L21
16. Pérez-Torres, M.A., et al. 2009, MNRAS, 399, 1641
17. Ryder S., et al. 2010, CBET, 2189, 1
18. Thomasson M., et al. 1989, A&A, 211, 25
19. Ulvestad J.S. 2009, AJ, 138, 1529
20. Väisänen P., et al. 2008, MNRAS, 384, 886
21. Väisänen P., Ryder S., Mattila S., Kotilainen J., et al. 2008, ApJ, 689, L37
22. Väisänen P., Mattila S., Ryder, S. 2010, Galaxy Wars, In: eds B. Smith et al., ASPC, 423, 323 (arXiv:0908.3495)

Structure, Mass, and Stability of Galactic Disks

Pieter van der Kruit

Abstract In this review I concentrate on three areas related to structure of disks in spiral galaxies. First I will review the work on structure, kinematics, and dynamics of stellar disks. Next I will review the progress in the area of flaring of HI layers. These subjects are relevant for the presence of dark matter and lead to the conclusion that disks are in general not "maximal", have lower M/L ratios than previously suspected, and are locally stable w.r.t. Toomre's Q criterion for local stability. I will end with a few words on "truncations" in stellar disks.

1 Stellar Disks

The structure and general properties of stellar disks have been reviewed in some detail by Ken and myself at previous honorary symposia (e.g., [27, 68]). The radial distribution of surface brightness can be approximated by an exponential [25] and the vertical distribution with an isothermal sheet [76] with a scaleheight that is independent of galactocentric distance. A more general form is [66]

$$L(R,z) = L(0,0)\,\mathrm{e}^{-R/h}\,\mathrm{sech}^{2/n}\left(\frac{nz}{2h_z}\right) \tag{1}$$

This ranges from the isothermal distribution ($n = 1$) to the exponential function ($n = \infty$) and allows for more realistic stellar distribution that are not exactly isothermal in z. From actual fits in I and K' de Grijs et al. [17] found

$$2/n = 0.54 \pm 0.20. \tag{2}$$

A detailed study by de Grijs and Peletier [16] has shown that the constancy of the vertical scale parameter h_z is very accurate in late-type spiral disks, but in early-types it may increase, in an extreme case by as much as 50% per scalelength h.

P. van der Kruit (✉)
Kapteyn Astronomical Institute, University of Groningen, 9700AV Groningen, The Netherlands
e-mail: vdkruit@astro.rug.nl

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_14,

The origin of the exponential nature of stellar disks is still uncertain. Freeman [25, 26] already pointed out that the distribution of angular momentum in a self-gravitating exponential disk resembles that of the uniform, uniformly rotating sphere [44]. This also holds for an exponential density distribution with a flat rotation curve [28, 65], so that a model with collapse with detailed conservation of angular momentum (see also [22]) would give a natural explanation for the exponential nature of disks (and maybe their truncations; see below). However, bars or other non-axisymmetric structures may induce severe redistribution of angular momentum; the work on the effects of nonaxisymmetric instabilities on the secular evolution of disks and their structural parameters by Debattista et al. [13] shows the potential of such approaches.

The distribution of the scale parameters is most easily studied in edge-on galaxies. Following on from the studies by van der Kruit and Searle [76–78], an extensive sample of edge-on galaxies has been studied by de Grijs [15] and has been re-analyzed by Kregel et al. [36]. The scale parameters correlate well with the rotation velocity of the galaxy, e.g., for the scaleheight

$$h_z = (0.45 \pm 0.05)\,(V_{\rm rot}/100\ {\rm km\ s^{-1}}) - (0.14 \pm 0.07)\ {\rm kpc} \tag{3}$$

with a scatter of 0.21 kpc. This could be useful to find a statistical estimate of the thickness of disks in galaxies that are not seen edge-on. The flattest galaxies (largest ratio of h and h_z) appear to be those with late Hubble type, small rotation velocity, and faint (face-on) surface brightness.

2 Stellar Kinematics

At the basis of the analysis of the vertical dynamics of a stellar disk we have the Poisson equation for the case of axial symmetry and at low z (e.g., [51])

$$\frac{\partial K_{\rm R}}{\partial R} + \frac{K_{\rm R}}{R} + \frac{\partial K_z}{\partial z} \approx 2(A-B)(A+B) + \frac{\partial K_z}{\partial z} = -4\pi G\rho(R,z) \tag{4}$$

For a flat rotation curve $A = -B$ and $2(A-B)(A+B) = 0$, so the plane-parallel case becomes an excellent approximation [74] . The equation of hydrostatic equilibrium

$$\sigma_z(R) = \sqrt{c\pi G \Sigma(R) h_z}, \tag{5}$$

relates the vertical velocity dispersion $\sigma_z(R)$ of the old stars to the surface density Σ so that if the mass-to-light ratio M/L is constant with radius, the exponential radial surface brightness distribution implies that $\sigma_z(R)$ should decline as the square root of Σ or also as an exponential with radius, but with an e-folding of twice the scalelength h (the constant c in Eq. (5) varies between 3/2 for an exponential with $n = \infty$ in Eq. (1) to 2 for an isothermal distribution with $n = 1$). This was first tested

in face-on spirals by van der Kruit and Freeman [73, 74], where the prediction was verified in detail in NGC 5347; in fact, the e-folding of σ_z was 2.4±0.6 photometric scalelengths. Many studies have since confirmed this decrease of σ_z with radius (e.g. [5, 37, 38], and references therein).

There are two recent developments that have a very strong impact on this issue. The first is the use of integral field units that enable a more complete sampling of the disks. The so-called *Disk Mass Project* [81, 83] aims at a mapping of the stellar vertical velocity dispersion in this manner in about 40 face-on spiral galaxies. As above this will provide a kinematic measurement of the mass surface density of stellar disks. Not many results have appeared in the literature yet, but recent conference presentations show that the 'kinematics follows the light', i.e., the velocity dispersions drop off according to the manner described above with constant M/L. The actual values indicate relatively low mass-to-light ratios that are well below those required for maximum disk fits (see below).

This result is also obtained in the other recent development, which is the use of planetary nebulae as test particles in disks [29] of five face-on spirals. This method allows the velocity dispersion of these representative stars of the old disk population to be measured out to large radii. In general the findings are the same: except for one system, the M/L is constant out to about three radial scalelengths of the exponential disks. Outside that radius the velocity dispersion stops declining and becomes flat with radius. Possible explanations these authors put forward for this behavior include an increase in the disk mass-to-light ratio, an increase in the importance of the thick disk, and heating of the thin disk by halo substructure. They also find that the disks of early type spirals have higher values of M/L and are closer to maximum disk than later-type spirals.

There is certainly support from stellar dynamics that in general there are no substantial gradients in mass-to-light ratios in disks. The rather low M/L values that are obtained currently do not require large amounts of material unaccounted for, as was found originally by Kapteyn [31] and Oort [50].

The stellar velocity dispersions *in* the plane are more complicated to determine from observations. The radial and tangential ones are not independent, but governed by the local Oort constants: $\sigma_\theta/\sigma_R = \sqrt{-B/(A-B)}$. This results from the axis ratio of the epicyclic motion that describes stellar orbits deviating little from circular. The frequency in the epicycle is $\kappa = 2\sqrt{-B(A-B)}$ and its axis ratio $\sqrt{-B/(A-B)}$ [51]. For a flat rotation curve $A = -B$, so σ_θ/σ_R (the "axis ratio of the velocity ellipsoid") is 0.71 and $\kappa = \sqrt{2}V_{rot}/R$.

In highly inclined or edge-on systems the dispersions can be measured both from the line profiles and the asymmetric drift equation

$$V_{rot}^2 - V_\theta^2 = \sigma_R^2 \left\{ \frac{R}{h} - R\frac{\partial}{\partial R}\ln(\sigma_R) - \left[1 - \frac{B}{B-A}\right]\right\} \tag{6}$$

where the circular velocity V_{rot} can be measured with sufficient accuracy from the gas (optical emission lines or HI observations), which have velocity dispersions of order 10 km/s or less and therefore very little asymmetric drift ($V_{rot} - V_\theta$).

The radial dispersion plays an important role in the Toomre [61] Q-criterion for local stability in galactic disks

$$Q = \frac{\sigma_{\rm R}\kappa}{3.36G\Sigma} \tag{7}$$

with Σ the local mass surface density. On small scales local stability results from a Jeans-type stability, where the tendency to collapse under gravity is balanced by the kinetic energy in random motions, but only up to a certain (Jeans) scale. On large scales, shear as a result of galactic differential rotation provides stability. In the Toomre Q-criterion the smallest scale for this is just equal to the Jeans scale, so that that local stability exists on *all* scales. According to Toomre [61], local stability requires $Q > 1$. Numerical simulations suggest that galaxy disks have $Q = 1.5 - 2.5$ and are on the verge of instability [2, 8, 30, 45, 57].

The first study where an attempt was made to measure these velocity dispersions was by van der Kruit and Freeman [74] on the highly inclined galaxy NGC 7184. They fitted their data using two different assumptions for the radial dependence of the radial velocity dispersion, one being that the axis ratio of the velocity ellipsoid is the same everywhere, and the other that Toomre Q is constant with radius. Over the range from the center of one or two scalelengths the assumptions work out to similar variations (see [67], p. 196).

More extensive observations on a sample of 12 galaxies (including the Milky Way Galaxy from [39]) by Bottema [5] resulted in the discovery of a relation between a fiducial value of the velocity dispersion (either the vertical one measured at or extrapolated to the center or the radial velocity dispersion at one scalelength) and the integrated luminosity or the rotation velocity (equivalent through the Tully-Fisher relation). This has been confirmed (see Fig. 1) by Kregel and van der Kruit [35] and Kregel et al. [38]:

$$\sigma_{\rm z|0} = \sigma_{\rm R|1h} = (0.29 \pm 0.10)V_{\rm rot}. \tag{8}$$

It actually extends to small dwarf galaxies, e.g., $19\,{\rm km\,s^{-1}}$ in UGC 4325 ([60], Chapter 7). Interestingly, the scatter in this relation is not random. Galaxies below the relation (with lower velocity dispersions) have higher flattening, lower central surface brightness, or dynamical mass ($4hV_{\rm rot}^2/G$) to disk luminosity ratio.

The linear $\sigma - V_{\rm rot}$ relation follows from straightforward arguments [5, 67, 72]. We evaluate now properties at one radial scalelength ($R = 1h$). For a flat rotation curve and eliminating h using a Tully–Fisher relation $L_{\rm disk} \propto \mu_\circ h^2 \propto V_{\rm rot}^4$ results in

$$\sigma_{\rm R} \propto Q\left(\frac{M}{L}\right)_{\rm disk}\mu_0^{1/2}V_{\rm rot} \tag{9}$$

This shows that when Q and M/L are constant among galaxies, the Bottema relation itself results and galaxy disks with lower (face-on) central surface brightness $\mu_\circ$ have lower stellar velocity dispersions than the mean.

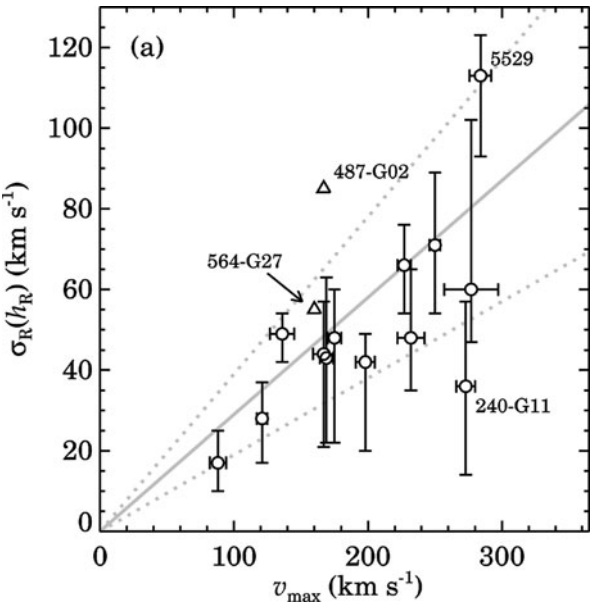


Fig. 1 Stellar disk velocity dispersion, measured at one scalelength in edge-on galaxies versus the maximum rotational velocity. The *gray lines* indicate the relation $\sigma_R(h) = (0.29 \pm 0.10)\ V_{rot}$ [5]. (From [38])

3 Mass-to-Light Ratios and "Maximum Disk"

For a disk that is exponential in both the radial and the vertical direction (cf. [66]):

$$\sigma_z(R,z) = \sqrt{\pi G h_z (2 - e^{-z/h_z})(M/L)\mu_0}\, e^{-R/2h} \tag{10}$$

Assuming a constant axis ratio of the velocity ellipsoid σ_z/σ_R, we find

$$\sigma_R(R,z) = \sqrt{\pi G h_z (2 - e^{-z/h_z})(M/L)\mu_0}\, \left(\frac{\sigma_z}{\sigma_R}\right)^{-1} e^{-R/2h} \tag{11}$$

The distribution of the products $\sqrt{M/L}\ (\sigma_z/\sigma_R)^{-1}$ in the Kregel et al. [38] sample is shown in Fig. 2. Thirteen of the fifteen disks have $1.8 \lesssim \sqrt{M/L}\ (\sigma_z/\sigma_R)^{-1} \lesssim 3.3$. The values of the outliers may have been overestimated (see Kregel et al., [38]). Excluding these, the average is $\langle\ \sqrt{M/L}\ (\sigma_z/\sigma_R)^{-1}\ \rangle = 2.5 \pm 0.2$ with a 1σ scatter of 0.6. The near constancy of the product can be used with M/L based on stellar population synthesis models to estimate the axis ratio of the velocity ellipsoid. Conversely, the upper scale of Fig. 2 indicates that a typical (I-band) M/L of a galactic stellar disk is of order unity and varies for the majority systems between 0.5 and 2.

It is possible to relate the axis ratio of the velocity ellipsoid to the flattening of the stellar disk h/h_z [72]. In the radial direction the velocity dispersion is related

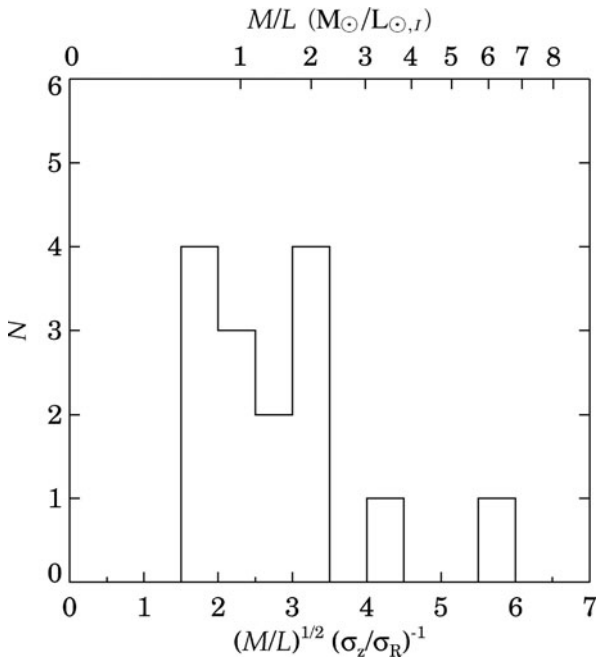


Fig. 2 Histogram of the product $\sqrt{M/L}\ (\sigma_z/\sigma_R)^{-1}$ from stellar kinematics in edge-on galaxies. Except for two outliers the distribution of $\sqrt{M/L}\ (\sigma_z/\sigma_R)^{-1}$ is rather narrow. The outliers are ESO 487-G02 and 564-G27; data for these galaxies are less complete than for the other ones. Along the *top* we show the values of M/L implied by $\sigma_z/\sigma_R = 0.6$. (From [38])

to the epicyclic frequency through the Toomre parameter Q for local stability. The Tully–Fisher relation then relates this to the integrated magnitude and hence to the disk scalelength. In the vertical direction the scaleheight and the velocity dispersion relate through hydrostatic equilibrium.

Equation (9) shows that when Q and M/L are constant among galaxies, disks with lower (face-on) central surface brightness $\mu_\circ$ have lower stellar velocity dispersions. Combining Eq. (9) with Eq. (5) for hydrodynamic equilibrium and using Eq. (8) gives [38, 72]

$$\frac{h}{h_z} \propto Q\left(\frac{\sigma_R}{\sigma_z}\right)\sigma_z^{-1}V_{rot} \propto Q\left(\frac{\sigma_R}{\sigma_z}\right) \tag{12}$$

The observed constancy of $\sqrt{M/L}\ (\sigma_z/\sigma_R)^{-1}$ implies that the flattening of the disk h/h_z is proportional to $Q\sqrt{M/L}$.

The mass-to-light ratio is a crucial measure of the contribution of the disk to the rotation curve and the relative importance of disk and dark halo mass in a galaxy. In the "maximum disk hypothesis" the disk contribution is optimized such that the amplitude of the disk-alone rotation curve is as large as the observations allow. In a maximal disk, the ratio between the disk-alone rotation curve and the observed one

will be a bit lower than unity to allow a bulge contribution and let dark halos have a low density core. A working definition is $V_{\rm disk}/V_{\rm rot} = 0.85 \pm 0.10$ [55].

For an exponential disk, the ratio of the peak rotation velocity of the disk to the maximum rotation velocity of the galaxy ($V_{\rm disk}/V_{\rm rot}$) is

$$\frac{V_{\rm disk}}{V_{\rm rot}} = \frac{0.880\ (\pi G\ \Sigma_0\ h)^{1/2}}{V_{\rm rot}} \tag{13}$$

Using Eqs. (5) and (8) this can be rewritten as

$$\frac{V_{\rm disk}}{V_{\rm rot}} = (0.21 \pm 0.08)\sqrt{\frac{h}{h_{\rm z}}} \tag{14}$$

So we can estimate the disk contribution to the rotation curve from a statistical value for the flattening (see also [5, 7, 68]). For the sample of Kregel et al. [36] this then results in $V_{\rm disk}/V_{\rm rot} = 0.57 \pm 0.22$ (rms scatter). In the dynamical analysis of Kregel et al. [38], the ratio $V_{\rm disk}/V_{\rm rot}$ is known up to a factor $\sigma_{\rm z}/\sigma_{\rm R}$ and distance-independent. For $\sigma_{\rm z}/\sigma_{\rm R} = 0.6$, $v_{\rm disk}/v_{\rm rot} = 0.53 \pm 0.04$, with a 1σ scatter of 0.15. Both estimates agree well. Thus, at least for this sample, the average spiral has a submaximal disk.

The values obtained for individual galaxies are illustrated in Figs. 3 and 4. Most galaxies are not "maximum-disk". The ones that may be maximum disk have a high surface density according to Fig. 2. From the panels we also note that disk that are maximal appear to have more anisotropic velocity distributions or are less stable according to Toomre Q,

Originally, van der Kruit and Freeman [74] used the property Y defined as a criterion for global stability [20]: $Y = V_{\rm rot}\sqrt{h/GM_{\rm disk}} \gtrsim 1.1$ for stability. This is – up to a factor 0.62 for an infinitesimally thin exponential disk – the reciprocal of the ratio $V_{\rm disk}/V_{\rm rot}$. Then they found $Y = 1.0 \pm 0.3$ [74], which corresponds to $V_{\rm disk}/V_{\rm rot} = 0.6 \pm 0.2$. It is remarkable – and probably significant – that all values obtained for $V_{\rm disk}/V_{\rm rot}$ are close to 0.6.

4 Flaring of HI Layers and Disk Masses

The thickness of the gas layer in a disk galaxy can be used to measure the surface density of the disk. Assume the density distribution of the exponential, locally isothermal disk (that in Eq. (1) with $n = 1$). If the HI velocity dispersion $\langle V_{\rm z}^2\rangle_{\rm HI}^{1/2}$ is independent of radius – as, e.g., in the face-on spiral NGC 628, [59] – *and* isotropic, and if the stars dominate the gravitational field, the HI layer has a full width at half maximum (to $\lesssim 3\%$) of

$$W_{\rm HI} = 1.7\langle V_{\rm z}^2\rangle_{\rm HI}^{1/2}\left[\frac{\pi G(M/L)\mu_\circ}{z_\circ}\right]^{-1/2} \exp\,(R/2h) \tag{15}$$

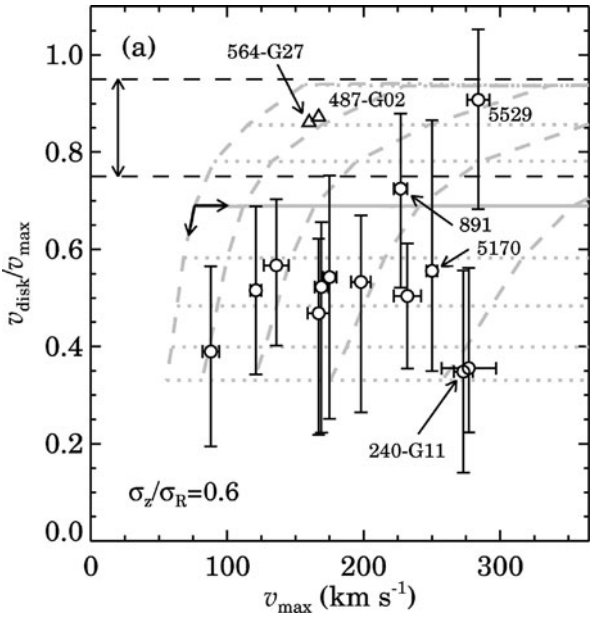


Fig. 3 The contribution of the disk to the amplitude of the rotation curve $V_{\rm disk}/V_{\rm rot}$ for a sample of 15 edge-on galaxies as a function of the rotation velocity itself. The *horizontal dashed lines* are the limits of 0.85 ± 0.10 from Sackett [55], which would indicate maximal disks. The axis ratio of the velocity ellipsoid is assumed to be 0.6. The *gray lines* correspond to collapse models of Dalcanton et al. [12]. The two without error bars are the same galaxies as the outliers in Fig. 2. (From [38])

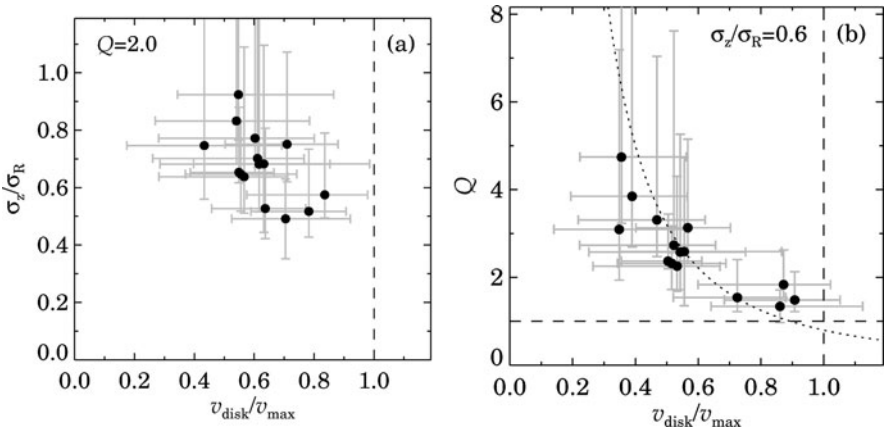


Fig. 4 Stellar dynamics parameters for edge-on galaxies. **a** The axis ratio of the velocity ellipsoid as a function of $V_{\rm disk}/V_{\rm rot}$ for $Q = 2.0$. (b) $V_{\rm disk}/V_{\rm rot}$ as a function of Q for an assumed axis ratio of the velocity ellipsoid of 0.6. (From [38])

So the HI layer increases exponentially in thickness with an e-folding $2h$. This has first been derived and applied to HI observations of NGC 891 by van der Kruit [64]. One has to be careful to distinguish signatures for flaring from those of residual inclination away from exactly edge-on. Such studies can determine whether galaxies have in general *maximum disks* or not. The modeling, using photometry from van der Kruit and Searle [77], indicated that $V_{\rm rot,disk}$ of NGC 891 is $\sim$ 140 km/s. The observed value is 225 $\pm$ 10 km/s, so the ratio in Eq. (1) is $\sim$ 0.6, and NGC 891 is clearly sub-maximal.

For our Galaxy my preferred values are $V_{\rm rot,disk} \sim 155 \pm 30$ and $V_{\rm rot,obs} \sim 225 \pm 10$ km/s, so that the ratio is 0.69 ± 0.14 and the Milky Way also is sub-maximal. In other systems similar results were found; e.g., in NGC 4244 Olling [48] deduced from the flaring a disk-alone rotation of 40–80% of that of observed rotation. Actually, the flaring of the HI layer in NGC 4244 was used by Olling [47] to infer that the dark matter is highly flattened (but see Olling and Merrifield [49], who found for the Galaxy halo closer to spherical).

In a recent study by O'Brien et al. [46] on the superthin edge-on galaxy UGC 7321, the rotation curve was decomposed using constraints from the thickness and flaring of the HI layer. This study was aimed at a determination of the shape of the dark matter halo, which was found to be close to spherical. The disk (I-band) M/L was found to be only about 0.2 and the galaxy is very far from maximum disk. This M/L is even somewhat lower than the range 0.5–2 indicated above. In fact, Banerjee et al. [3] also concluded that the dark matter dominates the gravitational field everywhere in the disk of this galaxy.

Work on decompositions of rotation curves and analysis involving observations of the "baryonic" Tully–Fisher relation has progressed as well. In particular, McGaugh [42] has analyzed a sample of galaxies with extended HI rotation curves and finds that high surface brightness galaxies are closer to maximum disk than low surface brightness ones. This may be the same effect as described above in terms of surface density. On the other hand, Weiner et al. [82] find from detailed fluid dynamical gasflows in the barred galaxy NGC 4123 that this system must be close to maximum disk. So, almost certainly some (massive) disks *are* maximal.

Finally I note the following clever argument of Courteau and Rix [11] that makes use of the *scatter* in the Tully–Fisher relation. The amplitude of the rotation curve of the self-gravitating exponential disk is

$$V_{\rm max} \propto \sqrt{M_{\rm disk}/h} \tag{16}$$

For fixed disk-mass $M_{\rm disk}$ we then get from differentiation

$$\frac{\partial \log V_{\rm max}}{\partial \log h} = -0.5 \tag{17}$$

So at a given absolute magnitude (or mass) a lower scalelength disks should have a higher rotation velocity. If all galaxies were maximum disk this then should be visible in the scatter of the Tully–Fisher relation. This is *not* observed and the estimate

is that on average $V_{\rm disk} \sim 0.6 V_{\rm total}$ and galaxies in general do not have maximal disks.

Recent reviews of disk masses in galaxies are by van der Kruit [71] and McGaugh [43] at the Kingston symposium. In summary, the overall state of affairs concerning the maximum disk hypothesis appears to be that in general galaxy disk are *not* maximal, except possibly the ones with the highest surface brightness and surface density.

5 Flatness and Truncations in Stellar Disks

It is important to realize that stellar disks are often remarkably flat. This can be studied in edge-on systems by determining the centroid of the light distribution in the direction perpendicular to the major axis at various galactocentric distances (e.g., [14, 23, 56], Chapter 5). Apart from some minor warps in the outer parts of the stellar disks, in the inner parts the systematic deviations are very small.

We may also look at the flatness of the layers of the ISM within them, such as dustlanes. In Fig. 5 I collected some images of edge-on disk galaxies. At the top are two "super-thin" galaxies; the disks are straight lines to within a few percent. The same holds for the dustlanes in NGC 4565 (allow for the curvature due to the imperfect edge-on nature) and NGC 891. In the third row the peculiar structure of NGC 5866 has no measurable deviation from a straight line, while for the Sombrero Nebula the outline of the dustlane fits very accurately to an ellipse. In the bottom row, NGC 7814 (right) is straight again to within a few percent, but NGC 5866 is an example of a galaxy with a large warp in the dust layer.

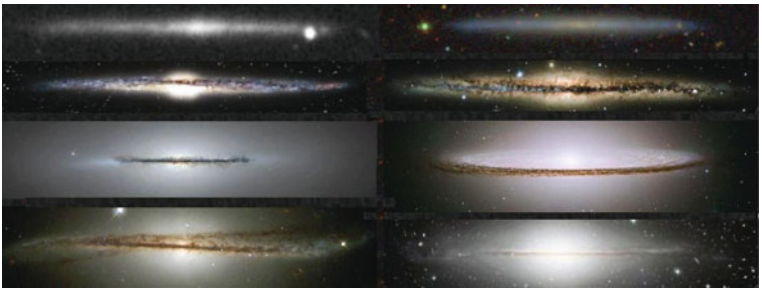


Fig. 5 Selected images of edge-on disks and dust lanes from various public Web-galleries. *Top*: 'Superthin' galaxies IC 5249 (from the Sloan Digital Sky Survey, [75]) and UGC 7321 (cosmo.nyu.edu/hogg/rc3/UGC_732_irg_hard.jpg); *second row*: NGC 4565 (www.cfht.hawaii.edu/HawaiianStarlight/AIOM/English/2004/Images/Nov-Image2003-CFHT-Coelum.jpg) and NGC 891 (www.cfht.hawaii.edu/HawaiianStarlight/Posters/ NGC891-CFHT-Cuillandre-Coelum-1999.jpg); *third row*: NGC 5866 (heritage.stsci.edu/ 2006/24/big.html) and M104 (heritage.stsci.-edu/2003/28/big.html); *bottom row*: ESO 510-G013 (heritage.stsci.edu/2001/23/big.html) and NGC 7814 (www.cfht.hawaii. edu/HawaiianStarlight/English/Poster50x70-NGC7814.html)

The HI kinematics provide probably the strongest indications for flatness. In three almost completely face-on spirals (NGC 3938, 628 and 1058), van der Kruit and Shostak [79, 80] and Shostak and van der Kruit [59] found that the residual velocity field after subtraction of that of the rotation field has an r.m.s. value of only 3–4 km/s (or a few pc per Myr) without any systematic pattern. A vertical oscillation with a similar period as that for stars in the Solar Neighborhood (10^7 years) or even of that of rotation around the Galactic Center (10^8 years) would correspond to a vertical amplitude ten to a hundred pc. The absence of such residual patterns shows that the HI layers and the stellar disks must be extraordinarily flat, except maybe in their outer regions or when they have recently been in interaction.

Recently, Matthews and Uson [40, 41] have found evidence for a pattern of corrugation in the disk of the edge-on galaxy IC 2233 with an amplitude up to 250 pc, especially in HI and young stars. IC 2233 is a rather small galaxy (radius 7 or 8 kpc and rotation velocity about 100 km/s) unlike the ones discussed in the previous paragraph.

I have indicated above that the flattening of the stellar disk h_z/h is smallest for systems of late Hubble type, small rotation velocity, and faint (face-on) surface brightness. It is of interest then to look more closely at systems at this extreme end of the range of flattening: "superthin" galaxies. A prime example is the galaxy UGC 7321, studied extensively by Lynn Matthews and collaborators ([3], and references therein). The picture that appears is that this is a very low surface brightness galaxy (the face-on B-band central surface brightness is $\sim$23.4 mag arcsec^{-2}) and a scale-length of about 2 kpc, but a projected vertical scaleheight of only 150 pc. It appears to have vertical structure since there is a color gradient (bluer near the central plane) and appears to consist of two components. Its HI is warped in the outer parts, starting at the edge of the light distribution.

Another good example of a superthin galaxy is IC 5249 [1, 9, 75]. This also is a faint surface brightness galaxy with presumably a small fraction of the mass in the luminous disk. However, the disk scaleheight is not small (0.65 kpc). It has a very long radial scalelength (17 kpc); the faint surface brightness then causes only the parts close to the plane to be easily visible against the background sky, while the long radial scalelength assures this to happen over a large range of R. Therefore it appears thin on the sky. The flattening h_z/h is 0.09 (versus 0.07 for UGC 7321). The stellar velocity dispersions are similar to those in the Solar Neighborhood; disk heating must have proceeded at a pace comparable to that in the Galaxy.

The flattest galaxies on the sky have indeed very small values of h_z/h. However, these two examples show that superthin galaxies share at least the properties of late type, faint face-on surface brightness, and small amounts of luminous disk mass compared to that in the dark halo.

Truncations in stellar disks were first found in edge-on galaxies, where the remarkable feature was noted that the radial extent did not grow with deeper photographic exposures [63]. Prime examples of this phenomenon of truncations are the large edge-on galaxies NGC 4565 and NGC 5907 (see Fig. 6). The truncations appear very sharp, although of course not infinitely so. Rather sharp outer profiles

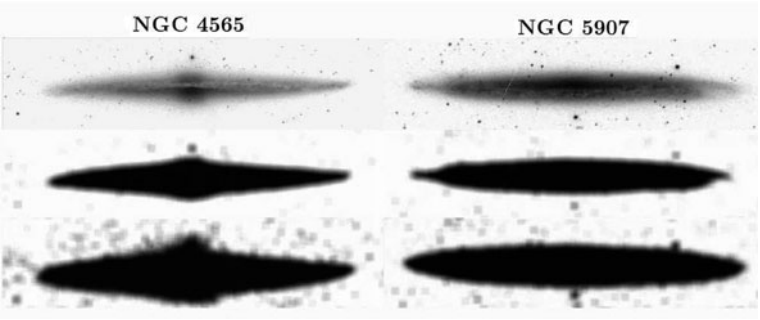


Fig. 6 NGC4565 and NGC 5907 at various light levels. These have been produced from images of the Sloan Digital Sky Survey, which were clipped at three different levels (*top to bottom*) and turned into two-bit images and subsequently smoothed (see [69], for an explanation of the details). Note that the disks grow significantly along the minor axes but not in radial extent

are actually obtained after deprojecting near-IR observations of edge-on galaxies (e.g., [24]).

Various models have been proposed for the origin of truncations (for a discussion see [69]). Truncations could be the current extent of the disks while they are growing from the inside out from accretion of external material. This predicts larger age gradients across disks than are observed [18]. Another possibility is that star formation is inhibited when the gas surface (or space?) density falls below a certain threshold for local stability [22, 32, 58]. The Goldreich–Lynden-Bell criterion for stability of gas layers gives a poor prediction for the truncation radii [78]. Another problem for the threshold hypothesis is that the rotation curves of some galaxies, e.g., NGC 5907 and NGC 4013 [5, 10], show features near the truncations that indicate that the *mass* distributions are also truncated. Schaye predicts an anti-correlation between $R_{\max}/h$ and h, which is not observed. Models have been proposed (e.g., [24]), in which a magnetic force breaks down as a result of star formation so that stars escape. The evidence for sufficiently strong magnetic fields needs strengthening.

Obviously, the truncation corresponds to the maximum in the current specific angular momentum distribution of the disk, which would correspond to that in the protogalaxy [65] if the collapse occurs with detailed conservation of specific angular momentum. As noted above, if this starts out as a Mestel [44] sphere (i.e., uniform density and angular rotation) in the force field of a dark halo with a flat rotation curve, a roughly exponential disk results. This disk has then a truncation at about 4.5 scalelengths, so this hypothesis provides at the same time an explanation for the exponential nature of disk as well as for the occurrence of the truncations. This requires the absence of substantial redistribution of angular momentum takes place. Bars may play an important role in this, as suggested by Debattista et al. [13] and Erwin et al. [21]. In fact a range of possible agents in addition to bars, such as density waves, heating, and stripping of stars by bombardment of dark matter subhalos, has been invoked [19].

Kregel and van der Kruit [34] derive correlations of the ratio of the truncation radius $R_{\max}$ and the disk scalelengths h with h itself and with the face-on

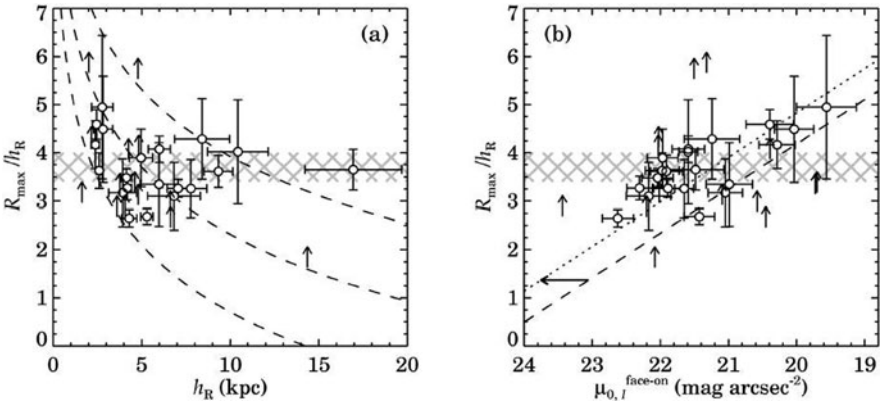


Fig. 7 Correlations of R_{max}/h with scalelength h and face-on central surface brightness $\mu_{\circ,fo}$ for a sample of edge-on galaxies. The cross-hatched regions show the prediction from a collapse model as in van der Kruit [65] and Dalcanton et al. [12]; the *dotted* and *dashed lines* show predictions from the star formation threshold model of Schaye [58] for three different values of the disk mass (from [34], see there for details)

central surface brightness $\mu_{\circ,fo}$ (Fig. 7). R_{max}/h does not depend strongly on h, but is somewhat less than the 4.5 predicted from the collapse from a simple Mestel-sphere. There is some correlation between R_{max}/h and $\mu_{\circ,fo}$, indicating approximate constant disk surface density at the truncations, as possibly expected in the star-formation threshold model. But this model predicts an anti-correlation between R_{max}/h and h [58], which is not observed. The maximum angular momentum hypothesis predicts that R_{max}/h should not depend on h or $\mu_{\circ,fo}$ and such a model therefore requires some redistribution of angular momentum or somewhat different initial conditions.

Due to line-of-sight integration, truncations will be more difficult to detect in face-on galaxies. The expected surface brightness at 4 scalelengths is about 26 B-mag arcsec^{-2} or close to sky. In face-on galaxies like NGC 628 [59, 66] an isophote map shows that the outer contours have a much smaller spacing than the inner ones. The usual analysis uses an inclination and major axis determined from kinematics (if available, otherwise this is estimated from the average shape of isophotes) and then determines an azimuthally averaged radial surface brightness profile. But this will smooth out any truncation if its radius is not exactly constant with azimuthal angle. The effects are nicely illustrated in the study of NGC 5923 ([52], their Fig. 9), which has isophotes in polar coordinates. The irregular outline shows that some smoothing out will occur contrary to observations in edge-on systems.

Pohlen and Trujillo [53] studied a sample of moderately inclined systems through ellipse-fitting of isophotes in SDSS data. They distinguish three types of profiles: *Type I*: no break; *Type II*: downbending break; *Type III*: upbending break. Pohlen et al. [54] have reported that the same types profiles occur among edge-on systems; however, of their 11 systems there were only one for each of the types I and III.

Various correlations have been reviewed in van der Kruit [71]. In general, the edge-on and face-on samples agree in the distribution of R_{max}/h; however, the fits

in moderately inclined systems result in small values of the scalelength compared to the edge-on sample. S. Peters, R. S. de Jong, and I have re-analyzed the Pohlen et al. data using two approaches: (1) mimic an edge-on view by collapsing the data onto the major axis and (2) calculate a radial profile using equivalent profiles. The luminosity profiles from ellipse fitting (Pohlen et al.) and that using equivalent profiles agree well, in spite of the difference that the first assumes a position for the center and the method with equivalent profiles does not. Often the "major-axis-collapse" method shows in Types I *and* II truncations when seen "edge-on." So, there are truncations in the stellar disks but less symmetric than one might expect. Finally, Type III galaxies do not show "edge-on truncations," but invariably evidence for interaction or other disturbances of the outer parts. A prime example of a Type III profile is NGC 3310, which is a well-known case of a disturbed, probably merging galaxy [33, 62].

There is a good correlation between $R_{\rm max}$ and the rotation velocity [70]. On average a galaxy like our own would have an $R_{\rm max}$ of 15–25 kpc (and a scalelength of 4–5 kpc). Now look at NGC 300, which has no truncation even at 10 scalelengths [4], so that $R_{\rm max} > 14.4$ kpc. In spite of that it is not outside the distribution observed in edge-on systems between $R_{\rm max}$ and $V_{\rm rot}$ (NGC 300 has $\sim$ 105 km/s and this would give an $R_{\rm max}$ of 8–15 kpc and an h of 2–4 kpc). So it has a unusually small h for its $V_{\rm rot}$; not an unusual $R_{\rm max}$ for its rotation! At least some of the Type I galaxies could have disks with normal truncation radii, but large $R_{\rm max}/h$ and small h so that the truncations occur at much lower surface brightness.

I note, but cannot discuss in detail, that truncations in stellar disks and warps of HI layers are often associated and refer to my discussion in van der Kruit [69].

Acknowledgments Ken Freeman is an expert in many area's of astronomy, but he is in particular known for his research in that of disks of spiral galaxies, and I feel fortunate to have been able to work with him on projects related to that. Many congratulations, Ken, and thanks for all the years of friendship and stimulating collaboration.

References

1. Abe, F. et al. 1999, AJ, 118, 261
2. Athanassoula, E., Sellwood, J.A. 1986, MNRAS, 221, 213
3. Banerjee, A., Matthews, L.D., Jog, C.J. 2010, NewA, 15, 89
4. Bland-Hawthorn, J., Vlajić, M., Freeman, K.C., Draine, B.T. 2005, ApJ, 629, 249
5. Bottema, R. 1993, A&A, 275, 16
6. Bottema, R. 1996, A&A, 306, 345
7. Bottema, R. 1997, A&A, 328, 517
8. Bottema, R. 2003, MNRAS, 344, 358
9. Byun, Y.-I. 1998, ChJAA, 36, 677
10. Casertano, S. 1983, MNRAS, 203, 735
11. Courteau, S., Rix, H.-W. 1999, ApJ, 513, 561
12. Dalcanton, J.J., Spergel, D.N., Summers, F.J. 1997, ApJ, 482, 659
13. Debattista, V.P., Mayer, L., Carollo, C.M., Moore, B., Wadsley, J., Quinn, T. 2006, ApJ, 645, 209

14. de Grijs, R. 1997, Ph.D. thesis, Univ. Groningen, The Netherlands
15. de Grijs, R. 1998, MNRAS, 299, 595
16. de Grijs, R., Peletier, R.F. 1997, A&A, 320, L21
17. de Grijs, R., Peletier, R.F., van der Kruit, P.C. 1997, A&A, 327, 966
18. de Jong, R.S. 1996, A&A, 313, 45
19. de Jong, R.S. et al. 2007, ApJ, 667, L49
20. Efstathiou, G., Lake, G., Negroponte, J. 1982, MNRAS, 199, 1069
21. Erwin, P., Pohlen, M., Beckman, J.E., Gutierrez, L., Aladro, R. 2008, ASP. Conference Series, 390, 251
22. Fall, S.M., Efstathiou, G. 1980, MNRAS, 193,189
23. Florido, E., Prieto, M., Battaner, E., Mediavilla, E., Sanchez-Saavedra, M.L. 1991, A&A, 242, 301
24. Florido, E. et al. 2006, A&A, 455, 467
25. Freeman, K.C. 1970, ApJ, 160, 811
26. Freeman, K.C. 1975, In: Stars & Stellar Systems, Chicago University Press, Chicago, p.9, Chap. 11
27. Freeman, K.C. 2007, Ap&SSP, 24, 3
28. Gunn, J.E., 1982, In: Astrophysical Cosmology, Pontifical Academy of Sciences, Vatican, 233
29. Herrmann, K.A., Ciardullo, R. 2009, ApJ, 705, 1686
30. Hohl, F. 1971, ApJ, 168, 343
31. Kapteyn, J.C. 1922, ApJ, 55, 302
32. Kennicutt, R.C. 1989, ApJ, 344, 685
33. Kregel, M., Sancisi, R. 2001, A&A, 376, 59
34. Kregel, M., van der Kruit, P.C. 2004, MNRAS, 355, 143
35. Kregel, M., van der Kruit, P.C. 2005, MNRAS, 358, 481
36. Kregel, M., van der Kruit, P.C., de Grijs, R. 2002, MNRAS, 334, 646
37. Kregel, M., van der Kruit, P.C., Freeman, K.C. 2004, MNRAS, 351, 1247
38. Kregel, M., van der Kruit, P.C., Freeman, K.C. 2005, MNRAS, 358, 503
39. Lewis, J.R., Freeman, K.C. 1989, AJ, 97, 139
40. Matthews, L.D., Uson, J.M. 2008a, AJ, 135, 291
41. Matthews, L.D., Uson, J.M. 2008b, ApJ, 688, 237
42. McGaugh, S.S. 2005, ApJ, 632, 859
43. McGaugh, S.S. 2009, In: Unveiling The Mass, www.astro.queensu.ca/GalaxyMasses09/data/McGaugh_GMasses09.pdf
44. Mestel, L. 1963, MNRAS, 126, 553
45. Mihos, J.C., McGaugh, S.S., de Blok, W.J.G. 1997, ApJ, 477, 79
46. O'Brien, J.C., Freeman, K.C., van der Kruit, P.C. 2010, A&A, 515, A63
47. Olling, R.P. 1996, AJ, 112, 481
48. Olling, R.P. 1996, AJ, 112, 457
49. Olling, R., Merrifield, M.R. 2000, MNRAS, 358, 503
50. Oort, J.H. 1932, BAN, 6, 249
51. Oort, J.H. 1965, In: Stars & Stellar Systems, University of Chicago Press, Chicago, IL, vol. 5, Chap. 21
52. Pohlen, M., Dettmar, R.-J., Lütticke, R., Aronica, G. 2002, A&A, 392
53. Pohlen, M., Trujillo, I. 2006, A&A, 454, 759
54. Pohlen, M., Zaroubi, S., Peletier, R.F., Dettmar, R.-J. 2007, MNRAS, 378, 594
55. Sackett, P.D. 1997, ApJ, 483, 103
56. Sanchez-Saavedra, M.L., Battaner, E., Florido, E. 1990, Ap&SS, 171, 239
57. Sellwood, J.A., Carlberg, R.G. 1984, ApJ, 282, 61
58. Schaye, J. 2004, ApJ, 609, 667
59. Shostak, G.S., van der Kruit, P.C. 1984, A&A, 132, 20
60. Swaters, R.A. 1999, Ph.D. Thesis, University of Groningen, Groningen
61. Toomre, A. 1964, ApJ, 139, 1217

62. van der Kruit, P.C. 1976, A&A, 49, 161
63. van der Kruit, P.C. 1979, A&As, 38, 15
64. van der Kruit, P.C. 1981, A&A, 99, 298
65. van der Kruit, P.C. 1987, A&A, 173, 59
66. van der Kruit, P.C. 1988, A&A, 192, 117
67. van der Kruit, P.C. 1990, In: eds G. Gilmore, I.R. King, P.C. van der Kruit, *The Milky Way as a Galaxy*, Univ. Science Books, Mill Valley, CA, Chap. 10
68. van der Kruit, P.C. 2002, ASP Conf Ser, 273, 7
69. van der Kruit, P.C. 2007, A&A, 466, 883
70. van der Kruit, P.C. 2008, ASP Conf Ser, 396, 173
71. van der Kruit, P.C. 2009, In: Unveiling the Mass, www.astro.queensu.ca/GalaxyMasses09/data/vanderKruit_GMasses09.pdf
72. van der Kruit, P.C., de Grijs, R., 1999, A&A, 352, 129
73. van der Kruit, P.C., Freeman, K.C. 1984, ApJ, 278, 81
74. van der Kruit, P.C., Freeman, K.C., 1986, ApJ, 303, 556
75. van der Kruit, P.C., Jiménez-Vicente, J., Kregel, M., Freeman, K.C. 2001, A&A, 379, 374
76. van der Kruit, P.C., Searle, L. 1981, A&A, 95, 105
77. van der Kruit, P.C., Searle, L. 1981, A&A, 95, 116
78. van der Kruit, P.C., Searle, L. 1982, A&A, 110, 61
79. van der Kruit, P.C., Shostak, G.S. 1982, A&A, 105, 351
80. van der Kruit, P.C., Shostak, G.S. 1984, A&A, 134, 258
81. Verheijen, M.A.W., Bershady, M.A., Swaters, R.A., Andersen, D.R., Westfall, K.B., 2007, Island Unicverses In: ed. R.S. de Jong, Springer New York, NY, p. 95
82. Weiner, B.J. Sellwood, J.A., Williams, T.B. 2001, ApJ, 546, 931
83. Westfall, K.B., Bershady, M.A., Verheijen, M.A.W., Andersen, D.R., Swaters, R.A. 2008, ASP Conf Ser 396, 41

What Can the Radial Surface Brightness Profiles of Galaxy Discs Tell Us About Their Evolution?

John E. Beckman, Leonel Gutiérrez, Peter Erwin, Ruyman Azzollini, and Inma Martínez-Valpuesta

Abstract Deriving the azimuthally averaged radial surface brightness profile of a galaxy is a schematic way of presenting the light and then the mass structure of disc galaxies. In this overview we will dwell on the outer discs, as it has become possible, with modern data to explore the light profile out to ever increasing galactocentric radii. We find that these profiles can be well classified into three main types: truncated profiles, untruncated profiles, and antitruncated profiles, denoted as Types II, I, and III, respectively, amplifying an original classification into Types I and II by Freeman. We present statistics of these types over the range of the Hubble galaxy classification for 187 nearby galaxies. We then analyse a sample of disc profiles for a sample of nearly 500 galaxies in three redshift bins, out to z = 1.1, from HST data. For the Type II discs we show that the "break" radii where the inner disc slope changes to a more abrupt outer slope have increased with epoch, but that at all epochs this radius has the bluest, i.e. the youngest stellar population. This can be explained via models which combine stellar migration due to non-axisymmetry with continuous star formation due to gas infall, one version of inside-out disc growth. We note that the observed features of Type I and Type III profiles are not explained in this scenario and show the need for even deeper profiles, especially in the IR, the UV, and Hα.

1 Profile Classification: A Brief Historical Summary

The radial surface brightness profile of a disc galaxy is a one-dimensional map of its structure giving us a quantitative summary of its properties. In the context of the present conference it is easy for me to pay a tribute to Ken Freeman by noting that he initiated the modern study of brightness profiles in the article [13] showing a systematic difference between profiles of discswith a single radially declining

J.E. Beckman (✉)
Instituto de Astrofísica de Canarias, Spain; Consejo Superior de Investigaciones Científicas, Spain; Departmento de Astrofísica, Universidad de La Laguna, Spain
e-mail: jeb@iac.es

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_15,

exponential and those with a second exponential with a lesser slope in the inner disc. Ken termed the former Type I profiles and the latter Type II. One explanation for Type II profiles is that the star-forming gas has been removed from the inner disc, possibly going into bulge formation. In recent years I have been involved in studies designed to obtain and interpret better observations of outer discs by measuring accurate radial brightness profiles. In these studies we have based much of our thinking on Ken's initial binomial classification, though we have amplified it to take in the wealth of newer observations. A second point of reference for our work was the presentation by Piet van der Kruit and Len Searle ([31, 32] and subsequent publications in the 1980s) of truncation as an important element of outer disc light profiles. They showed, using data from images of edge-on discs, that their light profiles were quite sharply truncated at the edges and that this truncation occurred at some 2–3 scale lengths from the centre. These papers were based on photographic observations taken at the end of the 1970s and some 20 years later, with the implied technical advances, it seemed useful to look again at disc edge profiles. This work was initiated by Michael Pohlen (see Pohlen et al. [22]). He showed that if edge-on brightness profiles were transposed, using appropriate radiative transfer models, to show how the profile from a disc observed face-on would look, the abruptness of the truncation would diminish and we would find that the outer disc showed an exponentially falling profile, with a shorter scale length, i.e. a steeper descent, than the profile of the inner disc, but not a sudden cut-off. He supported this with a few measurements of face-on discs. When he moved to our group at the IAC we began a systematic observational study of face-on discs, in which Pohlen collaborated with Peter Erwin and myself. A well-cited summary of this work until 2004, with a comparison of the basic theoretical scenarios for explaining the fall-off observed at the disc edge, can be found in Pohlen et al. [23] in the proceedings of a previous meeting organized by David Block, Ken Freeman, et al. Later these types of measurements were made on greatly extended samples by Pohlen and Trujillo ([24], hereafter PT06) of mainly late-type galaxies and by Erwin, Pohlen, and Beckman ([12], hereafter EPB08) for early-type barred galaxies, confirming that the best description of "truncated" profiles is that of a double exponential with a break between the two at a well-defined radius. This description is also valid for Freeman Type II profiles, though the break radius for these is in the inner disc, while the break point for a truncation is in the outer disc. PT06 and EPB08 suggested that the Type II definition is a useful way to summarize information from all discs with double exponential profiles. In both articles a significant fraction of discs, out to radii well is excess of 5 disc scale lengths, showed no change in profile slope. These were clearly akin to the Freeman Type I's, so that the Type I/ Type II dichotomy could be generalized to profile behaviour over a wide radial range. One may wonder whether, if a Type I profile could be measured out far enough, its slope would also change. However, in some very nearby galaxies single exponential profiles have been found to very large radii using star counts (not affected by sky subtraction errors), such as NGC 300 [6] out to 10 scale lengths, so the Type I/Type II description is well based. The only major departure from this classification was presented in Erwin, Beckman and Pohlen ([11], hereafter EBP05), who found that in the sample of

early-type barred galaxies a significant fraction showed outer disc exponentials with shallower slopes than those of the inner disc profiles. They termed this phenomenon antitruncation, and to extend the Type I/Type II description they termed the discs with antitruncations Type III.

2 A Detailed Radial Profile Classification

In EPB08 we presented a detailed radial profile classification for discs which can include all presently known varieties. This classification was anticipated (with acknowledgement to EPB) in PT06. Type I profiles are simple exponentials and need no sub-classification. Type II profiles have inner shallow exponentials and steeper outer exponentials, with a well-defined "break" radius at the transition. Barred galaxies then need a sub-classification: if the break radius is smaller than the bar length we term these profiles Type II-i ("inner"), and if the break radius is greater than the bar length they are Type II-o ("outer"). These latter are then sub-divided according to whether the break radius coincides with that of an observed outer ring, implying that the break is associated with the outer Lindblad resonance of the bar, so we use the term Type II-o-OLR. Where there is no such association and we cannot yet establish a physical basis for the change in the profile gradient we term these profiles Type II-o-CT (for "classical truncation"). The Type II-o profiles in unbarred galaxies are in general Type II-o-CT. PT06 also mention Type II-AB profiles, in which the break is due to an asymmetry in the outer disc, possibly due to an interaction or merger. We term those profiles with antitruncations Type III. These profiles have two modes, one in which the break occurs at a well-defined radius and one in which the slope starts to get shallower at a given galactocentric radius and gets progressively shallower out to the background limit. In the former case two-dimensional ellipse fits to the isophotes of the disc image show no ellipticity change as we cross the break radius. In the latter case the ellipticity falls steadily, tending to lower values at larger radii coinciding with the shallowing of the profile slope. In the first type of profiles we are seeing a change in slope due to some change in the disc component of the at the break radius, whereas in the second type the change in slope shows that in the outermost zone of the galaxy a spheroidal halo population of stars dominates the light profile. So we sub-classify the Type III profiles into Type III-d, (for "disc") and Type III-s (for "spheroid"). The full classification is visualized in Fig. 1, taken directly from EPB08.

3 Profiles of Unbarred Early-Type Galaxies and Statistics for a Representative Sample of Hubble Types

PT06 used profiles taken from SDSS images, mainly of late-type galaxies, while EPB08 concentrated on early-type barred galaxies. In order to complete the survey we have recently submitted a paper on early-type unbarred galaxies [15] allowing us

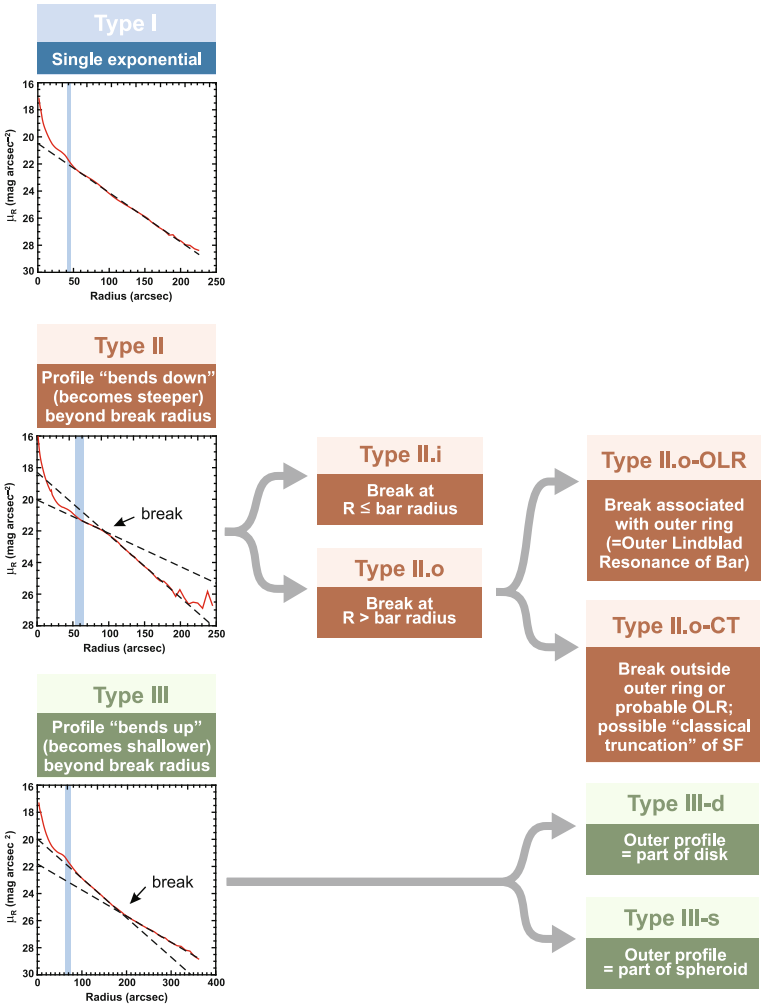


Fig. 1 The classification of radial surface brightness profiles of galaxy discs (From EPB08)

to make a statistical summary of all Hubble types, previewed in the present section. The selection criteria for the unbarred sample, as for our previous barred sample, include all galaxies in the UGC [19] with major axis diameter larger than 2 arcminutes (not too small), with major to minor axis ratios less than 2 (not too inclined), with redshift less than 2,000 km/s (not too distant), with RC3 morphological classification in the range S0 to Sb (not too late type), and with declination above $-10°$ (not too far south to observe from the northern hemisphere). The sample used here is the unbarred subset of the full sample, i.e. those with RC3 classifications of SA or S, plus those deemed by Erwin and Sparke [9] or Erwin [10] to be unbarred, giving a sample size of 49 objects. We observed 24 of them with the Wide Field

camera of the 2.5 m Isaac Newton telescope, La Palma, for 2 we used archive data from the same telescope, for 4 we used archive data from the Jakobus Kapteyn Telescope, La Palma, while 19 were measured using data from the Data Release 5 of the Sloan Digital Sky Survey (SDSS, [1, 36]). The technical details of how the profile is extracted can be found in EPB08. Here we give a summary of some of the key results from Gutiérrez et al. [15] and incorporate these into a statistical summary of all the local galaxies, including the data from PT06 and EPB08. We note that of the 49 initial unbarred galaxies in Gutiérrez et al. [15], 11 were found to be barred and excluded from the analysis.

In Table 1 we show the statistics of the numbers of galaxies with each type of profile in the full sample from the three articles cited. A histogram representation is shown in Fig. 2.

Surprisingly, the biggest category is Type III, though this is matched by the sum of the Type II's. Truncated galaxies are not in fact a majority, as the sum of the Type I's and Type III's is well in excess of the Type II's. Type II's occur preferentially

Table 1 Classification of galaxy profiles in the "unbarred" early-type sample of Gutiérrez et al. [15] (left-hand panel), and in the full sample combining objects from PT06, EPB08 and Gutiérrez et al. [15] (right-hand panel)

	This sample			Full sample		
Profile type	Unbarred	Barred	Total	Unbarred	Barred	Total
I	10	3	13	14	33	47
II–CT	2	1	3	11	30	41
III	24	3	27	34	28	62
II–OLR	0	2	2	0	23	23
II+III	4	2	6	4	9	13
II+II	0	0	0	0	1	1

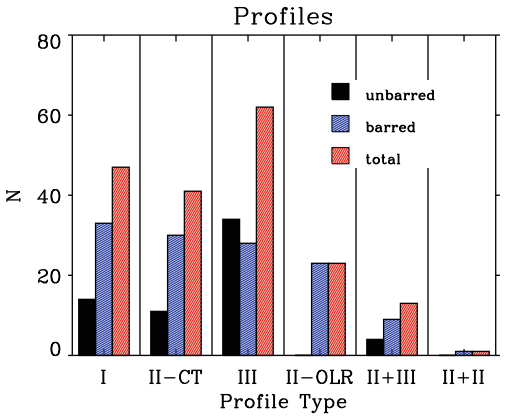


Fig. 2 The distribution of the profile types among all the galaxies of the local sample of 187 objects covering the full range of Hubble types. Each of the sub-samples was complete. Although the combined sample is not necessarily complete it gives a useful guide to the distribution. Types II+III and II+II show more than one clear break in their profiles (see PT06 and EPB08)

in barred galaxies, (this is almost a tautology for Type II-o-OLR) and so do Type I's, whereas the Type III's are evenly divided. This suggests strongly that there are physical differences underlying the classification. There are a few profiles which require classification as a combination of profile types, which are shown in the last two histograms of Fig. 2. In Fig. 3 we show how the profile types are distributed along the Hubble sequence.

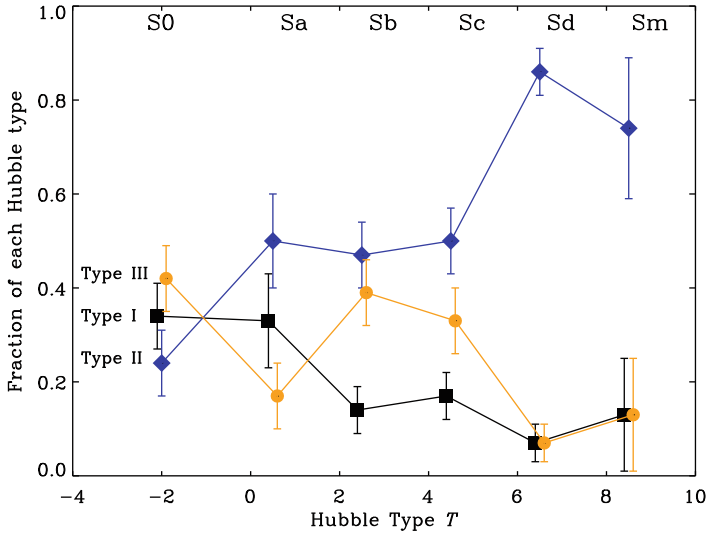


Fig. 3 Fractions of each profile type along the Hubble sequence for the full sample of 187 local galaxies

Truncated discs predominate in the latest types (Sc and especially Sd and Sm), the sum of Type I and Type III outweighs Type II in the earlier types, and at S0 it is Type III which is most representative. In Fig. 4, we plot histograms of the break radius in units of the effective disc radius R_{25}, for each profile type with a break and for barred and unbarred discs.

The distribution of break radii is similar for barred and unbarred galaxies, but the Type III breaks occur at larger radii, a median value of $1.1R_{25}$ compared with $0.85R_{25}$ for the Type II's. Other parameters of interest are the profile scale lengths. Type I's have a single scale length, whereas Type II's and Type III's have inner and outer scale lengths; in the former case the outer scale lengths are shorter and in the latter case they are longer. In Fig. 5 we show the distributions of the scale lengths, using the Type I scale lengths as units.

We can see that the internal scale lengths for Type III discs are in a similar range to those of Type Is, whereas the Type II discs have longer internal scale lengths. The external scale lengths of the Type IIs are comparable to the Type I scale lengths, while the external scale lengths of the Type IIIs are significantly bigger. These results give hints on the types of processes which cause the breaks, but we will leave discussion of this to a later section.

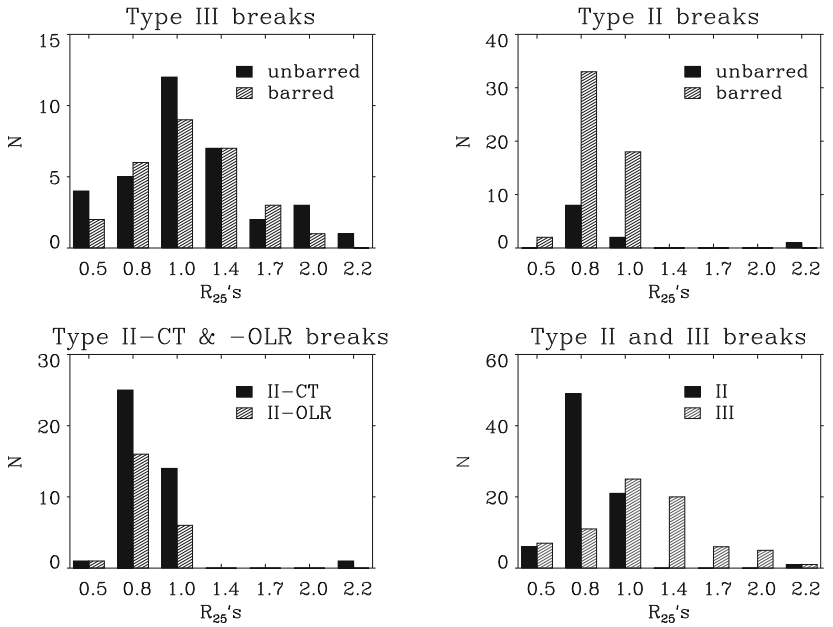


Fig. 4 *Upper panels*: The break radius in units of R_{25} for Type III and Type II profiles. The Type III breaks tend to occur at larger radii as shown. *Lower panels*: *left*: comparison of break radii for Type II-CT and Type II-OLR profiles, which have very similar distributions. *Right*: Direct comparison of distributions of Type II and Type III break radii

4 The Time Evolution of Disc Truncations

Well-defined break radii could be used as scale markers for evolutionary processes in discs. The first researcher to investigate how truncations have developed over cosmic time was Ken Freeman's student Isabel Pérez, who is at this meeting. In a pioneering paper [21], she showed how to use ACS images from the GOODS survey to measure truncations out to redshift 1, finding clear evidence for truncations at smaller radii than are typical for local galaxies. She noted the difficulties due to cosmological dimming. A comparable result was found by Trujillo and Pohlen [30], with a sample of 21 galaxies from the Hubble Ultra Deep field. They found that truncations in discs at $z = 1$ occur at radii some 25% smaller than those in local galaxies: evidence for inside-out build-up of the stellar disc populations. A major step forward was taken by Azzollini, Trujillo, and Beckman ([2], hereafter ATB08a), who used ACS observations of galaxies from the GOODS-South Field [14] included in the GEMS Survey [25]. The selection criteria imposed by ATB08a on those galaxies were (a) a Sérsic index of <2.5 to separate the disc galaxies, (b) a major/minor axis ratio <2 to avoid inclinations higher than 60°, (c) $M_B < -18.5$ mag (the criterion also adopted in [30]) to avoid signal-to-noise problems. The final sample whose profiles were measured and analysed was of 505 objects. Images were chosen through the V-band filter for galaxies in the range $0 < z < 0.5$, through the

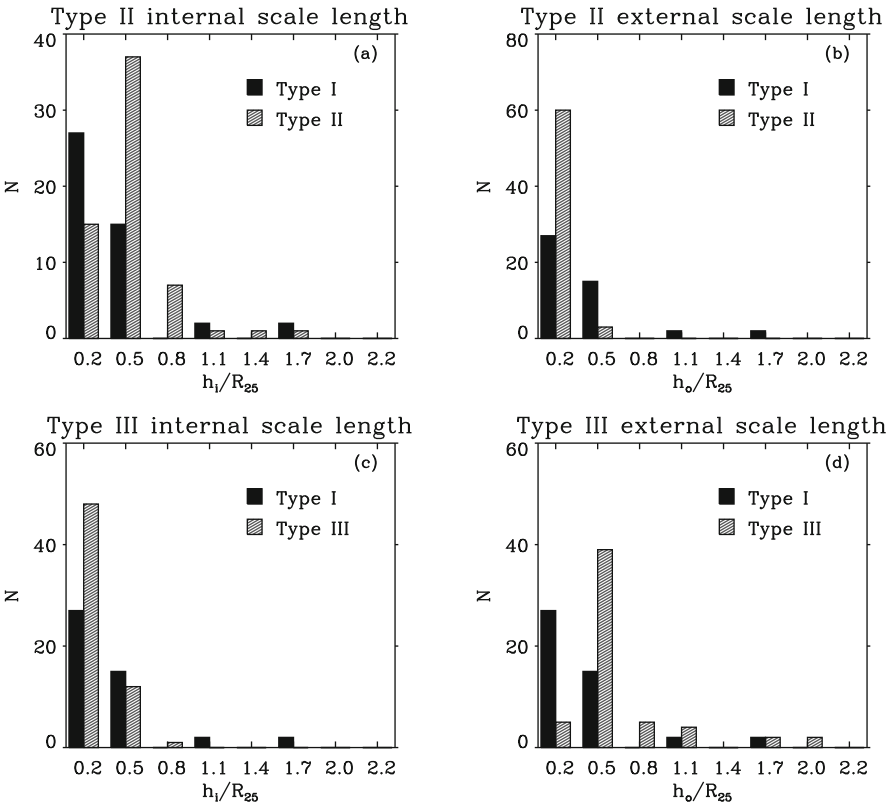


Fig. 5 Scale lengths, h, in units of R_{25} for the inner and the outer discs of Type II and Type III profiles for 187 galaxies. Measurements are based on R-band images

I-band filter for $0.5 < z < 0.8$, and through the Z-band for $0.8 < z < 1.1$, in to harmonize the rest-frame band. The profiles produced were reliable, as seen from the comparison, in Fig. 6. of 2 profiles obtained, from GOODS images, and of the same galaxies from the Hubble Ultra Deep field.

In the 15 galaxies common to this survey and to Trujillo and Pohlen [30] the break radii were equal within the error limits. The profiles were classified into Types I, II, and III, without finer detail, and the results are summarized in Table 2.

They show a slight predominance of Type II over the sum of Types I and III. We also note that compared with the local sample Type III profiles are less frequent, perhaps because cosmological dimming makes it harder to measure the outermost discs, but all three profile types are well represented. The key result of this chapter was the clear demonstration that the break radius was smaller at higher redshift, as shown in Fig. 7, where the results are means for each redshift bin, normalized to the break radius in local galaxies from TP06.

Plots are for galaxies of equal luminosity, M_B, in the top panel, and for galaxies of equal mass in the bottom panel. Using the mass we compare more closely like

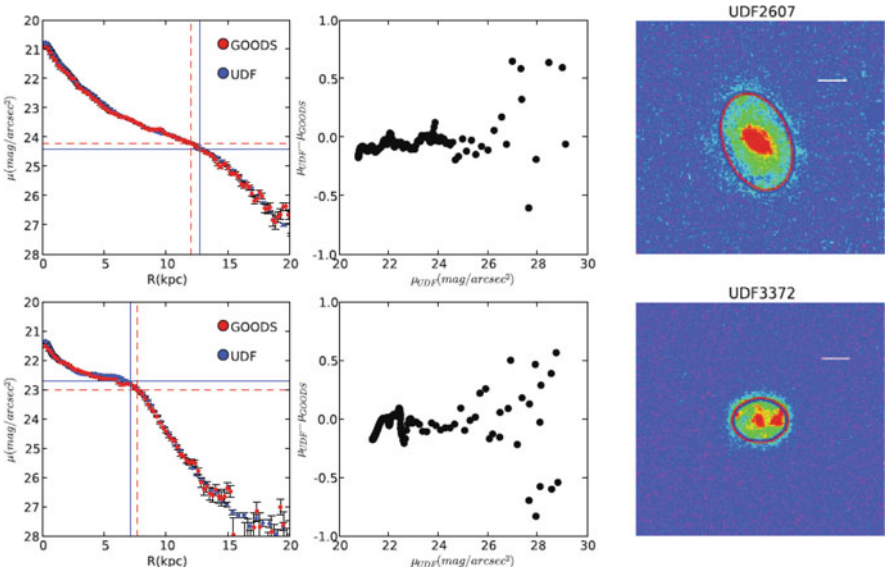


Fig. 6 Comparison of the profiles of two galaxies using data from the GOODS survey and from the Hubble Ultra Deep field, showing the consistency of the profiles out to 26 mag/arcsec^2

Table 2 Classification of the surface brightness profiles in the three redshift bins reported in ATB08a

	Results on classification of profiles					
Redshift	I	II	III	Total[a]	Discarded	Net total[b]
$0.1 < z_{0.5}$	17 (25%)	40 (59%)	10 (15%)	81	14	67
$0.5 < z_{0.8}$	77 (33%)	135 (58%)	22 (9%)	273	39	234
$0.8 < z_{1.1}$	52 (39%)	67 (50%)	15 (11%)	151	17	134
Sum	146	242	47	505	70	435

[a] Number of objects surveyed at each redshift range
[b] Number of objects analysed at each redshift range after rejection of those not well suited to our study

with like, since the enhanced SFR at higher z affects the luminosity distribution. The higher z galaxies are much more luminous, as seen by the plot in Fig. 8 (top) of the surface luminosity at the break radius vs. redshift. There is a decline by 3.2 mag/arcsec, a factor of over 20 in this parameter from z = 1 to the present epoch. A qualitative sketch of the disc profile change with redshift for truncated galaxies is given in Fig. 8 (bottom).

Important clues to the physics underlying the profile behaviour described above can be found in Azzollini, Trujillo and Beckman ([3], hereafter ATB08b), an analysis of the colour profiles of the truncated galaxies, obtained by measuring radial surface brightness curves in restframe B−V for the sample of galaxies from z = 0.1 to z = 1.1 used in ATB08a . A summary of the colour profiles is given in Fig. 9.

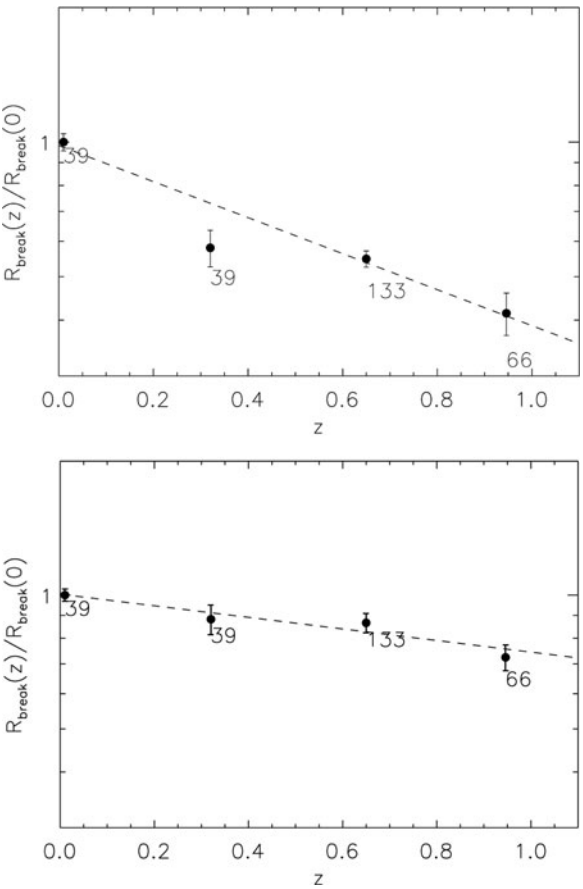


Fig. 7 Break radius vs. redshift for the Type II galaxies in ATB08a. *Upper panel*: for a sample of galaxies in a fixed luminosity bin. *Bottom panel*: for a sample of galaxies in a fixed mass bin

The Type II profiles are the most instructive. They are bluest close to the break radius, in all three redshift bins. ATB2008b showed that the minima are statistically significant at the 3σ level in each bin. The results for Type I profiles do not show this effect. They redden radially outwards, notably in the highest redshift bin, while the Type III profiles are bluer in the outer disc for the nearer redshifts but redder in the outer disc for the furthest bin. But the numbers of Type I's are quite small in the two nearer redshift bins, and of Type IIIs are small in all three bins, so more observations are needed. The complemetary zero redshift data set is found in Bakos, Trujillo, and Pohlen [4] who produced colour profiles of 85 local galaxies from the data of PT06. Their results may be summarized as follows: Type II profiles are bluest at the break radius, Type I profiles show constant colour radially in the disc, and Type III profiles, while redder than Type II's, also show constant radial colour. Bakos et al. [4] used the colours to transform the radial brightness profiles into mass

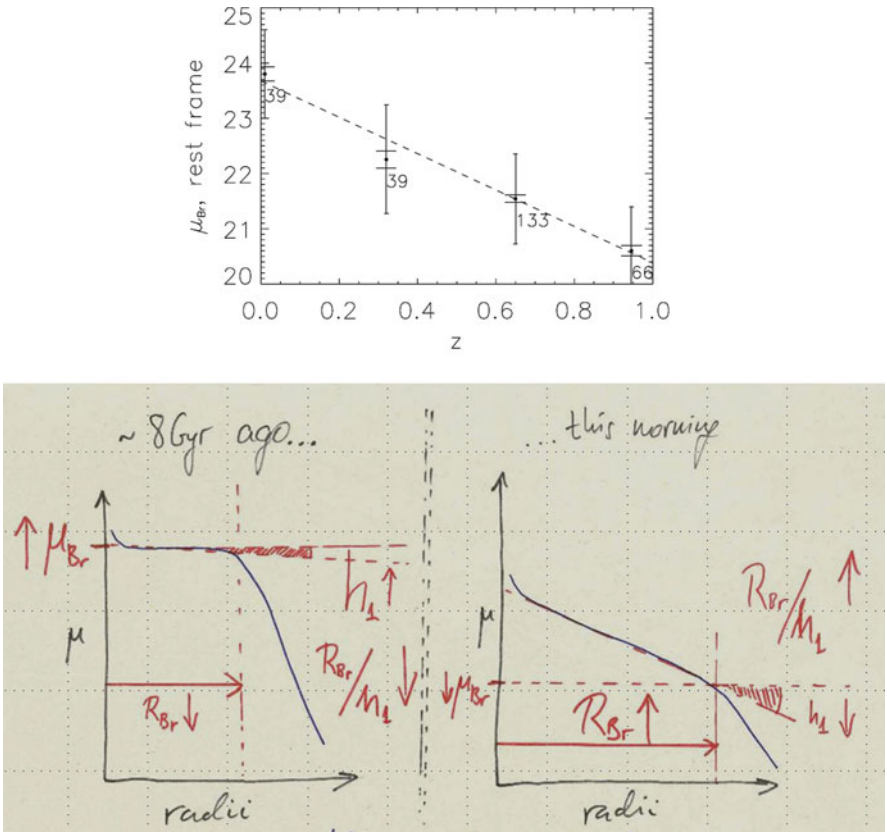


Fig. 8 *Upper panel*: surface brightness of the disc at the break radius as a function of redshift. *Bottom panel*: sketch showing qualitatively how a typical truncated profile has changed from redshift 1 to the present day

surface density profiles, finding that the breaks in both Type II and Type III profiles are smoothed out, though they do not (as hinted by [4]) disappear. The convexity of the Type II profiles and the concavity of the Type III profiles remain.

5 Theoretical Models and Scenarios for Outer Disc Profile Evolution

Two types of mechanisms were originally proposed to account for disc truncations: the effect of momentum conservation in the protogalactic cloud [33] or of star formation thresholds [7, 17, 28]. Elmegreen and Hunter [8] used a threshold model yielding a double exponential brightness profile and forming Type II or Type III galaxies. Magnetic fields have also been invoked to explain truncations [5], while

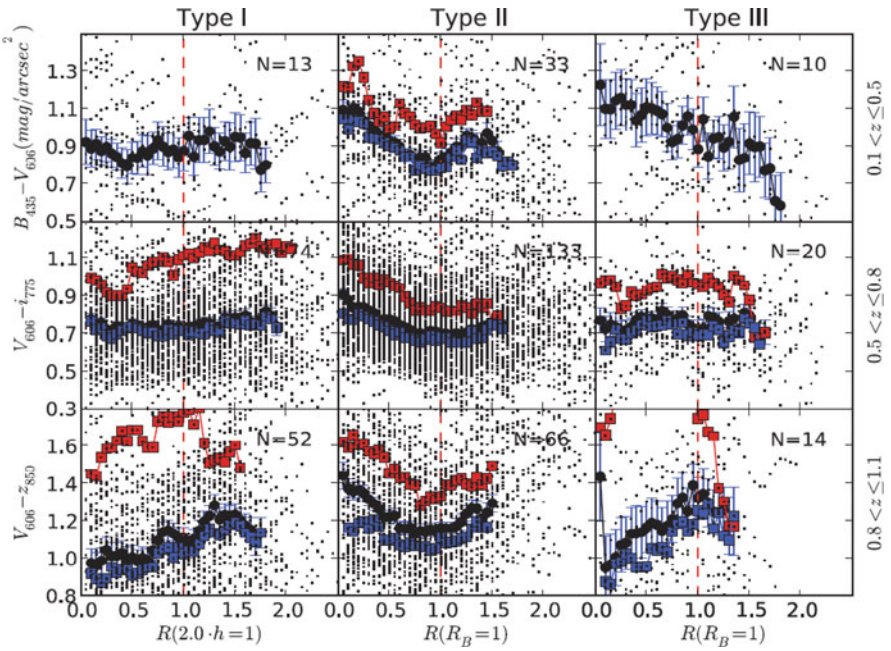


Fig. 9 Colour profiles of the galaxies studied in ATB08b. The *upper* 3 *panels* give data in the nearest redshift bin, the *middle panels* in the intermediate redshift bin and the *lower panels* in the furtherst bin (see ATB08b for full details)

Type III morphology has been explained by tidal stripping within a minor merger [20, 34], by the bombardment of the disc with dark matter subhalos [16] or by a high eccentricity flyby of a satellite galaxy [35]. Martínez-Valpuesta [18] has shown that the dynamical effects of a bar within a massive halo can lead to Type III profiles without recourse to interactions or mergers. I will pay most attention here to work by Roskar et al. [26, 27] because this relates directly to the evolution of the profiles and can be compared with both current and high redshift profiles. Roskar et al. [26] first discuss the effects of Galactic stellar migration on the metallicity distribution in the solar neighbourhood and on the metallicity gradient in the Galactic disc. Migration causes a number of effects which had not been taken considered in previous studies of the chemical evolution of the Galaxy, and in particular of the age–metallicity relationship. They show that the time evolution of the star formation rate in the Galaxy disc (and in other galaxy discs) depends on the Galactocentric radius. In the inner disc the rate is always highest, but has fallen steadily with time, whereas in the outer disc although the rate has been lower, it has peaked more recently and progressively more recently at increasing radius. Thus the average age of the stellar population should fall with increasing radius, out to a specific limit. In Roskar et al. [27] these considerations are applied to predict the evolution of azimuthally averaged light profiles, those whose observational parameters we have reviewed. Their predictions

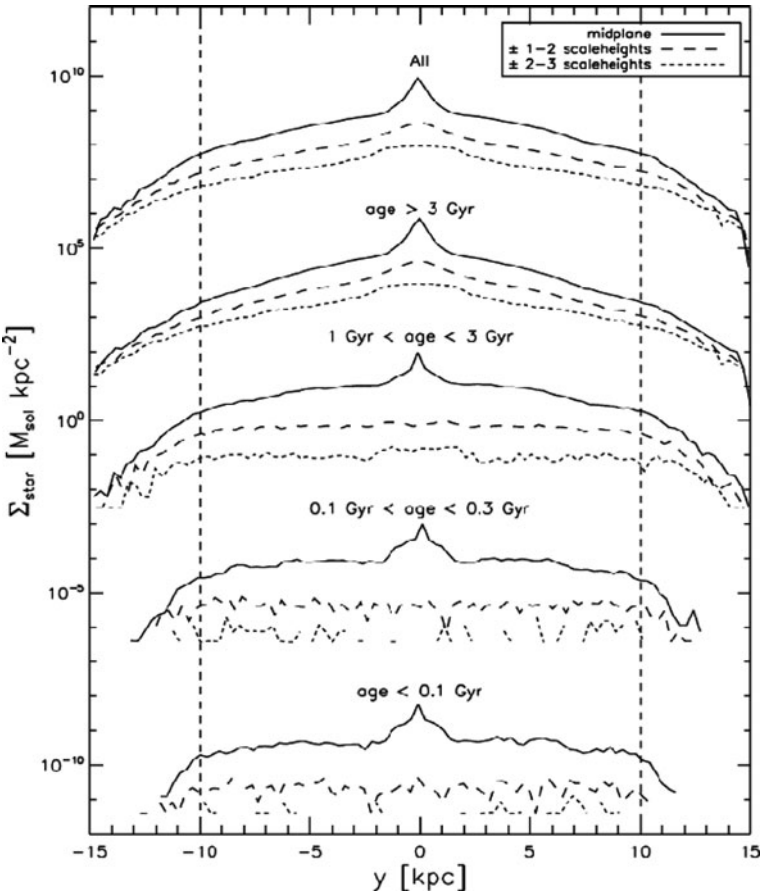


Fig. 10 Distribution of stars in different age ranges across a model galaxy disc from Roskar et al. [27], illustrating that their stellar migration models imply "inside-out" disc formation

can be summarized in two figures: Fig. 10, which gives the distribution of stars in different age ranges across the face of a galaxy disc at the present epoch. There is a break, where star formation fades away, such that the stars at larger radii have all diffused outwards from their birth positions. The model implies that the youngest population should be found at the break radius. Two effects operate to produce the disc profiles: the redistribution of the stellar population due to dynamical asymmetries (bars and spiral arms) and the continuous infall of fresh gas to the disc, which allows star formation to persist throughout its lifetime. The time evolution of the profiles is shown in Fig. 11.

We can see that this model predicts the existence of profile breaks, that the break should migrate outwards with epoch, and that the youngest (i.e. the bluest) mean stellar population should be at the break radius. These are all properties of the observed Type II profiles and their time development, and we note that Roskar et al.

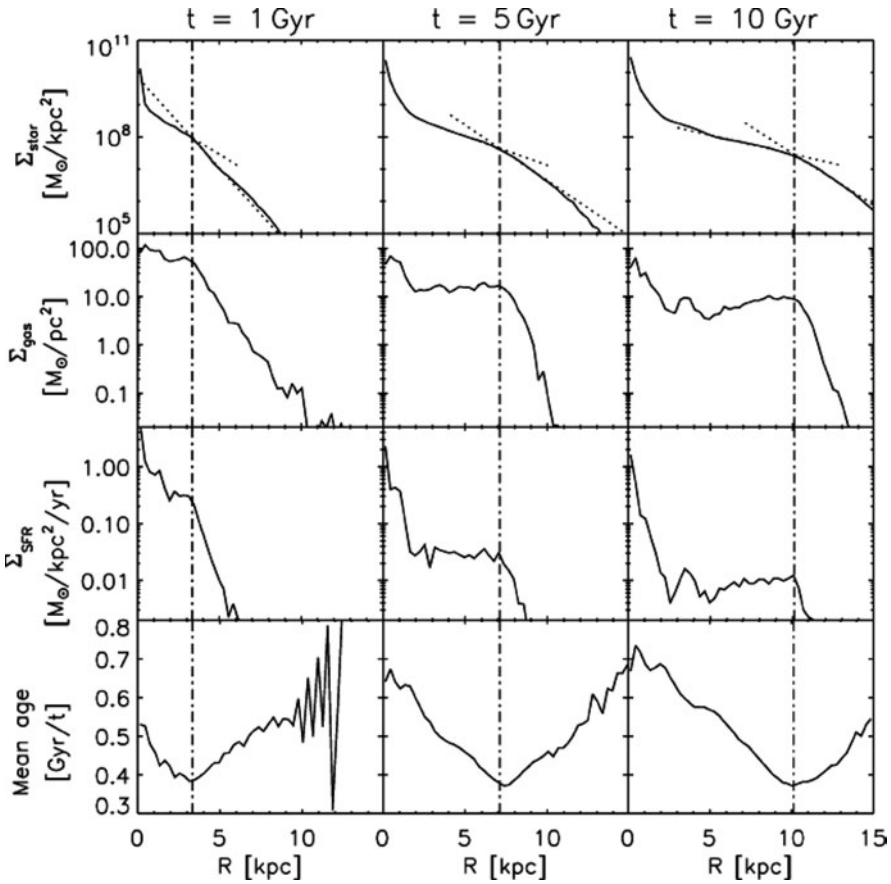


Fig. 11 Azimuthally averaged disc properties after 1, 5, and 10 Gyr in the models of Roskar et al. [27]. *Upper panel*: stellar surface density profiles, with *dotted lines* showing double exponential fits; their intersection gives the break radius. *Second panel*: surface density of cool gas. *Third panel*: SFR density. *Fourth panel*: stellar age

[26, 27] were written before ATB08a, ATB08b or Bakos et al. [4] were available. In these models discs which are more closely symmetric could yield Type I profiles. However, we know (Fig. 2) that the majority of Type I profiles are in fact found in barred galaxies (roughly in proportion to the fraction of barred galaxies in the full disc galaxy population), so that there is clearly work to be done here. Finally, we also note that no mention is made in Roskar [27] of Type III profiles. However, Martínez-Valpuesta et al. [18] have shown that in simulation models of bar evolution in dynamical models of galaxies with angular momentum exchange among the disc, the bar, and the dark halo, for cases with weak bars Type III profiles with properties comparable to those observed can be naturally produced. An example is shown in Fig. 12.

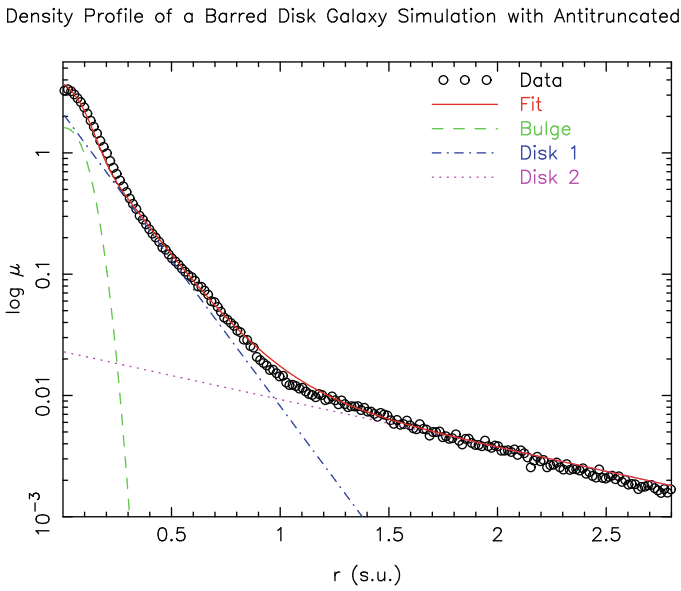


Fig. 12 Simulation of a disc profile in a weakly barred galaxy, with a live dark halo, from Martínez-Valpuesta et al. [18]. The break radius and the ratio of inner to outer disc scale lengths are in the range of observed parameters for Type III profiles in EBP05, PT06, and EPB08

6 Conclusions

As techniques have improved, deeper and more linear imaging has been used to better define the surface brightness profiles of outer galactic discs. One of the major worries, ensuring good sky subtraction for the low values of surface brightness beyond a few disc scale lengths, appears to be under control, as profiles measured with different instrumentation give results which agree within their errors down to quite low brightnesses. This enables us to measure face-on discs, where problems of extinction in the projected outer disc are greatly reduced. As a result we can use outer disc profiles as a practical tool to explore galaxy structure and evolution. Some of the results are presented here, and it is encouraging that the theorists are producing models which predict observations without excessive fine-tuning. But more observations, notably in the IR, for reliable exploration of the mass distribution, and in the UV and Hα, to follow recent star formation, are needed for a fuller picture on which the theorists can work. In fact, there have been very interesting advances in UV observations which, for length limitation, I cannot discuss here. I refer you to Thilker et al. [29] for a survey of outer discs in the UV with the GALEX satellite. It is perhaps ironic that we are in better shape with models which explain truncated profiles than with scenarios which predict single slope exponentials or antitruncations. These all add interest to a subject which was founded by two of the protagonists at this meeting (Ken Freeman and Piet van der Kruit), more than 30 years ago, but is still a live and developing field.

Acknowledgments The authors thank Michael Pohlen, Nacho Trujillo, Rebeca Aladro, Judit Bakos, and Armando Gil de Paz for collaborations, ideas, discussions, and other invaluable contributions to much of the work presented here.

References

1. Adelman-McCarthy, J.K., et al. 2007, ApJS, 172, 634
2. Azzollini, R., Trujillo, I., Beckman, J.E. 2008a, ApJ, 684, 1026 (ATB08a)
3. Azzollini, R., Trujillo, I., Beckman, J.E. 2008b, ApJ, 679, L69 (ATB08b)
4. Bakos, J., Trujillo, I., Pohlen, M. 2008, ApJ, 683, L103
5. Battaner, E., Florido, E., Jimenez-Vicente, J. 2002, A&A, 388, 213
6. Bland-Hawthorn, J., Vlajic, M., Freeman, K.C., Draine, B.T. 2005, ApJ, 629, 239
7. Elmegreen, B.G., Parravano, A. 1994, ApJ, 435, L121
8. Elmegreen, B.G., Hunter, D.A. 2006, ApJ, 636, 712
9. Erwin, P., Sparke, L.S. 2003, ApJS, 146, 299
10. Erwin, P. 2005, MNRAS, 364, 283
11. Erwin, P., Beckman, J.E., Pohlen, M. 2005, ApJ, 626, L81 (EBP05)
12. Erwin, P., Pohlen, M., Beckman, J.E. 2008, AJ, 135, 20 (EPB08)
13. Freeman, K.C. 1970, ApJ, 160, 811
14. Giavalisco, M., et al. 2004, ApJ, 600, L93
15. Gutiérrez, L., Erwin, P.E., Aladro, R., Beckman, J.E. 2010, AJ, Submitted
16. Kazantzidis, S., Bullock, J.S., Zentner, A.R., Kravtsov, A.V., Moustakas, L.A. 2008, ApJ, 688, 254
17. Kennicutt, R.C. 1989, ApJ, 344, 685
18. Martínez-Valpuesta, I., et al., 2010. In preparation
19. Nilson, P. 1973, Uppsala General Catalog of Galaxies, Uppsala Astron Obs Ann, 5, 1
20. Peñarrubia, J., McConnachie, A., Babul, A. 2006, ApJ, 650, L33
21. Pérez, I. 2004, A&A, 427, L17
22. Pohlen, M., Dettmar, R.-J., Lütticke, R., Aronica, G. 2002, A&A, 392, 807
23. Pohlen, M., Beckman, J.E., Hüttemeister, S., Knapen, J.H., Erwin, P., Dettmar, R.-J. 2004, In: Penetrating Bars Through Masks of Cosmic Dust: The Hubble Tuning Fork Strikes a New Note, eds. D.L. Block, I. Puerari, K.C. Freeman, R. Groess, E. Block, Kluwer, Dordrecht, p. 713
24. Pohlen, M., Trujillo, I. 2006, A&A, 454, 759 (PT06)
25. Rix, H.-W., et al. 2004, ApJS, 152, 163
26. Roskar, R., Debattista, V.P., Quinn, T.R., Stinson, G.S., Wadsley, J. 2008a, ApJ, 684, L79
27. Roskar, R., Debattista, V.P., Stinson, G.S., Quinn, T.R., Kaufmann, T., Wadsley, J. 2008b, ApJ, 675, L65
28. Schaye, J. 2004, ApJ, 609, 667
29. Thilker, D.A., et al. 2007, ApJS, 173, 538
30. Trujillo, I., Pohlen, M. 2005, ApJ, 630, L17
31. van der Kruit, P., Searle, L. 1981, A&A, 95, 105
32. van der Kruit, P., Searle, L. 1981, A&A, 95, 116
33. van der Kruit, P., 1987, A&A, 173, 59
34. Younger, J.D., Cox, T.J., Seth, A.C., Hernquist, L. 2007, ApJ, 670, 269
35. Younger, J.D., Besla, G., Cox, T.J., Hernquist, L., Robertson, B., Willman, B. 2008, ApJ, 676, L21
36. York, D.G., et al. 2000, AJ, 120, 1579

The Complex Interplay of Dust and Star Light in Spiral Galaxy Discs

Maarten Baes, Dimitri Gadotti, Joris Verstappen, Ilse De Looze, Jacopo Fritz, Edgardo Vidal Pérez, and Marko Stalevski

Abstract Interstellar dust grains efficiently absorb and scatter UV and optical radiation in galaxies and therefore can significantly affect the apparent structure of spiral galaxies. We discuss the effect of dust attenuation on the observed structural properties of bulges and discs. We also present some first results on modelling the dust content of edge-on spiral galaxies using both optical and Herschel far-infrared data. Both of these results demonstrate the complex interplay of dust and star light in spiral galaxies.

1 Introduction

In has been known for a long time that interstellar dust grains are an important component of the interstellar medium in galaxies: they efficiently absorb and scatter UV and optical radiation, play an important role in the chemistry of the ISM, and are the dominant source of far-infrared and submillimetre emission. Detailed knowledge of the quantity, spatial distribution, and physical properties of the dust in spiral galaxies is still controversial in many ways. For several decades after the work by Holmberg [21], it was assumed that spiral galaxies were optically thin for optical radiation. This conventional viewpoint was questioned in the late 1980s and early 1990s by several teams, including Disney et al. [13], Valentijn [18] and Burstein et al. [34]. These authors came to the conclusion that spiral galaxies are optically thick, even in the outer regions. Around the same time, other teams reached completely different conclusions, sometimes on the basis of identical data sets (e.g. [22]).

In retrospective, three issues conspired to complicate the discussion on the optical thickness of spiral galaxies. The first was observational biases in the classical tests, such as the expected variation of isophotal diameter, mean surface brightness with increasing inclination. It is very difficult to quantify these biases, let alone to circumvent them (e.g. [17]). A second issue that complicated the discussion was the unavailability of reliable far-infrared observations that traced the bulk of the dust

M. Baes (✉)
Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium
e-mail: maarten.baes@ugent.be

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_16,

mass in spiral galaxies. In particular, almost no far-infrared or submillimetre data were available at wavelengths longer than 100 μm, the longest IRAS wavelength. The final and possibly the most important reason why the contrary results were obtained was the simplicity of the models used to analyse or interpret the data. This was convincingly demonstrated by Disney et al. [18], who showed that simple optically thick models could reproduce the same observations on the basis of which Holmberg [21] had reached his conclusions.

One of the first to realize that realistic models for dusty galaxies were absolutely needed to investigate the dust content of spiral galaxies was Ken Freeman. He and his Ph.D. student Yong-Ik Byun embarked on the first systematic effort to investigate the complex interplay of dust and star light in realistic models of spiral galaxies, using detailed radiation transfer simulations including both absorption and scattering. The result of this seminal work, presented in Byun et al. [14], is still a monument in the extragalactic radiative transfer community.

In the past 15 years, the modelling of the interplay between dust and star light in spiral galaxies has been dealt with by many authors. In particular, several groups have developed the necessary numerical codes to solve the radiative transfer problem in extragalactic environments. The most advanced of these codes are not restricted to only absorption and scattering in 2D geometries, but take into account thermal emission by dust, polarization, kinematics, and multi-phase dust distributions (e.g. [11, 20, 25, 28, 32]). Our group has also developed a 3D radiative transfer code, SKIRT, based on the Monte Carlo method. Its original aim was to investigate the effects of dust absorption and scattering on the observed kinematics of elliptical galaxies [2, 5, 6]. It has now developed into a mature radiative transfer code that can be used to simulate images, spectral energy distributions, kinematics, and temperature maps of dusty systems, ranging from circumstellar discs to AGNs (e.g. [31, 35]). In particular, the code has been used as the main tool for a detailed investigation of radiative transfer in spiral galaxies [3, 4, 7]. In the remainder of this chapter we shortly describe two recent results that demonstrate the complex interplay of dust and star light in spiral galaxies. In Section 2 we discuss the effect of dust attenuation on the observed structural properties of bulges and discs, and in Section 3 we present some first results on modelling the dust content of edge-on spiral galaxies using both optical and far-infrared data.

2 Dust Effects on Bulge and Disc Parameters

It has been known for many decades that the presence of dust influences the observed, apparent photometric galaxy parameters (apparent scalelengths, surface brightnesses, luminosities, axial ratios, etc.) and makes it a challenge to recover the intrinsic unaffected parameters. Several authors have investigated these effects using radiative transfer modelling with varying degrees of sophistication and/or geometrical realism. In general, it was found that the importance of dust attenuation varies as a function of wavelength, galaxy inclination, and star–dust geometry. In particular,

Byun et al. [14] were the first to convincingly demonstrate that effects of scattering are often counterintuitive and crucial to properly interpret the effects of dust.

In the last few years, two independent teams have investigated the effects of dust attenuation in bulge and disc components, on their integrated properties, separately, using realistic models of spiral galaxies. Both Pierini et al. [26] and Tuffs et al. [33] presented attenuation functions for the individual disc and bulge components of dusty spiral galaxies. They clearly demonstrated that the effects of dust on the bulge and disc components can differ substantially, as a result of the different star–dust geometry. We have aimed to extend this work one stage further. We have embarked on a project to investigate the systematic effects of dust attenuation on the apparent detailed structural properties of discs and bulges simultaneously. We have created artificial galaxy images, using radiative transfer simulations, to mimic the observed structural properties of disc galaxies with classical and pseudo-bulges and include the effects of dust attenuation in the observed light distribution. By applying 2D bulge/disc decomposition techniques in this set of models, we were able to evaluate what are the effects of galaxy inclination and dust opacity on the results from such decompositions.

Rather surprisingly, we have found that the effects of dust on the structural parameters of bulges and discs obtained from 2D bulge/disc decomposition cannot be simply evaluated by putting together the effects of dust on the properties of bulges and discs treated separately. In particular, the effects of dust in galaxies hosting pseudo-bulges might be different from those in galaxies hosting classical bulges, even if their dust content is identical. Confirming previous results, we find that disc scale lengths are overestimated when dust effects are important. In addition, we also find that bulge effective radii and Sérsic indices are underestimated. Furthermore, the apparent attenuation of the integrated disc light is underestimated, whereas the corresponding attenuation of bulge light is overestimated. Dust effects are more significant for the bulge parameters and, combined, they lead to a strong underestimation of the bulge-to-disc ratio, which can reach a factor of 2 in the V band, even at relatively low galaxy inclinations and dust opacities (see Fig. 1). The reason for these, at first sight, counterintuitive results comes from the fact that such decompositions use specific models to fit bulges and discs which cannot accommodate the effects of a dust disc in the galaxy. Therefore, when the model for a component tries to adjust itself when dust is present, this has direct consequences on the model of the other component, even if the latter is not directly affected by dust. More details can be found in Gadotti et al. [19].

3 The Energy Balance in Spiral Galaxies

Apart from more advanced modelling techniques, one of the crucial steps forward. (compared to the early 1990s) in determining the dust content of spiral galaxies is the availability of observations in the far-infrared range between 100 and 1000 μm. Observations in this wavelength range are crucial because they directly trace the

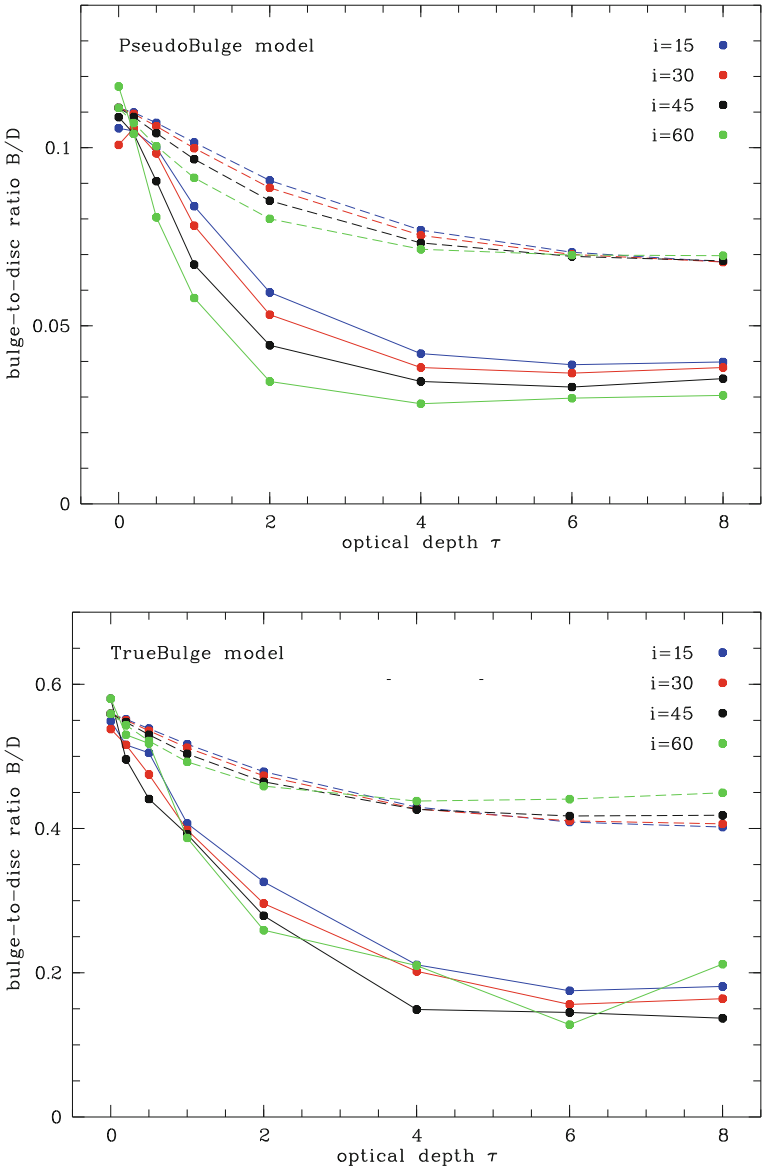


Fig. 1 Dependence of the bulge-to-disc ratio on the V-band optical depth τ. The *solid lines* represent the apparent bulge-to-disc ratio as derived from the BUDDA bulge/disc decompositions of the dust-affected images. The *dashed lines* represent the actual bulge-to-disc ratio as determined from the ratio of the input bulge and disc integrated fluxes

smoothly distributed cold dust in spiral galaxies, which is heated to temperatures of some 20 K by the general interstellar radiation field. While the ISO and Spitzer missions have opened up the far-infrared window outto 200 μm, the most important

step forward was the launch of Herschel in May 2009. The first results of Herschel imaging of nearby galaxies such as M81, M82, or NGC 4438 are spectacular (e.g. [9, 15, 29]).

We are focusing our attention to edge-on spiral galaxies, because they are an important class of galaxies in which the distribution and properties of interstellar dust grains can be studied in great detail. On the one hand, the dust in these systems shows prominently as dust lanes in optical images; on the other hand, surface brightness enhancements ensure that the far-infrared dust emission in edge-on spirals can be traced out to large radii. A self-consistent treatment of extinction and thermal emission, i.e. a study of the dust energy balance, gives the strongest constraints on the dust content of spiral galaxies.

For several edge-on spiral galaxies, the dust distribution has been modelled by fitting realistic radiative transfer models to such optical images [1, 10, 23, 36–38]. The conclusion of these works is that, in general, the dust disc is thinner (vertically) but radially more extended than the stellar disc and that the central optical depth perpendicular to the disc is less than one in optical wavebands, making the disc almost transparent when seen face-on. This result seems to be at odds with FIR per submillimetre emission studies, which indicate that spiral galaxies typically reprocess about 30% of the UV/optical radiation [27]. When applied to individual edge-on spiral galaxies, it is found that the predicted FIR fluxes of self-consistent radiative transfer models that successfully explain the optical extinction generally underestimate the observed FIR fluxes by a factor of about 3 [1, 16, 24, 28]. Several scenarios have been proposed to explain this discrepancy, but a major problem discriminating between these is that the number of edge-on galaxies for which such detailed studies have been done so far is limited, due to the poor sensitivity, spatial resolution, and limited wavelength coverage of the available FIR instruments.

We have used Herschel PACS and SPIRE observations of the edge-on spiral galaxy UGC 4754 to investigate its dust energy balance (for more details, see [8]). We build detailed SKIRT radiative models based on SDSS and UKIDSS maps and use these models to predict the far-infrared emission. We find that our radiative transfer model underestimates the observed FIR emission by a factor 2–3 (see Fig. 2). Similar discrepancies have been found for other edge-on spiral galaxies based on IRAS, ISO, and SCUBA data. Thanks to the good sampling of the SED at FIR wavelengths, we can rule out an underestimation of the FIR emissivity as the cause for this discrepancy. We argue that the most likely explanation for this energy balance problem is that a sizable fraction of the FIR per submillimetre emission arises from additional dust that has a negligible extinction on the bulk of the starlight, such as young stars deeply embedded in dusty molecular clouds. The presence of compact dust clumps can boost the FIR per submillimetre emission of the dust while keeping the extinction relatively unaltered (e.g. [11, 12, 30]). An indication that embedded star-forming clouds might be the solution to the case of UGC 4574 is that the discrepancy between our radiative transfer model and the observed FIR SED is stronger at shorter than at longer wavelengths. This implies that warmer dust (such as in star-forming regions) is necessary to bring the model in balance with the data.

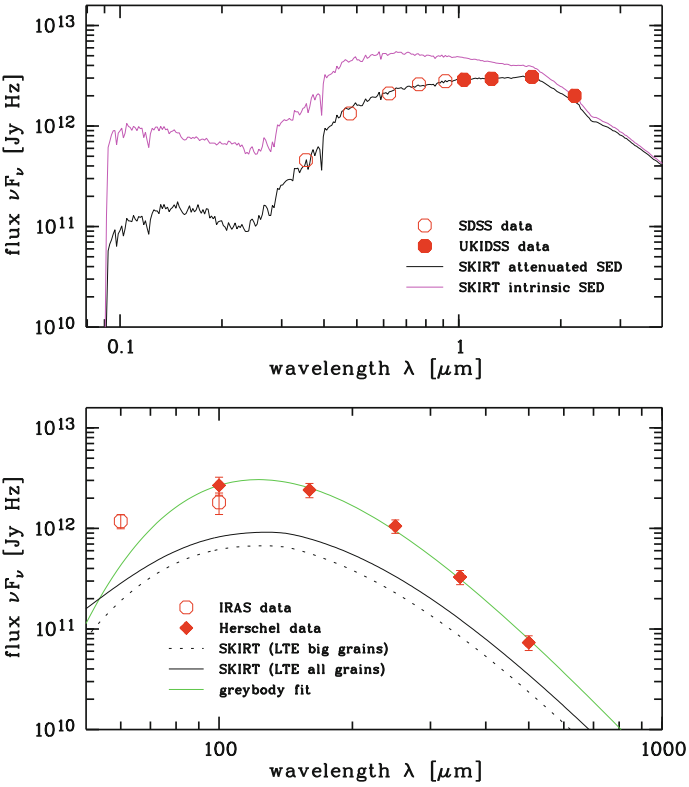


Fig. 2 The optical/NIR (*top*) and FIR per submillimetre (*bottom*) spectral energy distribution of UGC 4754. The *solid black line* in the *top panel* corresponds to the attenuated SED of the SKIRT model fitted to the SDSS and UKIDSS images, the *magenta line* is the unattenuated SED. In the *bottom panel*, the *solid black line* corresponds to the FIR emission of the model assuming LTE for all grains; the *dotted line* represents the contribution of the large grains only. Our model significantly underestimates the observed IRAS and Herschel fluxes

References

1. Alton, P.B., Xilouris, E.M., Misiriotis, A., et al. 2004, A&A, 425, 109
2. Baes, M., Dejonghe, H. 2000, MNRAS, 313, 153
3. Baes, M., Dejonghe, H. 2001, MNRAS, 326, 722
4. Baes, M., Dejonghe, H. 2001, MNRAS, 326, 733
5. Baes, M., Dejonghe, H. 2002, MNRAS, 335, 441
6. Baes, M., Dejonghe, H., De Rijcke, S. 2000, MNRAS, 318, 798
7. Baes, M., Davies, J.I., Dejonghe, H., et al. 2003, MNRAS, 343, 1081
8. Baes, M., Fritz, J., Gadotti, D.A., et al. 2010, A&A, 518, L39
9. Bendo, G.J., Wilson, C.D., Pohlen, M., et al. 2010, A&A, 518, L65
10. Bianchi, S. 2007, A&A, 471, 765
11. Bianchi, S. 2008, A&A, 490, 461
12. Bianchi, S., Davies, J.I., Alton, P.B. 2000, A&A, 359, 65
13. Burstein, D., Haynes, M.P., Faber, M. 1991, Nature, 353, 515

14. Byun, Y.I., Freeman, K.C., Kylafis, N.D. 1994, ApJ, 432, 114
15. Cortese, L., Bendo, G.J., Boselli, A., et al. 2010, A&A, 518, L63
16. Dasyra, K.M., Xilouris, E.M., Misiriotis, A., Kylafis, N.D. 2005, A&A, 437, 447
17. Davies, J.I., Phillips, S., Boyce, P.J., Disney, M.J. 1993, MNRAS, 260, 491
18. Disney, M., Davies, J., Phillipps, S. 1989, MNRAS, 239, 939
19. Gadotti, D.A., Baes, M., Falony, S. 2010, MNRAS, 403, 2053
20. Gordon, K.D., Misselt, K.A., Witt, A.N., Clayton, G.C. 2001, ApJ, 551, 269
21. Holmberg, E. 1958, Meddelanden fran Lunds Astronomiska Observatorium Serie II, 136, 1
22. Huizinga, J.E., van Albada, T.S. 1992, MNRAS, 254, 677
23. Kylafis, N.D., Bahcall, J.N. 1987, ApJ, 317, 637
24. Misiriotis, A., Popescu, C.C., Tuffs, R., Kylafis, N.D. 2001, A&A, 372, 775
25. Misselt, K.A., Gordon, K.D., Clayton, G.C., Wolff, M.J. 2001, ApJ, 551, 277
26. Pierini, D., Gordon, K.D., Witt, A.N., Madsen, G.J. 2004, ApJ, 617, 1022
27. Popescu, C.C., Tuffs, R.J. 2002, MNRAS, 335, L41
28. Popescu, C.C., Misiriotis, A., Kylafis, N.D., et al. 2000, A&A, 362, 138
29. Roussel, H., Wilson, C.D., Vigroux, L., et al. 2010, A&A, 518, L66
30. Silva, L., Granato, G.L., Bressan, A., Danese, L. 1998, ApJ, 509, 103
31. Stalevski, M., Fritz, J., Baes, M., Nakos, Th., Popescu, L.C., 2010, in prep.
32. Steinacker, J., Henning, T., Bacmann, A., Semenov, D. 2003, A&A, 401, 405
33. Tuffs, R.J., Popescu, C.C., Völk, H.J., et al. 2004, A&A, 419, 821
34. Valentijn, E.A. 1990, Nature, 346, 153
35. Vidal, E., Baes, M. 2007, BaltA, 16, 101
36. Xilouris, E.M., Kylafis, N.D., Papamastorakis, J., et al. 1997, A&A, 325, 135
37. Xilouris, E.M., Alton, P.B., Davies, J.I., et al. 1998, A&A, 331, 894
38. Xilouris, E.M., Byun, Y.I., Kylafis, N.D., et al. 1999, A&A, 344, 868

Galaxy Morphology Revealed By SDSS: Blue Elliptical Galaxies

Hong Bae Ann

Abstract The Sloan Digital Sky Survey (SDSS) reveals many new features of galaxy morphologies. Among others, the discovery of blue elliptical galaxies provides some insights into the formation and evolution of galaxies. There seems to be two types of blue elliptical galaxies. One type shows globally blue colors suggesting star formations over the entire galaxy whereas the other type shows blue core that indicates enhanced star formation in the nuclear regions. The former seems to be currently forming galaxies, while the latter is thought to be in transition stage from the blue cloud to the red sequence due to AGN feedback.

1 Introduction

SDSS opens a new era of galaxy studies by providing $ugriz$ images and medium resolution spectra of roughly 1 million galaxies. This enables us to take into account the population information derived from the colors in the morphology classification and to examine the environmental dependence of the physical parameters of galaxies from the scales of galactic satellite systems [1] to the large-scale structures [2, 7].

The discovery of blue elliptical galaxies [10] demonstrates the usefulness of color images in the study of galaxy morphologies. Elliptical galaxies have been thought to have simple structures with red colors. They are dead and boring galaxies with little or no star-forming activities. Blue elliptical galaxies are not. They show a variety of morphologies with active star formations [3, 5, 8, 11].

However, the morphological properties of blue elliptical galaxies are not well understood because of the difficulties to segregate them from other types of galaxies. They are likely to be confused with BCDs in the blue cloud and blue outliers of the red sequence. The purpose of the present study is to examine the properties of blue elliptical galaxies selected from the SDSS DR7 that provides color images and spectra of roughly 1 million galaxies.

H.B. Ann (✉)
Pusan National University, Busan, Korea
e-mail: hbann@pusan.ac.kr

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_17,

2 Morphology and Photometric Properties of SDSS Galaxies

Accurate determination of the morphological types of galaxies is essential for the analysis of the morphology-related properties of galaxies. To do this, we used the morphological types determined by the automated classifier of Park and Choi [6] which is optimized for the SDSS galaxies. It divides galaxies into early types (E/S0) and late types (Sp/Irr) with $\sim 90\%$ accuracy. Moreover, it can detect blue early-type galaxies by using colors and color gradient along with central concentration of light as the criteria of morphology classifications.

Figure 1 shows the $u - r$ versus M_r diagram of SDSS galaxies with redshift $z < 0.05$ where late-type galaxies (Sp/Irr) are plotted in the left panel and early-type galaxies (E/S0) are plotted in the right panel. The distribution of late-type galaxies shows a cloud-like feature while that of early-type galaxies shows a well-defined sequence of red galaxies with two groups of outliers in the left and right sides of the sequence.

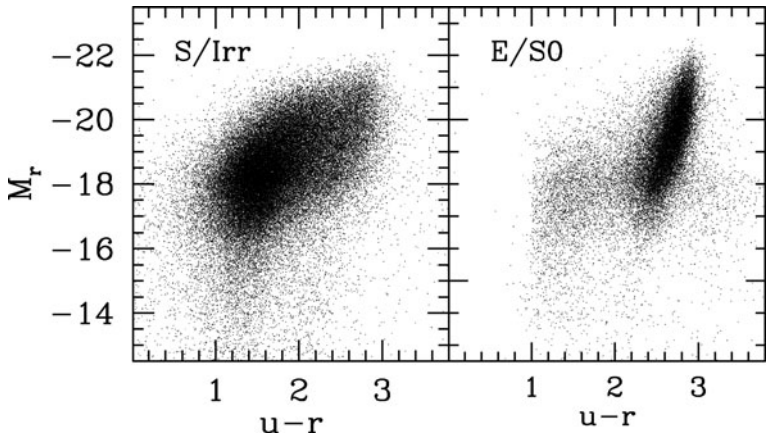


Fig. 1 Color–magnitude diagram of galaxies with redshifts $z < 0.05$

The early-type galaxies in the left side of the red sequence are called as blue elliptical galaxies which are difficult to be distinguished from the galaxies in the blue cloud in the color–magnitude diagrams. The lack of early-type galaxies at $u - r < 1.0$ is due to the color limit of blue elliptical galaxies set by Park and Choi [6] to distinguish them from BCDs. The galaxies in the right side of the red sequence have different origins. The majority of them are S0 galaxies with pronounced dustlanes seen edge-on. Some of the red outliers are merging galaxies where dust obscuration is the main cause of the red colors. There is a family of red outliers whose extremely red colors are not due to dust extinction but due to a large UV dropout. Thus, it is of interest to see the morphological features of the extremely red galaxies as well as the blue elliptical galaxies.

Figure 2 shows the fractional distribution of $u - r$ colors for early- and late-type galaxies. There is a clear difference between the two distributions. The distribution

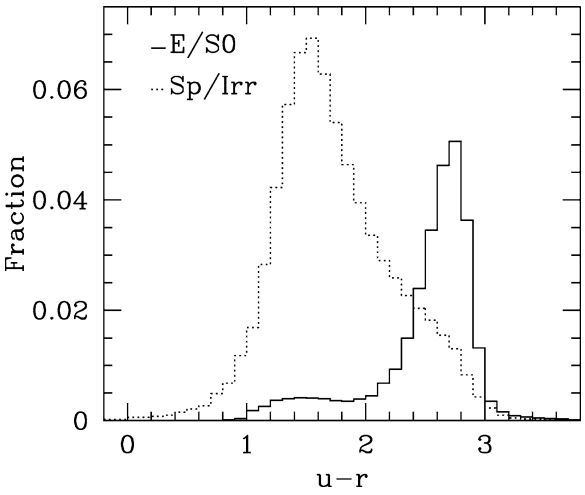


Fig. 2 Frequency distribution of early and late-type galaxies as a function of $u - r$ colors

of early-type galaxies shows a peak at $u - r \approx 2.65$ with the secondary peak at $u - r \approx 1.45$, while that of late-type galaxies shows a peak at $u - r \approx 1.55$ with a slow decline to the red colors. The secondary peak of the early type galaxies is caused by the blue early-type galaxies that comprise about 10% of early-type galaxies. The distribution of $u - r$ colors of late-type galaxies suggests that there are two populations of blue elliptical galaxies. One population has blue colors centered on $u - r \approx 1.45$ and the other population has red colors similar to those of E/S0 galaxies. The latter population is consisted mostly of Sa galaxies whose photometric properties are similar to those of S0 galaxies.

3 Blue Elliptical Galaxies

Blue elliptical galaxies are elliptical galaxies that have $u - r$ colors similar to those of late-type galaxies in the blue cloud. Thus, it is necessary to know their morphological types beforehand to segregate them from the late-type galaxies. We consider all the elliptical galaxies that have $u - r$ colors smaller than 2.2 as the blue elliptical galaxies although a considerable fraction of the galaxies in the green zone ($0.18 < u - r < 2.2$) are thought to be contaminated by the blue tail of the red sequence galaxies.

3.1 Morphology

Figure 3 presents the three typical images of blue elliptical galaxies that display different morphological properties. The blue elliptical galaxy in the left panel of

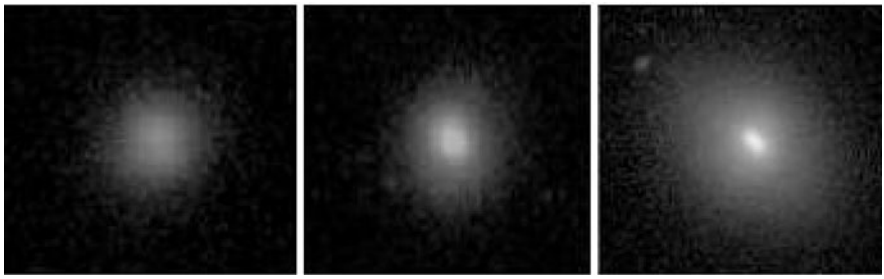


Fig. 3 Examples of blue elliptical galaxies. From the *left* to the *right*, SDSS J11440.70+165358.7, J123223.61-014424.4, J122417.04+672623.9

Fig. 3 is characterized by the bluer colors due to global star formation, whereas the blue elliptical galaxies in the middle and right panels of Fig. 3 are characterized by the blue cores of young stellar populations, but have different global $u - r$ colors. The majority of blue elliptical galaxies outside the local volume ($z > 0.01$) are the blue-cored galaxies with global colors similar to that in the right panel of Fig. 3. They are mostly the green valley galaxies.

Galaxies with globally blue colors and no blue core are easily confused with blue compact dwarfs (BCDs) in visual inspections. Owing to the central concentration of light as well as $u - r$ colors and color gradient used in the automated classifier of Park and Choi [6], we can segregate the blue ellipticals from the BCDs. However, a significant fraction of BCDs have morphological properties similar to those of blue elliptical galaxies except for the extremely blue colors. These BCD-like blue ellipticals and the blue-cored ellipticals with blue global colors seem to be the most dominant populations of the local universe. The absence of globally blue elliptical galaxies at high redshifts is due to the dwarf nature of blue elliptical galaxies.

3.2 Spectral Properties

The blue colors of the elliptical galaxies with $u - r < 1.8$ are likely to be due to the global star formation while those of the elliptical galaxies in the green valley seem to be mostly due to active star formation at cores. A significant fraction ($\sim 25\%$) of blue-cored galaxies seem to be related with the AGN activity since about a quarter of emission line galaxies show AGN characteristics in their spectra according to AGN criteria of Kauffmann et al. [4]. The reason for the higher AGN fractions for the elliptical galaxies in the green valley is that they are more likely to host supermassive black holes than the elliptical galaxies in the blue cloud regions because the former are generally more massive than the latter. The spectral properties of blue elliptical galaxies with $u - r < 1.8$ and those of the late-type galaxies with $u - r < 1.0$ are very similar. Thus, the blue elliptical galaxies with global star formation seem to be the sites of on-going galaxy formation.

3.3 Fraction of Blue Elliptical Galaxies

We derived the fractions of blue elliptical galaxies as a function of redshift, along with N_{blueE}/N_{redE}, and the total elliptical fractions (Fig. 4). As shown in Fig. 4, the blue elliptical fractions decrease with redshifts due to flux limit imposed by SDSS, $r = 17.77$, while total elliptical fractions increase with redshifts owing to the red sequence ellipticals at high redshifts. The blue elliptical fractions derived in the present study are somewhat smaller than those of Kannappan, Guie and Baker [3] and Schawinski et al. [9] who reported blue elliptical fractions of ~ 0.15. But if we consider BCD-like galaxies that have morphological properties similar to those of blue ellipticals except for the extremely blue colors as blue ellipticals, the fractions are similar to each other. The ratio of blue ellipticals relative to the red sequence ellipticals decrease rapidly from ~ 0.6 at $z < 0.01$ to ~ 0.1 at $z \approx 0.05$. This is mainly due to the faint dwarf nature of globally blue elliptical galaxies.

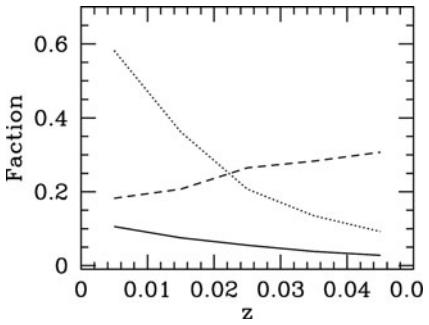


Fig. 4 Blue elliptical fractions as a function of redshifts. *Solid line* represents blue elliptical fractions while *dotted* and *short-dashed lines* indicate the N_{blueE}/N_{redE} and the total elliptical fractions, respectively

4 Discussion and Conclusions

There seems to be three types of blue elliptical galaxies. The first type of blue elliptical galaxies has elliptical morphology with blue colors due to global star formation. The BCDs are thought to be the faintest and bluest extremes of this population. They are highly abundant in the local universe and are likely to be found in the under-dense regions. The second type of blue elliptical galaxies shows blue global colors and bluer cores of massive star formation. They are most abundant population and usually found in the under-dense regions. The third type of blue elliptical galaxies is characterized by moderately blue colors and blue cores showing strong emission lines. A significant fraction of these galaxies have spectra with AGN characteristics.

The blue ellipticals showing blue global colors are dwarf galaxies which are too faint to be observed outside the local volume. They contribute to the faint end of the galaxy luminosity function significantly. They are much fainter than the dwarf

elliptical galaxies that have red colors and are preponderant in the high-density regions. The luminosity difference between the genuine blue elliptical galaxies that are bluer than $u - r = 1.8$ and the classical dwarf elliptical galaxies dictates the morphology–luminosity relation which is closely related with the local background density.

The blue elliptical galaxies, especially those with blue global colors are thought to be the sites of on-going galaxy formation. The reason for the delayed formation for the blue elliptical galaxies is the dependence of dynamical times on the local background density. Since blue elliptical galaxies are located in the low-density regions, the collapsing and accretion timescales are much longer than those for the high-density regions where red sequence galaxies including the dwarf ellipticals are likely to be located. It is consistent with the down-sizing scenario of galaxy formation which predicts late formation of less massive galaxies. The bright blue ellipticals, mostly those in the green valley ($1.8 < u - r < 2.2$), seem to be in the transition stage where active star formation is quenched by AGN feedback.

Acknowledgments We thank to the KIAS to provide value-added catalog SDSS DR7 which includes redshifts of bright galaxies prepared by C. Park and Y.-Y Choi. This work was supported by KOSEF grant through ARCSEC.

References

1. Ann, H.B., Park, C., Choi, Y.-Y. 2008, MNRAS, 389, 86
2. Fukugita, M. et al. 2004, AJ, 127, 3155
3. Kannappan, S.J., Guie, J.M., Baker, AJ. 2009, AJ, 138, 579
4. Kauffmann, G. et al. 2003, MNRAS, 346, 1055
5. Lee, J.H., Lee, M.G., Park, C., Choi, Y.-Y. 2010, MNRAS, 403, 1930
6. Park, C., Choi, Y.-Y. 2005, ApJ, 635, 29
7. Park, C. et al. 2005, ApJ, 633, 11
8. Schawinski, K. et al. 2007, ApJS, 173, 512
9. Schawinski, K. et al. 2009, MNRAS, 396, 818
10. Strateva, I. et al. 2001, AJ,122, 1861
11. Suh, H. et al. 2010, ApJS, 187, 374

Rings and Bars: Unmasking Secular Evolution of Galaxies

Johan H. Knapen

Abstract Secular evolution gradually shapes galaxies by internal processes, in contrast to early cosmological evolution which is more rapid. An important driver of secular evolution is the flow of gas from the disk into the central regions, often under the influence of a bar. In this chapter we review several new observational results on bars and nuclear rings in galaxies. They show that these components are intimately linked to each other and to the properties of their host galaxy. We briefly discuss how upcoming observations, e.g., imaging from the *Spitzer* Survey of Stellar Structure in Galaxies (S^4G), will lead to significant further advances in this area of research.

1 Introduction

The general topic of galaxy evolution is enjoying widespread attention in the literature. The aim is to answer, through observations and numerical modelling, how galaxies have evolved from the earliest stages of the Universe, or from a subsequent epoch of formation, to the shape in which we observe them in the local Universe. The two main strands in this wide topic are cosmological evolution of galaxies, which deals with their formation and early evolution, and secular evolution, with which we mean the internal evolution of galaxies, under the influence of the dynamical actions of, e.g., bars or spiral arms. Secular evolution, as comprehensively reviewed by Kormendy and Kennicutt [55], is a relatively slow process compared to the more rapid evolution undergone by galaxies in the early Universe, the latter often due to mergers and galaxy–galaxy interactions.

Because the early stages of galaxy formation and evolution are hard or sometimes impossible to observe due to the combined effects of distance, redshift, and dust extinction, the detailed study of nearby galaxies is one of the very few ways to confirm the detailed predictions of models of large-scale galaxy formation and

J.H. Knapen (✉)
Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain; Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain
e-mail: jhk@iac.es

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_18,

evolution. This kind of study can be nicknamed "galactic palaeontology", because in local galaxies we study the "fossil record" of billions of years of galaxy evolution – both cosmological and secular. To read and interpret this fossil record, a combination of many different observational and interpretational techniques must be used, from observations of individual stars in our own Milky Way and the most nearby galaxies to photometric and kinematic observations across many different wavelengths in external galaxies, all combined with a wide range of interpretational, analytical, and numerical tools.

One of the major drivers of the internal evolution is the flow of gaseous material, from the disk to the central regions of the galaxy. To move inwards, rotating gas must lose angular momentum, and it can do so by shocking and under the influence of a non-axisymmetric potential. That in turn gets set up by a bar, by interactions or minor mergers, or even by less obvious deviations from axisymmetry like spiral arms, ovals, or lenses (e.g., [25, 49, 55, 87, 94, 95]). Bars are very common in galaxies, about two thirds of local galaxies have a bar [29, 38, 41, 43, 52, 56, 58, 61, 67, 72, 74, 75, 90, 91]. As we will see below, in Section 3, nuclear rings can occur in unbarred galaxies, and they seem to prove that the non-axisymmetry induced by a spiral or oval may well be enough to induce gas inflow and thus lead to secular evolution. Although most galaxies are barred, secular evolution is thus not dependent on the presence of a bar.

Galactic interactions are rare, occurring in only about 2% of local galaxies (up to 4% if merely bright galaxies are considered [50]). Close companions to local galaxies are much more common, and Knapen and James [50] found that some 15% of local galaxies have a companion not more than 3 mag fainter than itself within a radius of five times the diameter of the galaxy under consideration and within a range of $\pm 200\,\mathrm{km\,s^{-1}}$ in systemic velocity. The effects on the star formation rate in galaxies of the presence of a close companion, and even of interactions, are, perhaps surprisingly, limited (observations by Bushouse [12]; Smith et al. [96]; Woods & Geller [100]; Li et al. [65]; Knapen & James [50]; Jogee et al. [45]; Rogers et al. [81]; Ellison et al. [33], corroborated by numerical simulations by Mihos & Hernquist [73]; Kapferer et al. [46]; Di Matteo et al. [30, 31]; Cox et al. [27]). Statistically, the star formation rate is raised by a factor of just under two by the presence of a close companion, but the Hα equivalent width is hardly increased at all [50]. This implies that even though galaxies with close companions tend to form stars at a higher rate, they do so over extended periods of time, and not as a burst. Even the majority of the truly interacting galaxies in Knapen and James's [50] sample have unremarkable star formation properties. The reason that extreme star-forming galaxies, such as ultra-luminous infrared galaxies (ULIRGs), are found to be almost ubiquitously interacting must surely be a selection effect. In general, interactions do not always cause starbursts, and starbursts do not always occur in interacting galaxies.

In fact, Knapen and James [50] used their sample of 327 nearby disk galaxies to explore how one might best define the term "starburst". They concluded that *none* of the definitions that are in common use in the literature can be considered to be objective and generally discriminant. For instance, selecting galaxies on the basis of

their high star formation rate yields large star-forming disk galaxies. Selecting those with high equivalent widths (apparently a bona fide starburst discriminator as this selects galaxies with a much enhanced current star formation rate as compared to the average rate in the past) yields primarily late-type galaxies of very small mass, whose star formation activity is caused by one or a few HII regions (and which will have very low impact on the intergalactic medium through, e.g., stellar winds). And selecting galaxies with the shortest gas depletion timescales does not only select galaxies with very high current star formation rates but also gas-poor early-type galaxies with a very small star formation rate. The conclusion of Knapen and James [50] is that starbursts are very hard to define properly, and the use of the term should be restricted to well-described small numbers of objects.

This review deals with aspects of secular evolution, and how we can trace its actions back through the detailed study of structural components, particularly bars and rings, in nearby galaxies. To interpret the effects of these agents, the tools we will use here are primarily optical and infrared imaging, and two-dimensional kinematic mapping. Other authors have presented reviews on galactic evolution, and in particular the paper by Kormendy and Kennicutt [55] presents an authoritative review of the subject of secular evolution. We will supplement that by presenting selected recent results on bars and rings that highlight the intricate and detailed connections that exist between the different structural components of a galaxy and the overall galactic evolution.

In Section 2 of this chapter, we will review how the strength of a bar is connected to many of the basic properties of the bar, such as its length, or the shape of its dust lanes. We will also see that bars are indeed connected to spirals, and how bars in S0 galaxies may be different from those in spirals. Section 3 describes how the basic physical properties of the host galaxy and, where present, its bar condition the location and morphology of a nuclear ring, thus highlighting the close physical connections between these components. Section 4 discusses how the *Spitzer* Survey of Stellar Structure in Galaxies (S^4G; [92]) will deliver the data which should allow us to make significant further progress in the study of galaxy evolution by means of detailed analyses of the stellar component in a large sample of nearby galaxies. We briefly present our conclusions in Section 5.

2 Bars

Bars are common in galaxies: around one third of all disk galaxies have a strong bar, and when including weak bars and ovals, the fraction of barred galaxies rises to well over two-thirds (references in the introductory section). The bar fraction is roughly invariant with morphological type across the spirals [50], but lower in the S0s [9, 61]. There is an interesting debate in the literature over whether the fraction of bars is lower [1, 91, 97] or not [36, 44, 102] at the highest redshifts where bars can be reliably recognised, around $z \sim 1$. Nair and Abraham [76] find that the fraction of barred spiral galaxies between redshifts of 0.01 and 0.1 is a strong function of stellar

mass and star formation history. They suggest that the discrepancy in the reported bar fraction evolution with redshift may have its origin in observational biases and selection criteria, because the strong bar fraction is sensitive to the mass range that is being probed (see also [21], who find that the early- and late-type bar fractions vary with redshift and galaxy mass).

Bars are efficient agents of angular momentum transfer in galaxies and are expected to evolve as they transfer angular momentum to the disk and the halo and allow the radial inflow of gas (e.g., [6, 7, 28, 66, 70, 89, 94, 95]). As a result, bars should slow down and grow over time, and become longer, thinner, and stronger [7]. As bars are relatively easy to observe, they can be used as effective probes of secular evolution in galaxies.

2.1 Basic Properties of Bars

To study their basic properties and the relations to their host galaxies, bars are parametrised by determining their main parameters, such as length, thickness, and strength. Many different definitions of such parameters can be devised and have been presented in the literature, but modern analyses use techniques such as image decomposition, Fourier analysis, and the determination of bar strength parameters from deep high-quality images. This is done preferably in the infrared where light traces mass better than in the optical, and where the effects of dust extinction are greatly reduced. As an illustration of the kind of relations that are being found, we show in Fig. 1 the measurements that have been obtained by Comerón et al. [25] as part of their Atlas of Images of Nuclear Rings (AINUR), of which we will later review more results on nuclear rings. This is not an unbiased survey, but its results do illustrate the points we wish to highlight here.

Figure 1 shows various measures of the length of the bar, as well as ellipticity, as a function of the bar strength parameter Q_g, and indicating three ranges of morphological types. Q_g is a non-axisymmetric torque parameter, quantifying the impact of non-axisymmetries in a galaxy by measuring the maximum value of the tangential forces normalised by the axisymmetric force field (e.g. [11, 13, 24, 57, 79, 81]).[1] Comerón et al. [25] determined Q_g values using azimuthal Fourier decomposition and polar integration (following [84] and [57]).

The data points are from a sample of 107 galaxies hosting nuclear rings (Comerón et al. [25]). They confirm the general trend, already noted by Martinet and Friedli [68] and Laurikainen et al. [64], among others, for bars to become longer as they become stronger, but the scatter is rather large. This conclusion can be drawn from any of the three bar length measures plotted, directly in units of

[1] Note that Q_g is measured across a whole galaxy and may include contributions from the spiral arms as well as the bar. In earlier work, what we now refer to as Q_g was called Q_b. In more modern work, we refer to Q_b as the non-axisymmetric torque from the bar only, after separating the contributions of bar and spiral arms (e.g., Buta et al. [17]). In early-type galaxies, including S0s, the spiral arm contribution is very small and Q_g can be used as Q_b.

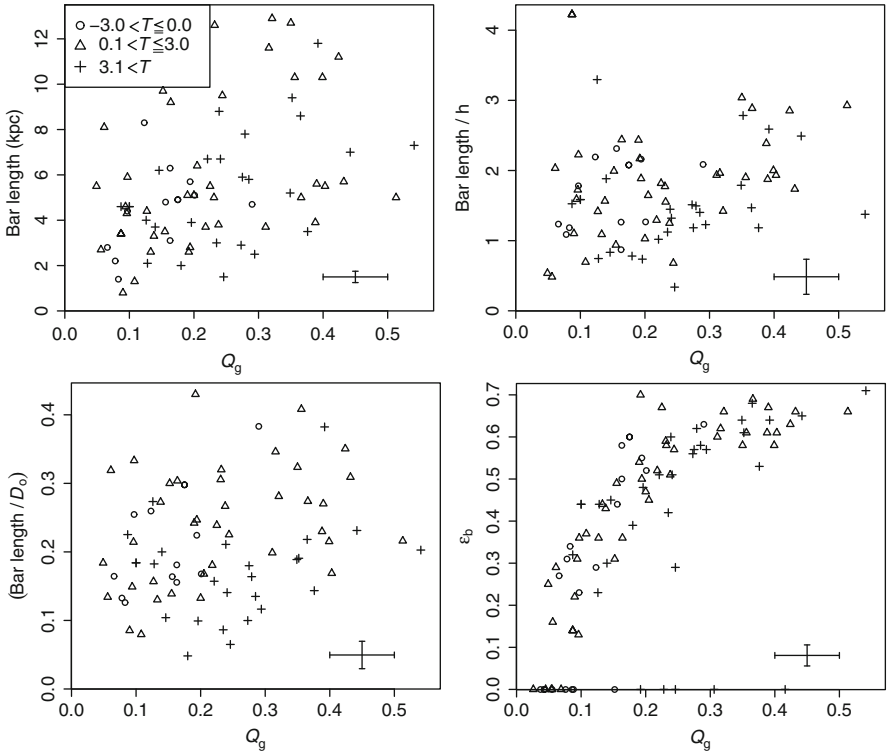


Fig. 1 Bar length and bar ellipticity (the latter in the *lower right panel*) as a function of the bar strength parameter Q_g, for 107 nuclear ring host galaxies. The *top left panel* shows the absolute bar length, the *top right panel* the bar length normalised by the scale length of the exponential disk, and the *bottom left panel* the bar length normalised by the disk size. Typical uncertainties are indicated in each panel, and the *symbol* styles indicate different morphological types as shown in the *top left*. Reproduced with permission from Comerón et al. [25]

kpc, and normalised by the disk scale length h or the disk size D_0. Two regions of the various diagrams in Fig. 1 are empty, which seems to indicate that there are no very strong short bars, nor any very weak and long ones. In fact, some late-type galaxies do have short bars with large Q_g, due to the fact that they have very little, or no, bulge to dilute the non-axisymmetric contribution of the bar to the potential (e.g., Laurikainen et al. [64]). This effect also explains why Q_g decreases but the bar length increases towards the earlier Hubble types ([19, 60]; Fig. 1). The bar ellipticity is seen to increase monotonically with Q_b [16], with some scatter but very few outliers.

From deep K_s-band images of 20 galaxies, Elmegreen et al. [37] confirmed independently that longer bars are stronger (as indicated in their paper by the peak amplitude of the normalised $m = 2$ Fourier component). They also found an interesting correlation between the bar length and an increased density in the central parts of the disks. Elmegreen et al. note that, as dense galaxies evolve faster, these

results indicate that bars grow in length and amplitude, with the densest galaxies showing the fastest evolution. Numerical modelling provides theoretical support for correlations such as the one between bar strength and length (see, e.g., the recent simulations by Villa-Vargas et al. [98], with varying gas fractions and gas spatial resolution).

Other aspects of bars which are directly related to secular evolution include the frequency of ansae-type morphology, significantly enhanced in early-type galaxies [60, 69]; the finding that bars with double-peaked Fourier amplitude profiles are stronger than those with single-peaked profiles [17] and occur more frequently in early-type galaxies than in spirals [60], which might indicate that they are more evolved [7, 8]; and the fact that most S0s have lenses [60].

2.2 Bar Dust Lanes

A further hint at the internal physical processes in bars is provided by the study of dust lanes of Comerón et al. [26]. Dust lanes have been recognised in relation to shocks in the flow of gas in barred galaxies for almost 50 years now [4], and Athanassoula [78] in fact predicted from numerical modelling that the degree of curvature of the main dust lanes in a bar should decrease as the bar gets stronger: stronger bars have straighter dust lanes. As dust lanes are clearly visible in barred galaxies (e.g., [77, 85]), this is in principle an easily observable tracer of the fundamental physics and dynamics of galactic bars, yet with the exception of the preliminary study by Knapen et al. [51] of only nine galaxies which confirmed the predictions, it had never been tested observationally. Comerón et al. [26] collected images of 55 bars of which the shape of the dust lanes could be measured and for which the bar strength $Q_{\rm b}$ was available from the literature.

The predicted correlation can indeed be recognised in their results, reproduced here in Fig. 2 (top panel). But the line which would indicate a linear relationship between stronger bars and less curved dust lanes is, in fact, only the upper envelope to the distribution of points. This is not scatter, as indicated by the typical uncertainties of each data point. The strength of the bar does not prescribe the degree of curvature that the dust lanes can have, but instead only provides an upper limit. Hence, strong bars can only have straight dust lanes, whereas weak bars allow their dust lanes to be either curved or straight.

To investigate what factor other than the bar strength might cause the degree of curvature of the dust lanes, Comerón et al. [26] analysed a set of 238 simulated galaxies, in 88 of which the dust lane curvature could be measured. The results of this, as shown in the bottom panel of Fig. 2, show that the scatter can be greatly reduced when the bar strength is described as a linear combination of $Q_{\rm b}$ and a/b, the quotient of the major and minor axes of the bar. These are both bar parameters, and since no other parameters (describing, e.g., the bulge) could be identified that reduce the scatter in the original diagram, this is proof that indeed the dust lane curvature is predominantly determined by the parameters of the bar. Just how the linear combination of $Q_{\rm b}$ and a/b, as plotted in Fig. 2, must be interpreted is an interesting question that remains to be explored further.

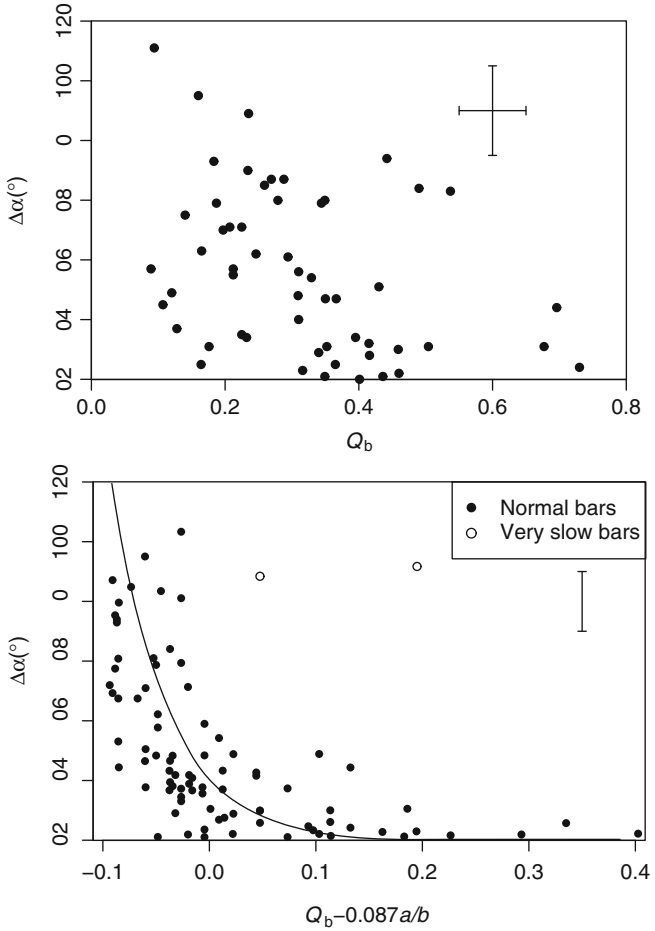


Fig. 2 Dust lane curvature as a function of bar strength parameter Q_b for 55 observed galaxies (*top panel*), and, for a set of 88 modelled galaxies, as a function of a combination Q_b and the bar ellipticity parameter a/b which minimizes the spread (*bottom panel*). The *curve* indicates the best fit to the "normal" bars. Reproduced with permission from Comerón et al. [26]

2.3 Bars and Spirals

Bars also affect the disk regions outside them. This is perhaps best seen in the way they drive spiral density waves. That they do is expected from theory and numerical modelling, but also in this case it has been hard to confirm observationally. Previously, various observational links between bars and spirals had been reported (e.g., [34, 54]), but more recent works have reported either good [10], some [15, 19], or no correlation [19, 32, 88] between bar and spiral strengths. The reasons for the discrepant results include sample size, methodology, and data quality, as summarised by Salo et al. [83]. The latter authors re-investigated this problem using the same

data which had led Buta et al. [15, 19] to inconclusive results, but using a novel analysis approach.

Salo et al. [83], rather than compare maxima of bar strength and spiral density amplitude, compare the *local* bar forcing and spiral amplitude as a function of distance. As reproduced in Fig. 3, this yields a correlation between bar forcing and spiral amplitude, strongest near the ends of the bar, but statistically significant up to one and a half times the bar radius. This confirms that the stellar spirals represent a continuation of the bar mode or are driven by the bar through some mechanism. Outside the region of influence of the bar, the spirals may be independent modes or transient. As the correlation between bar and spiral is similar for early- and late-type spirals, and for small and large bars, Salo et al. [83] conclude that the forcing of the spiral by the bar is a general occurrence.

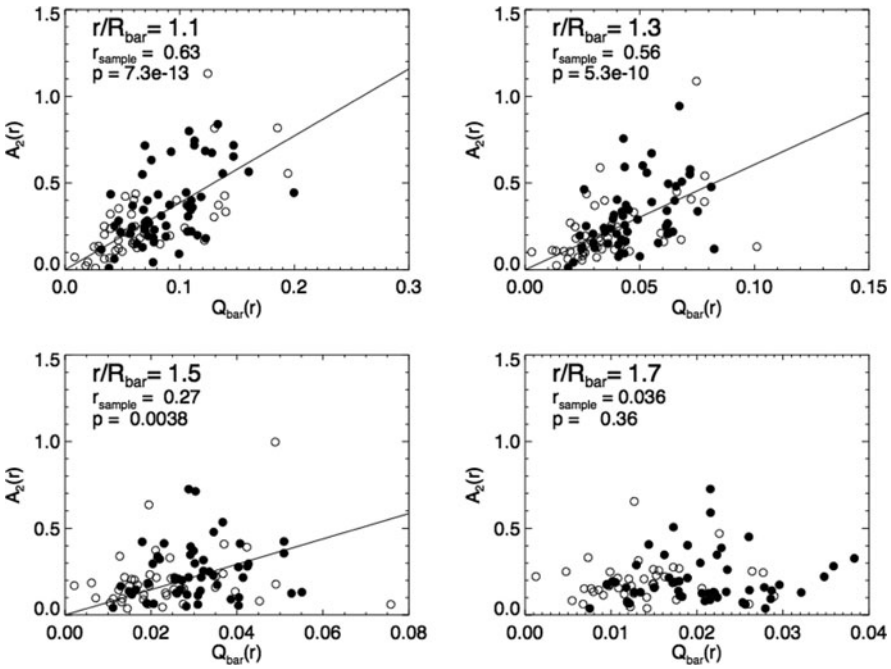


Fig. 3 Relation between the local bar forcing ($Q_{\rm bar}(r)$) and the local spiral amplitude ($A_2(r)$) for 103 barred galaxies, and at four distances $r/R_{\rm bar}$, normalised to the bar length. The values $r_{\rm sample}$ and p indicate the correlation coefficient and significance; the best fit is indicated by a *line* in the three cases of highest significance. *Open* and *filled* symbols indicate *short-* and *long bars* and yield the same distribution. Reproduced with permission from Salo et al. [83]

2.4 Bars and Secular Evolution

As a final illustration of the possible use of bars as tracers of secular evolution in galaxies, we cite the recent work by Buta et al. [18]. These authors use images from the Near-Infrared S0 Survey (NIRS0S; [17, 59, 63]) with similar images of spirals

from the literature to study the bar strength *distribution* in lenticular as compared to spiral galaxies. This is a pertinent question because lenticular galaxies remain somewhat of an enigma. They have been positioned between ellipticals and spirals in galaxy classification schemes since the earliest work by Hubble [42], and one of the main questions is whether the S0 galaxies are more closely related to elliptical or spiral galaxies.

In modern Lambda Cold Dark Matter (ΛCDM) cosmology, ellipticals and the bulges of spirals are formed early in the evolution of the Universe by mergers, and their properties were thus established already early on. In this framework, S0s may either be formed as ellipticals, or alternatively they may be transformed spirals, formed as the disks have lost their gas by some stripping mechanism. In a recent paper also based on NIRS0S images, Laurikainen et al. [62] confirm from photometric scaling relations that the formative processes of bulges and disks in S0s are coupled, and that the bulges of S0s are similar to those of spirals with bright bulges. They conclude that spiral galaxies with bulges brighter than $M_K\,(\text{bulge}) < -20$ mag can evolve directly into S0s, due to stripping of gas followed by truncated star formation. This is prima facie evidence for secular evolution.

The bar strength distribution of Buta et al. [18] shows (Fig. 4) that S0 galaxies on average have weaker bars than spiral galaxies in general, and even than early-type spirals, of type S0/a and Sa. Several studies, mostly based on contrast, have shown that bars in early-type galaxies are stronger and longer than those in later-type galaxies (e.g., [35]), other studies report weaker bars in lenticulars as compared to spirals (e.g., [2], who used a strength measure based on bar ellipticity). When considering the gravitational forcing of the bar, as measured, e.g., through the $Q_{\rm b}$ parameter, the non-axisymmetric bar forcing is to a large extent cancelled in early-type galaxies by the axisymmetric contribution to the potential of their more massive

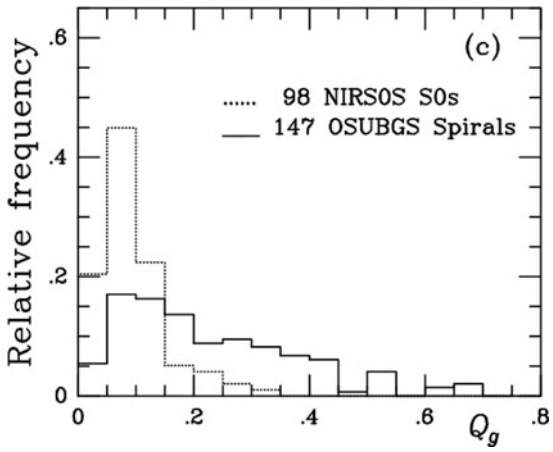


Fig. 4 Histograms of the bar strength distribution in S0 and spiral galaxies. Reproduced with permission from Buta et al. [18]

bulges (e.g., [64]), so that in early-type galaxies bars are longer and have larger $A2$-amplitudes, but have lower Q_g and ellipticity than spiral galaxies [60].

The differences found by Buta et al. [18] are significant, and only partly due to the dilution of the bar torques by the large bulges of the S0s, or the thicker disks in S0s. They tell us that if indeed S0s have evolved from spirals, the bar evolution must have continued after the gas depletion, which might also be suggested by the lower frequency of bars [61] and the higher frequency of lenses [53, 60] in S0s as compared to spirals. Buta et al. [18] speculate that the bars in early-type galaxies can be slightly skewed, so that a potential-density phase shift [101] can evolve the stellar distribution, leading to continued bulge building and bar weakening.

3 Rings

Rings in galaxies are common and are thought to trace resonances in the disk. This means that they can be used as easily observable tracers of the underlying dynamical structure of the galaxy. Of the three main categories of "resonance rings", outer, inner, and nuclear rings, the former two types can be recognised on standard optical imaging of galaxies and have thus been classified as part of the main galaxy catalogues. In particular, the RC3 catalogue [29] assigns the categories "R" for outer ring, and "r" for inner ring (pseudorings were also classified but are outside the scope of this review). The classification scheme has been perfected in the De Vaucouleurs Atlas of Galaxies [14], and most recently by Buta et al. [20]. In the latter paper, we used *Spitzer* mid-IR images from the S^4G survey to classify 207 galaxies, and introduced also the third class of rings, nuclear rings, into the classification ("nr").

In contrast to inner and outer rings, it is much harder to recognise nuclear rings in the images typically used for galaxy classification. As a result, there has not been a complete inventory of nuclear rings. This we have remedied with AINUR [25], where 113 bona fide nuclear rings in 107 galaxies were catalogued and studied. The main aims of AINUR were, first, to make an inventory as complete as possible of nuclear rings in the local Universe, and second, to study in a statistical sense their properties in relation to the properties of their host galaxies. The importance of nuclear rings lies primarily in the fact that they are assumed to be tracers of recent gas inflow to the circumnuclear region [5, 22, 23, 40, 49, 80, 94]. They form significant quantities of stars [47] and may thus help the secular building of a bulge and are close to the region where non-stellar activity occurs in many galaxies [48].

As seen in Fig. 5, the radii of nuclear rings vary from a few tens of pc (the limit allowed by space-based imaging) to over 3 kpc. We can reliably recognise nuclear rings in galaxies up to a distance of some 80 Mpc, although only the larger rings are seen in galaxies over some 20 Mpc away. In AINUR, we define a nuclear ring primarily as a star-forming ring-like feature in the proximity of the nucleus, employing a number of additional criteria relating to the ring width and radius. It is usually easy to decide whether rings are nuclear or inner, because inner rings occur near the end of a bar [87], whereas nuclear rings occur well within the bar. In the

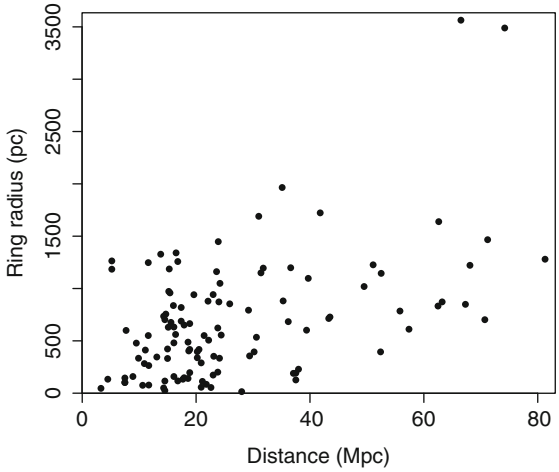


Fig. 5 Distribution of nuclear ring radius as a function of distance to the host galaxy for the 113 nuclear rings in AINUR. Reproduced with permission from Comerón et al. [25]

absence of a bar, and if only one ring is present, distinguishing between nuclear and inner rings may be impossible. We do not include pseudorings (intermediate between nuclear spiral and nuclear ring) but the dividing line is somewhat fuzzy.

AINUR catalogues 113 nuclear rings in a total of 107 galaxies, 18 of which are unbarred and 78 of which are barred disk galaxies. Most of these rings had been reported before and were confirmed from archival *Hubble Space Telescope* (*HST*) or other imaging (a significant number of other features called nuclear rings in the literature could not be confirmed and were not included; this category includes inner rings, pseudorings, or galaxies for which no high-quality imaging is available). A total of 17 previously unreported nuclear rings were discovered. The AINUR catalogue can be considered a complete list of all bona fide nuclear rings known at present, but as more high-quality (mainly *HST*) imaging comes available, it is likely that a few more nuclear rings are discovered.

On the basis of a complete sample of galaxies within AINUR, the fraction of disk galaxies with morphological types in the range $-3 < T < 7$ that host a star-forming nuclear ring is confirmed to be $20 \pm 2\%$ (see also [48]). This is a high fraction and considering that the star formation activity is mostly seen in emission from massive stars (UV, Hα) and thus quite possibly short-lived [3, 86], one can conclude that nuclear rings are very frequent indeed.

Using basic nuclear ring parameters such as size determined from the AINUR imaging (mostly *HST*), and host galaxy and bar parameters (e.g., bar length, bar ellipticity, $Q_{\rm g}$) determined from near-IR images from the 2 Micron All-Sky Survey (2MASS), Comerón et al. [25] explored relations between the nuclear rings and their host galaxies and bars (where present).

Figure 6 shows two examples. In the left panel, it is seen that for small values of $Q_{\rm g}$ (the non-axisymmetric torque of the galaxy, which in most cases is dominated

by the bar – small Q_g values thus generally denote weak bars, and high values strong bars) a wide range of relative nuclear ring sizes is allowed, where galaxies with high Q_g can only have small rings. This had been found from a much smaller sample of nuclear rings by Knapen [48] and seen in simulations by Salo et al. [84]. Large nuclear rings can only occur in weak bars. Analogous to what we found for the bar dust lanes (Section 2.2), the bar strength does not prescribe the nuclear ring size, but does set an upper limit. What determines the exact size of a nuclear ring within the range allowed by its bar is not clear, but must be related to the shape of the gravitational potential which conditions the location of the Inner Lindblad Resonances (ILRs).

The right panel of Fig. 6 shows how the length of the bar limits the radius of the nuclear ring (to around $r_{ring} = r_{bar}/4$), while rings smaller than the upper limit are apparently allowed. Both these relations confirm that the size of the nuclear ring is limited by basic parameters of the bar, and thus that the rings must be closely related dynamically to the bar.

Further evidence for the intricate links between host galaxy, bar, and nuclear ring is provided by Mazzuca et al. [71] from a study of the inner rotation curves of 13 nuclear ring host galaxies. They obtained Hα velocity fields using the DensePak integral field unit on the 3.5 m Wisconsin, Indiana, Yale & NOAO (WIYN) telescope and the TAURUS Fabry-Perot instrument on the 4.2 m William Herschel Telescope. From these, they derived rotation curves after assuring that the non-circular motions were small enough to allow such a derivation. As the rotation curve is, theoretically, intricately linked to the gravitational potential and to the location of the ILRs which give rise to the nuclear rings [49], Mazzuca et al. [71] then parametrised the inner part of the rotation curve and compared the results with the location of the nuclear ring, as well as the properties of the bar.

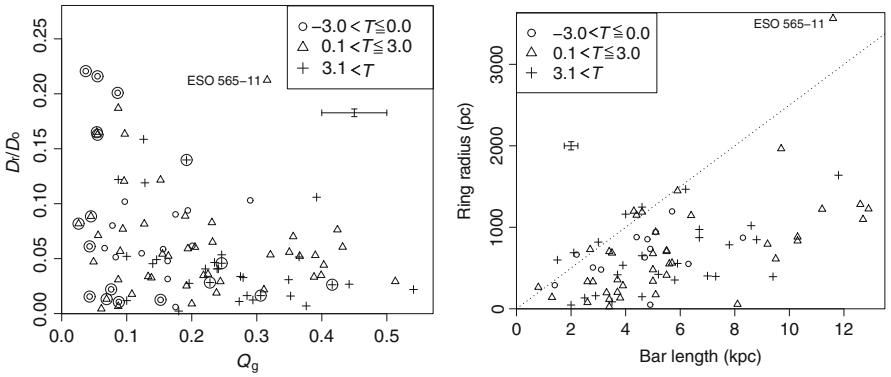


Fig. 6 *Left panel*: relative nuclear ring size, normalised by the host galaxy disk size, as a function of the bar strength parameter Q_g, and separated by morphological type. *Encircled symbols* denote unbarred galaxies. *Right panel*: Absolute nuclear ring size versus the bar length. *Dashed line* indicates bar lengths four times the ring radii. Reproduced with permission from Comerón et al. [25]

The main rotation curve parameter considered by Mazzuca et al. [71] is its rise rate, defined as the ratio between the difference of the velocity at the turnover point (where the rotation curve stops rising and flattens out) and that at the origin, and the difference in radius between those two points. As can be seen in Fig. 7, plotting the rise rate of the rotation curve as a function of the relative nuclear ring size or the nuclear ring width yields diagrams reminiscent of the ones we have presented earlier in this chapter: they outline an upper limit, indicating that ring size and width limit the rotation curve rise rate or vice versa. Large nuclear rings, and wide nuclear rings, can only occur when the rotation curve rises slowly. In the case of a rapidly rising rotation curve, the nuclear ring can only be small and narrow.

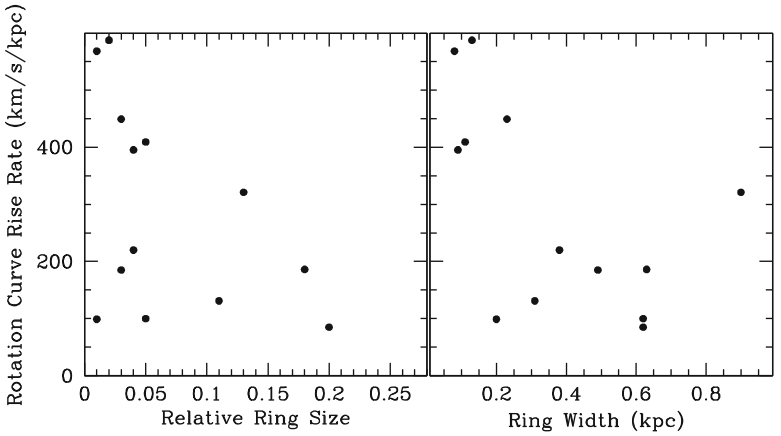


Fig. 7 Rise rate of the inner part of the rotation curve as a function of the nuclear ring size relative to the disk size (*left panel*), and the nuclear ring width (*right panel*). Data from Mazzuca et al. [71]

Mazzuca et al. [71] thus present a neat observational confirmation that the rotation curve and the metric parameters of the nuclear ring are intricately linked. The physical background of this link must include the location of the resonances: linear theory[2] predicts that ILRs, and nuclear rings, occur where the rate of change of the circular velocity is highest – this is where the rotation curve turns over. The rise rate of the rotation curve is thus a measure of where the ILRs are located: as rotation curves tend to flatten out at circular velocities of between 100 and 200 km s^{-1}, a high rise rate means that the ILRs are close to the nucleus, and thus that the nuclear ring must be small. It also implies that the ILRs are close together radially, which leads to narrower nuclear rings. All this is exactly as seen observationally. A slowly rising rotation curve implies that the ILRs can be located further out from the nucleus, but also that the distance between the inner and outer ILRs (which in linear theory limit

[2] In the linear approximation the gravitational potential of the bar is considered to be axisymmetric. While this may be reasonable for weak bars, it is not correct in the case of a strong bar [90, 93]. Nevertheless, as an intuitive tool it can provide a useful illustration.

the radial range where the nuclear ring can occur) can increase. Depending on the precise shape of the gravitational potential in the inner region, the inner ILR can still be relatively close to the nucleus, which might explain that in galaxies with slowly rising rotation curves large and wide nuclear rings can occur, but small and narrow ones are not excluded.

Figure 8 shows that the rotation curve rise rate is also located to the bar strength, measured here through the non-axisymmetric torque parameter Q_g. The figure indicates a general trend of more steeply rising rotation curves as Q_g increases (the one notable exception is NGC 1530 which is a galaxy with a bar that is exceptional in many more ways). It is not clear whether the two parameters plotted here are directly or indirectly related – we have already seen that stronger bars tend to host smaller rings. But as a general conclusion, it is beyond doubt now that the underlying dynamics of a host galaxy influences its appearance and kinematics, as well as detailed aspects of some of the structural components such as bars and rings.

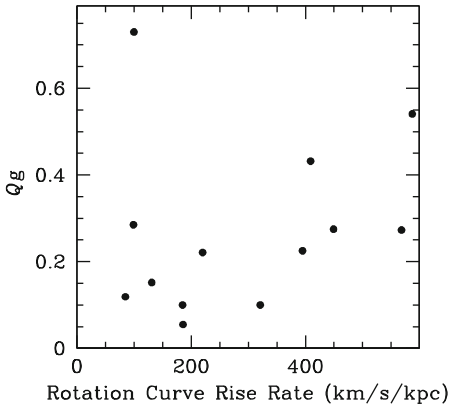


Fig. 8 Non-axisymmetric torque parameter Q_g, which can be interpreted as an indicator of the strength of the bar, plotted against the rotation curve rise rate. Data from Mazzuca et al. [71]

4 Future Work: The S^{4}G Survey

The future of the field of secular evolution in galaxies is in no small measure infrared. The *Herschel* satellite has just started producing spectacular data in nearby galaxies, the *James Webb Space Telescope* (*JWST)* is being prepared for launch, and the *Spitzer Space Telescope* (*SST*; [99]) is still producing wonderful data, even now its coolant has been depleted and the telescope is "warm".

The *Spitzer* Survey of Stellar Structure in Galaxies (S^{4}G; [92]) is designed to be the definitive survey of the distribution of stellar structure in the nearby Universe. Over the 2 years of the warm mission of the *SST*, the S^{4}G will observe the stellar mass distribution in a volume-, magnitude-, and size-limited ($d < 40$ Mpc, $m_B < 15.5$, $D_{25} > 1$ arcmin) sample of 2,331 galaxies using the Infrared Array Camera (IRAC; [39]) at 3.6 and 4.5 μm. This survey will provide an unprecedented set of

imaging data which will be used to study the stellar structure in galaxies. Among many other topics, the S^{4}G data will allow the study of how outer disks and halos are formed, how the formation and evolution of galactic structures are affected by galaxy interactions, or which structural parameters govern secular galaxy evolution. The large sample, ranging from dwarfs to spirals to ellipticals will allow for such structural studies both as a function of stellar mass and as a function of environment, vital to test cosmological simulations predicting the mass properties of present-day galaxies.

The S^{4}G survey was awarded 637.2 h of time on the *Spitzer* telescope, and the observations are in progress. It is expected that the survey will be completed by mid-2011. The data are being analysed using dedicated pipeline software mostly developed by the S^{4}G consortium, which will perform a basic reduction, sky subtraction and mosaicing of the images, will produce masks of foreground stars, will measure basic parameters of all galaxies, and will deliver multi-component decompositions (see [92] for a detailed description).

As an example of the data quality, including depth of imaging and field of view related to the size of the galaxy, that will be delivered by the S^{4}G for *all* survey galaxies, we present in Fig. 9 a 3.6 μm image of the well-known galaxy M51. The image is reproduced from a paper by Buta et al. [20], which discusses the morphology of the first 207 S^{4}G galaxies. Not only does the image show the intricate outer tidal structure caused by the interaction of M51 and its companion NGC 5194 (see [82]), it also illustrates how a high-quality dust-free vision of a galaxy can lead to a better insight into its nature. On the basis of this new image, Buta et al. [20] were able to revise the morphological classification of M51 from SA(s)bc pec (RC3) to S$\underline{\text{A}}$B(rs,nr)bc, but much more spectacular is the revision of the type of the companion of M51, NGC 5195, which was classified as an I0 pec galaxy in de Vaucouleurs et al. [29], but which Buta et al. [20] could classify as SAB(r)0/a pec. The bar and ring giving rise to this new classification can be easily seen in Fig. 9.

Fig. 9 *Spitzer* 3.6 μm S^{4}G image of the galaxy M51 (NGC 5194, *right*) and its companion NGC 5195. From Buta et al. [20]

The combination of deep, wide-field, mid-infrared imaging, and a large sample offered by the S^{4}G will allow significant progress on most of the issues discussed earlier in this chapter. For instance, the survey will allow a comprehensive and definitive study of bar and spiral arm properties across a wide range of host galaxy parameters. The new images will allow a more complete study of the host galaxies of not only nuclear but also inner and outer rings, including the interesting cases of unbarred ring hosts. The combination of the *Spitzer* images with ancillary data, for instance *GALEX* UV imaging, Hα imaging, atomic and molecular gas observations, and kinematic observations offers almost limitless opportunities for scientific progress. The S^{4}G should thus lead to a much more complete understanding of galaxy evolution and of the precise role of secular evolution.

5 Conclusions

This review has highlighted some recent work which further illuminates the tight physical links between bars, nuclear rings, and their host galaxies. These relationships are shedding light on the underlying dynamical structure and on the overall process of secular evolution of galaxies. The dynamics of bars is now rather well understood. They indeed stimulate inflow of gaseous material from the disk to the central region of a galaxy, and thus inflowing gas can help build the bulge and drive evolution, even across morphological types. Nuclear rings are common, and transform significant amounts of gas into stars, again increasing the bulge mass. Nuclear rings occur mostly in barred galaxies, but the fraction occurring in unbarred galaxies is similar to the overall fraction of unbarred disk galaxies, so a causal link is not proven. Ovals, interaction events, or possibly even strong spiral arms can cause enough non-axisymmetry in the gravitational potential to stimulate the formation of a nuclear ring, even in the absence of a bar. This shows how ubiquitously secular evolution in galaxies can occur.

The future for this area of research is bright, because significant amounts of new data are being collected which will allow us to increase sample sizes and progress to study secular evolution on a more fundamental basis. We explicitly mentioned the S^{4}G survey, which will deliver deep, mid-IR imaging covering the complete disks of more than 2,300 galaxies in the local Universe.

Acknowledgments I thank my co-workers Ron Buta, Sébastien Comerón, Eija Laurikainen, and Heikki Salo for comments on an earlier draft of this chapter. I thank them and my other collaborators, including those on the S^{4}G team, for stimulating discussions and scientific progress.

References

1. Abraham, R.G., Merrifield, M.R., Ellis, R.S., Tanvir, N.R., Brinchmann, J. 1999, MNRAS, 308, 569
2. Aguerri, J.A.L., Méndez-Abreu, J., Corsini, E.M. 2009, A&A, 495, 491
3. Allard, E.L., Knapen, J.H., Peletier, R.F., Sarzi, M. 2006, MNRAS, 371, 1087

4. Athanassoula, E. 1992, MNRAS, 259, 345
5. Athanassoula, E. 1994, In: Mass-Transfer Induced Activity in Galaxies, ed. I. Shlosman, Cambridge University Press, Cambridge, p. 143
6. Athanassoula, E. 2002, ApJ, 569, L83
7. Athanassoula, E. 2003, MNRAS, 341, 1179
8. Athanassoula, E., Misiriotis, A. 2002, MNRAS, 330, 35
9. Barazza, F.D., Jogee, S., Marinova, I. 2008, ApJ, 675, 1194
10. Block, D.L., Buta, R., Knapen, J.H., Elmegreen, D.M., Elmegreen, B.G., Puerari, I. 2004, AJ, 128, 183
11. Block, D.L., Puerari, I., Knapen, J.H., Elmegreen, B.G., Buta, R., Stedman, S., Elmegreen, D.M. 2001, A&A, 375, 761
12. Bushouse, H.A. 1987, ApJ, 320, 49
13. Buta, R., Block, D.L. 2001, ApJ, 550, 243
14. Buta, R.J., Corwin, H.G., Odewahn, S.C. 2007, The de Vaucouleurs Atlas of Galaxies, Cambridge University Press, Cambridge
15. Buta, R.J., Knapen, J.H., Elmegreen, B.G., Salo, H., Laurikainen, E., Elmegreen, D.M., Puerari, I., Block, D.L. 2009, AJ, 137, 4487
16. Buta, R., Laurikainen, E., Salo, H. 2004, AJ, 127, 279
17. Buta, R., Laurikainen, E., Salo, H., Block, D.L., Knapen, J.H. 2006, AJ, 132, 1859
18. Buta, R.J., Laurikainen, E., Salo, H., Knapen, J.H. 2010, ApJ, 721, 259
19. Buta, R., Vasylyev, S., Salo, H., Laurikainen, E. 2005, AJ, 130, 506
20. Buta, R.J., et al. 2010, ApJS, 190, 147
21. Cameron, E., et al. 2010, MNRAS, in press
22. Combes, F. 2001, In: Advanced Lectures on the Starburst-AGN Connection, eds. I. Aretxaga, D. Kunth, R. Mújica, World Scientific, Singapore, p. 223
23. Combes, F., Gerin, M. 1985, A&A, 150, 327
24. Combes, F., Sanders, R.H. 1981, A&A, 96, 164
25. Comerón, S., Knapen, J.H., Beckman, J.E., Laurikainen, E., Salo, H., Martínez-Valpuesta, I., Buta, R.J. 2010, MNRAS, 402, 2462
26. Comerón, S., Martínez-Valpuesta, I., Knapen, J.H., Beckman, J.E. 2009, ApJ, 706, L256
27. Cox, T.J., Jonsson, P., Somerville, R.S., Primack, J.R., Dekel, A. 2008, MNRAS, 384, 386
28. Debattista, V.P., Sellwood, J.A. 1998, ApJ, 493, L5
29. de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H.G., Jr., Buta, R.J., Paturel, G., Fouque, P. 1991, Third Reference Catalogue of Bright Galaxies, Springer, Berlin
30. Di Matteo, P., Bournaud, F., Martig, M., Combes, F., Melchior, A.-L., Semelin, B. 2008, A&A, 492, 31
31. Di Matteo, P., Combes, F., Melchior, A.-L., Semelin, B. 2007, A&A, 468, 61
32. Durbala, A., Buta, R., Sulentic, J.W., Verdes-Montenegro, L. 2009, MNRAS, 397, 1756
33. Ellison, S.L., Patton, D.R., Simard, L., McConnachie, A.W., Baldry, I.K. 2010, MNRAS, in press
34. Elmegreen, D.M., Elmegreen, B.G. 1982, MNRAS, 201, 1021
35. Elmegreen, B.G., Elmegreen, D.M. 1985, ApJ, 288, 438
36. Elmegreen, B.G., Elmegreen, D.M., Hirst, A.C. 2004, ApJ, 612, 191
37. Elmegreen, B.G., Elmegreen, D.M., Knapen, J.H., Buta, R.J., Block, D.L., Puerari, I. 2007, ApJ, 670, L97
38. Eskridge, P.B., et al. 2000, AJ, 119, 536
39. Fazio, G.G., et al. 2004, ApJS, 154, 10
40. Heller, C.H., Shlosman, I. 1996, ApJ, 471, 143
41. Ho, L.C., Filippenko, A.V., Sargent, W.L.W. 1997, ApJ, 487, 591
42. Hubble, E.P. 1926, ApJ, 64, 321
43. Hunt, L.K., Malkan, M.A. 1999, ApJ, 516, 660
44. Jogee, S., et al. 2004, ApJ, 615, L105
45. Jogee, S., et al. 2009, ApJ, 697, 1971
46. Kapferer, W., Knapp, A., Schindler, S., Kimeswenger, S., van Kampen, E. 2005, A&A, 438, 87

47. Kennicutt, R.C., Lee, J.C., Akiyama, S., Funes, J.G., Sakai, S. 2005, AIPC, 783, 3
48. Knapen, J.H. 2005, A&A, 429, 141
49. Knapen, J.H., Beckman, J.E., Heller, C.H., Shlosman, I., de Jong, R.S. 1995, ApJ, 454, 623
50. Knapen, J.H., James, P. A. 2009, ApJ, 698, 1437
51. Knapen, J.H., Pérez-Ramírez, D., Laine, S. 2002, MNRAS, 337, 808
52. Knapen, J.H., Shlosman, I., Peletier, R.F. 2000, ApJ, 529, 93
53. Kormendy, J. 1979, ApJ, 227, 714
54. Kormendy, J., Norman, C.A. 1979, ApJ, 233, 539
55. Kormendy, J., Kennicutt, R.C., Jr. 2004, ARA&A, 42, 603
56. Laine, S., Shlosman, I., Knapen, J.H., Peletier, R.F. 2002, ApJ, 567, 97
57. Laurikainen, E., Salo, H. 2002, MNRAS, 337, 1118
58. Laurikainen, E., Salo, H., Buta, R. 2004, ApJ, 607, 103
59. Laurikainen, E., Salo, H., Buta, R. 2005, MNRAS, 362, 1319
60. Laurikainen, E., Salo, H., Buta, R., Knapen, J.H. 2007, MNRAS, 381, 401
61. Laurikainen, E., Salo, H., Buta, R., Knapen, J.H. 2009, ApJ, 692, L34
62. Laurikainen, E., Salo, H., Buta, R.J., Knapen, J.H., Comerón, S. 2010, MNRAS, 405, 1089
63. Laurikainen, E., Salo, H., Buta, R., Knapen, J., Speltincx, T., Block, D. 2006, AJ, 132, 2634
64. Laurikainen, E., Salo, H., Buta, R., Vasylyev, S. 2004, MNRAS, 355, 1251
65. Li, C., Kauffmann, G., Heckman, T.M., Jing, Y.P., White, S.D.M. 2008, MNRAS, 385, 1903
66. Lynden-Bell, D., Kalnajs, A.J. 1972, MNRAS, 157, 1
67. Marinova, I., Jogee, S. 2007, ApJ, 659, 1176
68. Martinet, L., Friedli, D. 1997, A&A, 323, 363
69. Martinez-Valpuesta, I., Knapen, J.H., Buta, R. 2007, AJ, 134, 1863
70. Martinez-Valpuesta, I., Shlosman, I., Heller, C. 2006, ApJ, 637, 214
71. Mazzuca, L.M., Swaters, R.A., Veilleux, S., Knapen, J.H. 2010, ApJ, submitted
72. Menéndez-Delmestre, K., Sheth, K., Schinnerer, E., Jarrett, T.H., Scoville, N.Z. 2007, ApJ, 657, 790
73. Mihos, J.C., Hernquist, L. 1996, ApJ, 464, 641
74. Moles, M., Márquez, I., Pérez, E. 1995, ApJ, 438, 604
75. Mulchaey, J.S., Regan, M. W. 1997, ApJ, 482, L135
76. Nair, P.B., Abraham, R.G. 2010, ApJ, 714, L260
77. Pease, F.G. 1917, ApJ, 46, 24
78. Prendergast, K.H. 1962, In: Distribution and Motion of ISM in Galaxies, ed. L. Woltjer, Benjamin, New York, NY, 217
79. Quillen, A.C., Frogel, J.A., Gonzalez, R.A.1994, ApJ, 437, 162
80. Regan, M.W., Teuben, P. 2003, ApJ, 582, 723
81. Rogers, B., Ferreras, I., Kaviraj, S., Pasquali, A., Sarzi, M. 2009, MNRAS, 399, 2172
82. Salo, H., Laurikainen, E. 2000, MNRAS, 319, 377
83. Salo, H., Laurikainen, E., Buta, R. J., Knapen, J.H. 2010, ApJ, 715, L56
84. Salo, H., Rautiainen, P., Buta, R., Purcell, G.B., Cobb, M.L., Crocker, D.A., Laurikainen, E. 1999, AJ, 117, 792
85. Sandage, A. 1961, The Hubble Atlas of Galaxies, Carnegie Institution of Washington, Washington, DC
86. Sarzi, M., Allard, E.L., Knapen, J.H., Mazzuca, L.M. 2007, MNRAS, 380, 949
87. Schwarz, M.P. 1984, MNRAS, 209, 93
88. Seigar, M.S., Chorney, N.E., James, P.A. 2003, MNRAS, 342, 1
89. Sellwood, J.A. 1981, A&A, 99, 362
90. Sellwood, J.A., Wilkinson, A. 1993, RPPh, 56, 173
91. Sheth, K., et al. 2008, ApJ, 675, 1141
92. Sheth, K., et al. 2010, PASP, submitted
93. Shlosman, I. 2001, ASPC, 249, 55
94. Shlosman, I., Begelman, M.C., Frank, J. 1990, Nature, 345, 679
95. Shlosman, I., Frank, J., Begelman, M.C. 1989, Nature, 338, 45

96. Smith, B.J., Struck, C., Hancock, M., Appleton, P.N., Charmandaris, V., Reach, W.T. 2007, AJ, 133, 791
97. van den Bergh, S., Abraham, R.G., Whyte, L.F., Merrifield, M.R., Eskridge, P.B., Frogel, J.A., Pogge, R. 2002, AJ, 123, 2913
98. Villa-Vargas, J., Shlosman, I., Heller, C. 2010, ApJ, 707, 218
99. Werner, M.W., et al. 2004, ApJS, 154, 1
100. Woods, D.F., Geller, M.J. 2007, AJ, 134, 527
101. Zhang, X., Buta, R.J. 2007, AJ, 133, 2584
102. Zheng, X.Z., Hammer, F., Flores, H., Assémat, F., Rawat, A. 2005, A&A, 435, 507

Bars and Bulges Through Masks of Time

Isabel Pérez, Patricia Sánchez-Blázquez, A. Zurita, G. Popping, Bard K. Gibson, and Pierre Ocvirk

Abstract We present here some of the results of a project devoted to understand the influence of bars in the evolution of galaxy disks through the study of their stellar content. First, we present the results on the derived stellar parameters for intermediate age populations, traced by Lick/IDS indices along the bar for a sample of 20 galaxies as well as the disk population parameters obtained from the whole spectra of 4 galaxies. Second, we present the results of the detail current and recent (<1 Gyr) star formation properties of two barred galaxies, NGC 2903 and NGC 1530. For the intermediate populations, we have obtained high S/N long-slit spectra of 20 barred galaxies, ranging in morphological types between S0 and Sb galaxies. We have derived Lick/IDS indices along the bar for all the sample galaxies. We find three types of bars according to their metallicity gradient: bars with negative metallicity gradient, bars with null metallicity gradient, and galaxies with positive metallicity gradient. Regarding the bulge region we find that the ages and metallicities are younger and more metal rich, at a given σ, than the bulges of unbarred galaxies. The bulges of the barred galaxies in our sample have suffered a different chemical enrichment history compared to unbarred galaxies of the same morphological types, same σ distribution, and same inclinations. We have also analyzed the kinematics star formation history of the disk region (2–3 disk scale lengths) for a sample of four barred galaxies. Most of the stellar mass in all the four galaxies is composed of old stars, this also true at large radii. However, a larger fraction of young stars is present in the external parts compared to the inner disk, compatible with a mild inside-out growing. The young stars all across the galaxy show lower metal content, favoring a scenario where stars are forming from recently accreted gas from the halo. Finally, we find some curious properties regarding the massive star forming regions on the two galaxies analyzed, NGC 2903 and NGC 1530. There are indications that massive stars are forming on the trailing side of the bar dust lane, and they age as they cross the bar of NGC 1530. We have found the presence of a large area of the bar of NGC 2903 containing older, with respect to the other star-forming regions present

I. Pérez (✉)
Departmento de Física Teórica y del Cosmos, Universidad Granada, Granada, Spain;
Instituto Carlos I de Física Teórica y Computación, Granada, Spain
e-mail: isa@ugr.es

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_19,

in the bar (a few hundred of Myrs) ultraviolet-emitting knots without significant Hα nor 24 μm emission, indicating regions that are not currently favored by the bar to form stars.

1 Introduction

During disk assembly, secular evolution must have played a role in shaping the structure of disk galaxies as we see them at $z = 0$. Non-axisymmetric structures, particularly bars, drive a substantial redistribution of the mass and angular momentum of the disk. This inflow of gas accumulates in central mass concentrations (CMC), fueling the central black hole (e.g. [37]) and, possibly, forming a stellar bulge. It is reasonable to expect that bars can be responsible for significant chemical evolution as well, since mixing by global flows will clearly change abundance profiles in the disk [14, 32, 34]. Several observational studies have found that barred galaxies show a shallower gas-phase metallicity gradient than non-barred ones (e.g. [9, 21, 26, 33, 40, 42]). However, an equivalent study in the stellar phase remains to be done.

Despite their obvious importance in disk galaxy evolution, the fate of bars is still a matter of debate. Several mechanisms are able to destroy the bar: for example (1) the bar may dissolve in the presence of a sufficiently massive central component (e.g. a black hole) [1, 2, 6, 7, 11, 12, 15, 25, 35]. However, the masses of the CMC needed to destroy the bar are much larger than those of the most super massive black holes found so far in disk galaxies [1, 10, 35]. (2) Transfer of angular momentum from infalling gas to the bar has also been proposed as a mechanism that could strongly weaken the bar. The combined effect of this angular momentum transfer and the CMC can destroy the bar in Sb–Sc galaxies in 1–2 Gyr [6]. Then, if gas is present, a new bar could form with characteristics different from those of the parent bar. However, simulations of the effect of gas on the stellar evolution of a bar embedded in a live halo [3] have shown that, although there are some structural differences in bars evolving with and without gas, in both cases the stellar bar can survive at least 5 Gyr (the total computation time of these simulations)

Do all bars go through these processes? Do some bars die while others are robust over many Hubble times?

It is important to study the evolution of the two components, gas and stars, suffer from very different evolutionary processes; the gas is mainly dominated by the gravitational torque of the non-axisymmetric mass component, while the evolution of the stellar component is mainly affected by different orbital mixing [5], so one would not expect the same abundance trends in both components.

1.1 Exploring Masks of Time Older Than 1 Gyr

Detailed analysis of the stellar populations in the bar region of local galaxies can shed some light on the formation and evolution of bars. Stellar age and metallicity

gradients provide a powerful link between a galaxy's star formation history and the dynamical processes operating within. They can give us archaeological clues as to the formation and evolution of the bar and to explore its influence on bulge formation and chemical redistribution in the disk. As already mentioned, bars are very likely to drive to some extend abundance profiles in the disk. Numerical simulations of abundance gradients in bars developed by Friedli et al. [13] predicted, using N-body simulations of bars with pre-existing exponential abundances, a null evolution of the stellar abundance profile (although with a decrease in the mean metallicity) while the gas abundance profile flattened rapidly.

First attempts to study stellar populations in bars have focused on broadband optical images. Gadotti and de Souza [16] obtained the color and color gradients in the bar region of a sample of 18 barred galaxies. They interpreted the color differences as differences in stellar ages. They concluded that younger bars were hosted by galaxies of later types. However, the conclusions are hampered by the assumption that the color traces the age, and that the metallicity effect or dust extinction is not important. Clearly, a study attempting to break the age-metallicity degeneracy and taking into account the effect of dust is needed. In addition, the influence of dust in the line-strength indices has been shown to be minimum [20] while ages and metallicities from broadband colors are heavily affected by dust as well as by the age-metallicity degeneracy.

To date, it has been difficult to obtain stellar population parameters along the bar because a high signal-to-noise ratio is required for this analysis. Although bars are high surface brightness structures, this still implies long integration times on medium-size telescopes. Recent work has presented the radial distribution of line-strength indices along bars [28] for six galaxies.

We have started a long-term project to analyze the stellar content on the bar region and the disk of barred galaxies. To this aim, we first obtained kinematics, mean ages, metallicities, and chemical abundances along the bar of 20 galaxies with morphological types from SB0 to SBbc. Second, we have derived the kinematics, star formation history, mass and luminosity weighted, as well as single stellar population equivalent, ages and metallicities for a sample of four galaxies with morphological types ranging from S0 to Sc. The results from this work will be presented in Sections 2 and 3.

1.2 Exploring Masks of Time Younger Than 1 Gyr

While the study of stellar populations gives us archaeological clues on the disk, bar, and bulge history, studying the location and properties of recent and current massive star formation can help us understand the processes of star formation and mass redistribution triggered by a bar potential.

In the years after the launch of the *Spitzer Space Telescope* and the *Galaxy Evolution Explorer* (*Galex*), knowledge on the star formation (SF) in galaxies has grown considerably. The panchromatic view of nearby galaxies offered by large

surveys carried out from these space telescope missions gives an extraordinary database to study star formation in galaxies. In particular, it allows us to link general galaxy properties with the local interstellar medium (ISM) properties and the galaxy dynamics. These links are crucial for understanding SF in galaxies. In nearby galaxies these sets of multi-wavelength data give the opportunity to study in detail, with good spatial resolution, the location and properties of relatively young populations, recent massive star formation, and dust attenuation (e.g. [8, 31, 38]).

New dust and SF indicators have been developed and calibrated using all the available bands. The combination of 24 μm and Hα emission has become a reliable indicator of dust attenuation [19, 31]. Both bands are linked to star formation processes, with the Hα emission originating from the recombination of hydrogen in the surrounding medium of very recently formed (less than a few Myr) massive stars, and the 24 μm emission as tracing local star formation radiation obscured by dust [8]. Extinction-corrected UV emission can also be used to retrieve the star formation rate (SFR) [18] and to get some insight into the star formation occurred in last gigayear (e.g. [4, 17]). This feature makes a combination of UV, Hα, and 24 μm emission ideal to reconstruct the recent star formation history in a galaxy by studying the location and properties of individual clusters and those of the gas and dust emission around them.

A panchromatic view revealing the history of star formation in bars can give a unique insight into the links of star formation and the galaxy dynamics, which could also help to understand how bars form and evolve.

We have started a project to investigate the *recent* and *current* star formation in barred galaxies by detail analysis of multi-wavelength data. We present in Section 4 the results of the analysis of two nearby galaxies, NGC 2903 and NGC 1530.

2 Stellar Line-Strength Indices of Bars and Bulges

We have obtained high S/N long-slit spectra of 22 barred galaxies, with morphological types between S0 and Sb galaxies. We have obtained Lick/IDS indices along the bar for all the sample galaxies. We presented in Pérez et al. [29] the detailed derivation of the indices and the stellar population parameters. We refer the reader to this work for a further reading on the technical details and the analysis of the populations along the bar region. We present in this chapter a summary of the results of the analysis of the stellar population on the bar region and new results of the study of the stellar parameters of the bulge region. The extended version of the work on the bulge region will be shortly published.

2.1 Metallicity Gradient Along the Bar

According to the metallicity gradient in the bar we find three types of bars (see Fig. 1):

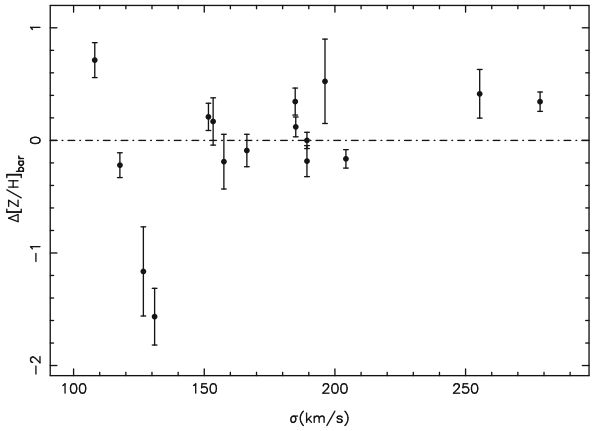


Fig. 1 Metallicity difference between the inner and the outer bar region as a function of maximum central velocity dispersion. Notice that galaxies with $\sigma < 170\,\mathrm{km\ s^{-1}}$ tend to have negative gradient values

- Bars with negative metallicity gradient (more metal poor at the ends of the bar): These bars show a mean young/intermediate population and have among the lowest velocity dispersion. They also tend to have positive age gradients, circumnuclear regions with young population. Since we expect the original metallicity gradient to flatten, this might be indicating that these bars formed recently.
- Bars with null metallicity gradients: These bars show gradients both in age and metallicity compatible with the models.
- Bars with positive metallicity gradients (more metal rich at the ends of the bar): These bars show older mean ages and higher velocity dispersion. Since the diffusion timescale is short [41], they are not likely to be a relic from the initial disk gradient. There is still not a conclusive explanation for the origin of this gradient yet. It is, however, possibly indicating an old bar.

2.2 Mean Values and Metallicity Gradients in the Bulge Region

The ages and metallicities derived show that the bulges of barred galaxies are younger and more metal rich, at a given σ , than the bulges of unbarred galaxies (comparing our sample to that of [24]). The E/Fe derived for the bulges of barred galaxies tend to be above the values of the unbarred galaxies at a given σ. The gradients of the metal indices in the bulge region tend to be negative, less metal rich toward the end of the bulge; the gradient values show a large scatter for galaxies with σ below 150 km/s^{-1}. The Balmer indices show the opposite trend with the scatter of the gradients also increasing for galaxies with σ below 150 km/s^{-1}.

In summary, the bulges of the barred galaxies in our sample have suffered a different chemical enrichment history compared to unbarred galaxies of the same

morphological types, same σ distribution, and same inclinations. This fact should be taken into account when doing statistical studies of central ages and metallicities and comparing the results to model predictions.

3 Stellar Content in the Disk Region of Barred Galaxies

We present here the kinematics, star formation history, mass weighted and luminosity weighted, and single stellar population equivalent ages and metallicities for a sample of four galaxies with types from S0 to Sc. The data extend to 2–3 scale lengths of the disk with S/N(Å)>50. Several techniques are explored to derive star formation histories and single stellar population equivalent parameters and all of them are compatible as far as they are interpreted properly. Our results are robust to the choices of stellar population models. The first results from this study indicate that the majority of the stellar mass in all galaxies is composed of old (around 10 Gyr) stars (see Fig. 2). This is true, not only in the bulge regions but also in the disk, at radius larger than 2 scale lengths. However, a larger fraction of young stars is present in the external parts of the disk compared with the inner disk. The metallicity gradients in the disk region are also very mild, flat, or negative in all cases. The young, dynamically cold, structures produced by the presence of the bar, as nuclear disks or rings are the responsible for shaping the bulge age and metallicity gradient, as suggested by Peletier et al. [27]. The young stars all across the galaxy show a lower metal content, favoring a scenario where stars are forming from recently accreted gas.

4 Massive Recent and Current Star Formation in Barred Galaxies

The places and parameters which provide suitable conditions for forming stars in bars are not yet well understood. Most of our current knowledge on the location, distribution, and properties of star-forming regions in bars comes from a handful of papers centered on their physical properties [23] and on the general morphology of the Hα emission (e.g. [30, 39]) and on its relation with the molecular gas (e.g. [36]) or with the stellar bar [22, 39]. We present below a detailed multi-wavelength analysis of the SF in the nearby barred galaxies NGC 1530 and NGC 2903.

4.1 NGC 1530

NGC 1530 has one of the strongest bars ever observed and recent star formation sites are distributed across its bar. Our aim is to study the photometric properties of the bar and its H II regions, to elucidate the conditions under which H II regions form and their spatial relation to the principal dynamical features of the bar.

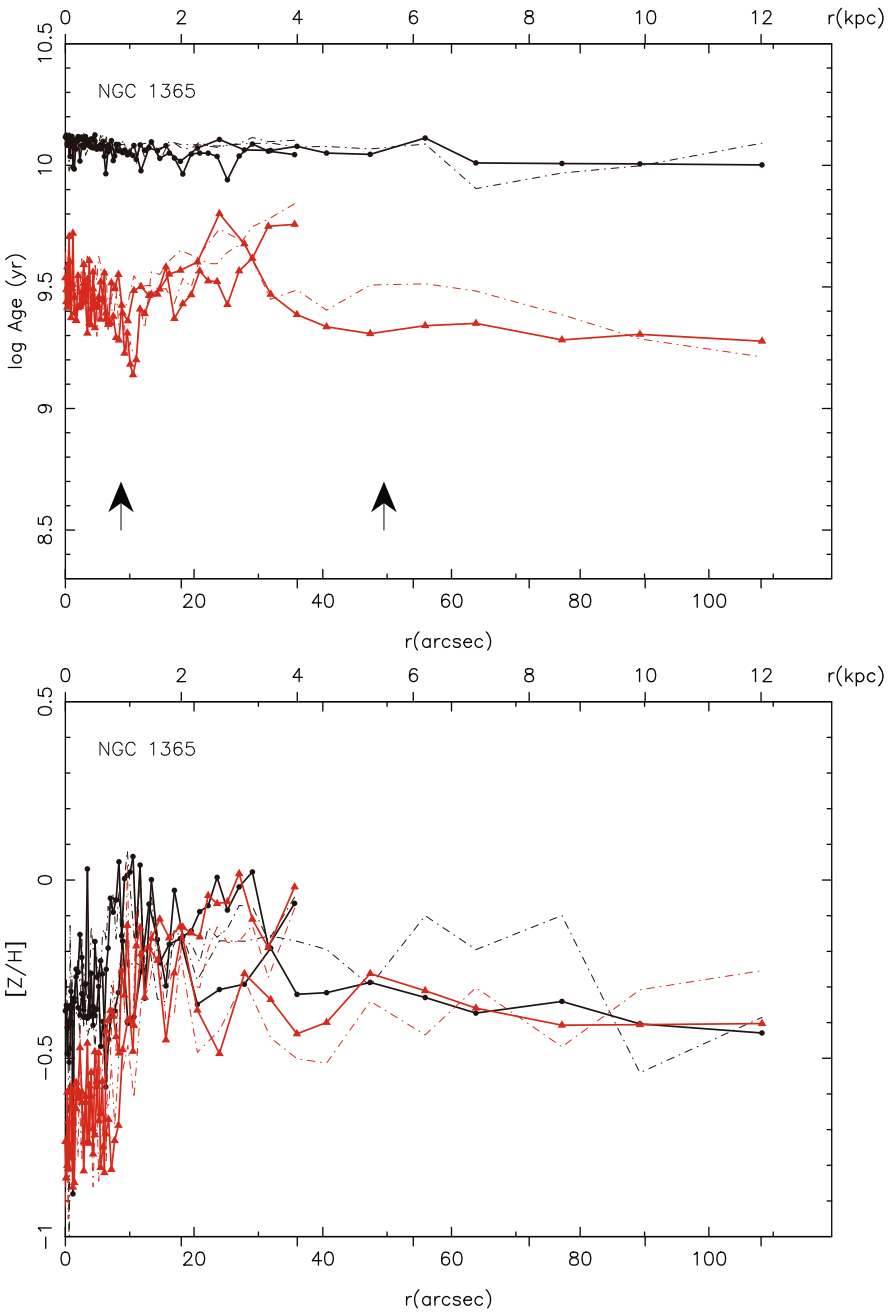


Fig. 2 Mass- and luminosity-weighted age and metallicity gradients derived from the recovered star formation history of the disk of NGC 1365. Different *lines* represent the solutions obtained with two different stellar populations models. The *arrows* at the *upper panel* show the effective radius of the bulge (*left-most arrow*) and the disk scalelength

We have obtained BVRKs and Hα photometry of the H II regions of the bar of NGC 1530. Broadband integrated colors and Hα equivalent widths have been carefully measured and analyzed as a function of their position with respect to the main dust lanes of the bar. We have found differences in the Hα equivalent widths of the H II regions located in the trailing and leading sides of the bar dust lane (see Fig. 3). The possible factors which could produce that difference ([NII] contamination, Lyman continuum photons dust extinction, escape of ionizing radiation, metallicity, IMF, age) have been carefully analyzed. The age has been confirmed as the most plausible explanation. This implies that H II regions located further away from the bar dust lane in its leading side (downstream from the main bar dust lane) are older than the rest by 1.5 − 2.5 Myr. In addition, a clear spatial correlation has been found between location of H II regions, dust spurs on the trailing side of the bar dust lane, and the loci of maximum velocity gradients parallel to the bar major axis (possibly tracing gas flow toward the main bar dust lane, see Fig. 4). These results support the hypothesis that massive stars are forming on the trailing side of the bar dust lane, and they age as they cross the bar, in a timescale which is compatible with the bar dynamical timescale. The detail analysis has been presented in Zurita and Pérez [43].

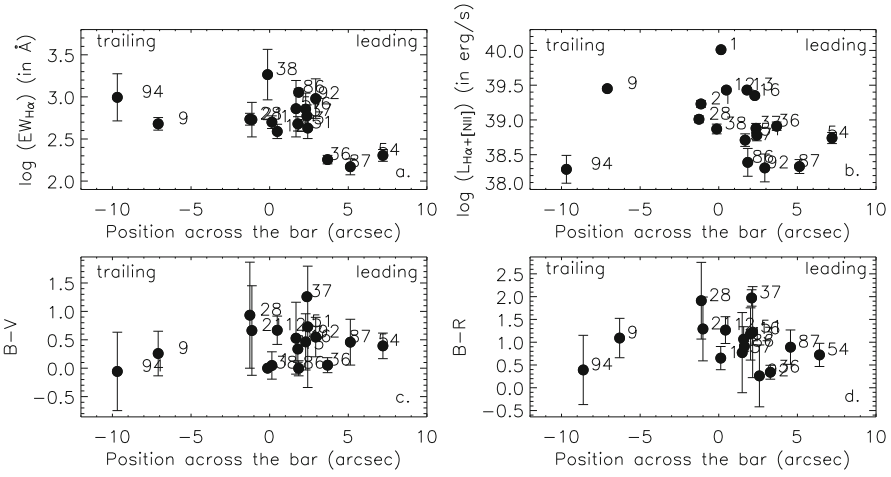


Fig. 3 Representation of (from *top* to *bottom* and from *left* to *right*) the Hα equivalent width, Hα luminosities, and $B - V$ and $B - R$ broadband colors of the bar H II regions of NGC 1530 as a function of their deprojected distance from the bar dust lane in arcseconds. Zero means that the H II region is located on the bar dust lane, negative and positive values refer to H II regions located on the trailing and leading side of the bar dust lane, respectively

4.2 NGC 2903

NGC 2903 is a nearby barred spiral with an active starburst in the center and H II regions distributed along its bar. We aim to analyze the star formation properties in

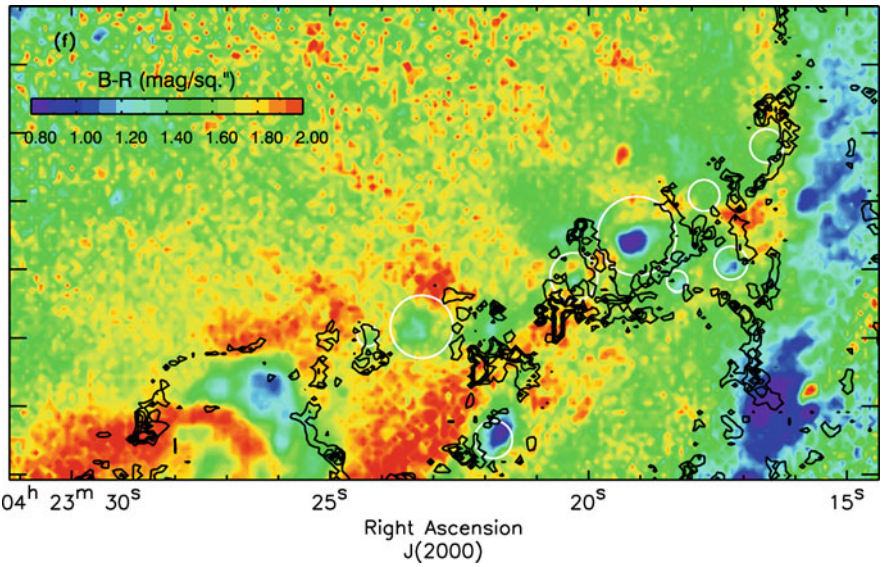


Fig. 4 $B - R$ color map of the NW side of the bar with overlaid contours of the velocity gradients in the direction of the bar. Notice the spatial correlation between the loci of maximum velocity gradient in the direction parallel to the bar major axis and the dust spurs in the trailing side of the bar (the information on the velocity gradients is spatially limited to zones with detectable Hα emission)

the bar region of NGC 2903 and study the links with the typical bar morphological features. A combination of space- and ground-based data from the far-ultraviolet to the sub-millimeter spectral ranges is used to create a panchromatic view of the NGC 2903 bar. We produce two catalogs: one for the current star formation regions, as traced by the Hα compact emission, and a second one for the ultraviolet (UV)-emitting knots, containing positions and luminosities. From them we have obtained ultraviolet colors, star formation rates, dust attenuation, and Hα EWs, and their spatial distribution has been analyzed. Stellar cluster ages have been estimated using stellar population synthesis models (*Starburst99*). NGC 2903 is a complex galaxy, with a very different morphology on each spectral band. The CO ($J = 1 - 0$) and the 3.6 μm emission trace each other in a clear barred structure, while the Hα leads both components and it has an s-shape distribution. The UV emission is patchy and does not resemble a bar. The UV emission is also characterized by a number of regions located symmetrically with respect to the galaxy center (see Fig. 5), almost perpendicular to the bar, in a spiral shape covering the inner $\sim$ 2.5 kpc. These regions do not show a significant Hα or 24 μm emission. We have estimated ages for these regions ranging from 150 to 320 Myr, being older than the rest of the UV knots, which have ages lower than 10 Myr. The SFR calculated from the UV emission is $\sim$0.4 $M_\odot$ yr^{-1}, compatible with the SFR as derived from Hα calibrations ($\sim$1 $M_\odot$ yr^{-1}).

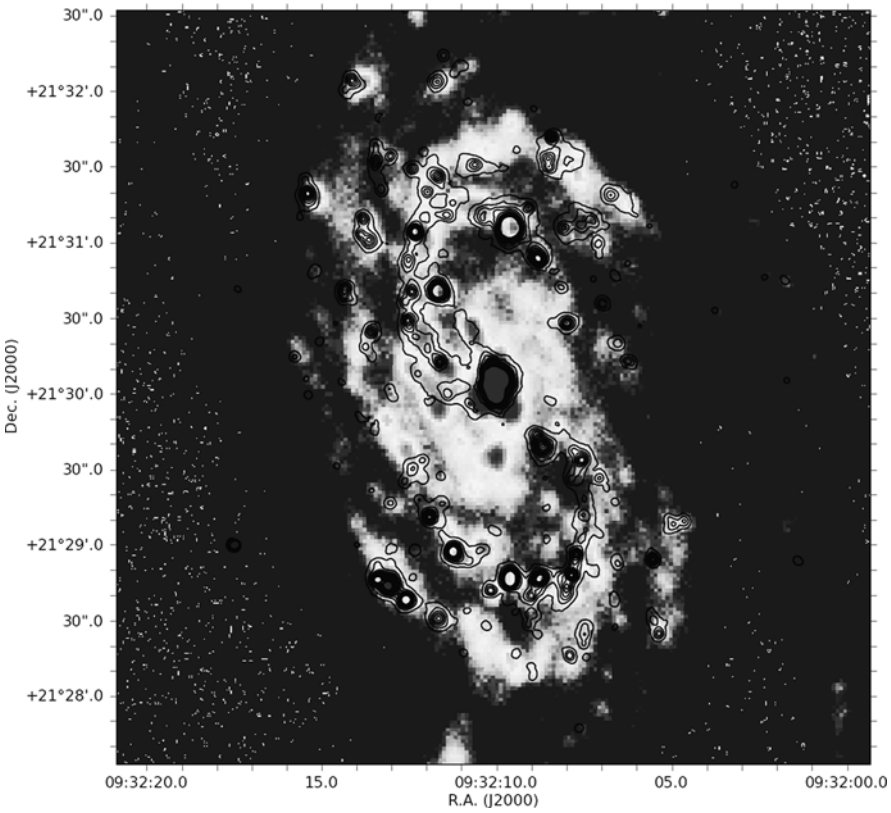


Fig. 5 FUV image of the NGC 2903 bar region with Hα contours overplotted. Notice the regions NW and SE of the center with bright UV emission and no Hα counterparts

Acknowledgments First of all, I thank the organisers for an amazing meeting held at an extraordinary location. I specially want to thank Ken for being such an inspiration to do science throughout my career. Pérez acknowledges support by the Netherlands Organisation for Scientific Research (NWO, Veni-Grant 639.041.511). I. Pérez and A. Zurita acknowledge support from the Spanish Plan Nacional del Espacio de Ministerio de Educación y Ciencia (via grant C-CONSOLIDER AYA 2007-67625-C02-02) . I. Pérez and A. Zurita also thank the Junta de Andalucía for support through the FQM-108 project. P.S.B. is supported by a Ramon y Cajal fellowship. P.S.B. also acknowledges a ERC within the 6th European Community Framework Programme

References

1. Athanassoula, E., Lambert, J.C., Dehnen, W. 2005, MNRAS, 363, 496
2. Berentzen, I., Heller, C.H., Shlosman, I., Fricke, K.J. 1998, MNRAS, 300, 49
3. Berentzen, I., Shlosman, I., Martinez-Valpuesta, I., Heller, C.H. 2007, ApJ, 666, 189
4. Bianchi, L., et al. 2005, ApJ, 619, L71
5. Binney, J. Tremaine, S. 1987, Galactic Dynamics, Princeton University Press, Princeton, NJ
6. Bournaud, F., Combes, F. 2002, A&A, 392, 83

7. Bournaud, F., Combes, F., Semelin, B. 1995, MNRAS, 364, L18
8. Calzetti, D., et al. 2005, ApJ, 633, 871
9. Edmunds, M.G., Roy, J.-R. 1993, MNRAS, 261, L17
10. Ferrarese, L., Ford, H. 2005, SSRv, 116, 523
11. Friedli, D., Pfenniger, D. 1991, Dynamics of Galaxies and Their Molecular Cloud Distributions, IAU Symp. 146, Kluwer, Dordrecht, p. 362
12. Friedli, D., Benz, W. 1993, A&A, 268, 65
13. Friedli, D., Benz, W., Kennicutt, R. 1994, ApJ, 430, L105
14. Friedli, D. 1998, The Central Regions of the Galaxy and Galaxies, IAU Coll. 184, ASP, San Francisco, CA, p. 269
15. Fukuda, H., Habe, A., Wada, K. 2000, ApJ, 529, 109
16. Gadotti, D.A., de Souza, R.E. 2006, ApJS, 163, 270
17. Hibbard, J.E., et al. 2005, ApJ, 619, L87
18. Kennicutt, R.C., Jr. 1998, ApJ, 498, 541
19. Kennicutt, R.C., Jr., et al. 2007, ApJ, 671, 333
20. MacArthur, L.A. 2005, ApJ, 623, 795
21. Martin, P., Roy, J.-R. 1994, ApJ, 424, 599
22. Martin, P., Friedli, D. 1997, A&A, 326, 449
23. Martin, P., Friedli, D. 1999, A&A, 346, 769
24. Moorthy, B.K., Holtzman, J.A. 2006, MNRAS, 371, 583
25. Norman, C.A., Sellwood, J.A., Hasan, H. 1996, ApJ, 462, 114
26. Pagel, B.E.J., Edmunds, M.G. 1981, ARAA, 19, 77
27. Peletier, R.F., et al. 2007, MNRAS, 379, 445
28. Pérez, I., Sánchez-Blázquez, P., Zurita, A. 2007, A&A, 465, L9
29. Pérez, I., Sánchez-Blázquez, P., Zurita, A. 2009, A&A, 495, 775
30. Phillips, A.C. 1996, Barred Galaxies, IAU Colloq. 157, ASP, San Francisco, CA, vol. 91, p. 44
31. Relaño, M., Kennicutt, R.C. 2009, ApJ, 699, 1125
32. Roşkar, R., Debattista, V.P., Stinson, G.S., Quinn, T.R., Kaufmann, T., Wadsley, J. 2008, ApJ, 675, L65
33. Ryder, S.D. 1995, ApJ, 444, 610
34. Sánchez-Blázquez, P., Courty, S., Gibson, B.K., Brook, C.B. 2009, MNRAS, 398, 591
35. Shen, J., Sellwood, J.A. 2004, ApJ, 604, 614
36. Sheth, K., Vogel, S.N., Regan, M.W., Teuben, P.J., Harris, A.I., Thornley, M.D. 2002, AJ, 124, 2581
37. Shlosman, I., Frank, J., Begelman, M.C. 1989, Nature, 338, 45
38. Tamura, K., Jansen, R.A., Windhorst, R.A. 2009, AJ, 138, 1634
39. Verley, S., Combes, F., Verdes-Montenegro, L., Bergond, G., Leon, S. 2007, A&A, 474, 43
40. Vila-Costas, M.B., Edmunds, M.G. 1992, MNRAS, 259, 121
41. Wozniak, H. 2007, A&A, 465, L1
42. Zaritsky, D., Kennicutt, R.C., Jr., Huchra, J.P. 1994, ApJ, 420, 87
43. Zurita, A., Pérez, I. 2008, A&A, 485, 5


Strength of the
NIGHT

Tidal Trails and Mass-Segregated Isothermal Clusters

Donald Lynden-Bell

Abstract New analytic approximations for central orbits show that in Kepler's problem the variation of the fourth power of the radius is sinusoidal in time. Related approximations are used to calculate the tidal trails generated in a logarithmic potential at the pericentric passage of a small victim. The structure of the simplest mass-segregated isothermal clusters is that of polytropes of negative index. These are at Maxwell–Boltzmann equilibrium under two body encounters but they can still evolve, either by multiple encounters to form binary or multiple stars or via Antonov's gravothermal instability.

1 Introduction

Ken Freeman's prolific output has centred on cosmic archaeology [7, 8], the dynamics and chemical evolution of galaxies and globular clusters. While he is best known for his observational work and its interpretation, he is also rightly famous for his theoretical work on barred spirals [5, 6] which earned him the distinction of a research fellowship at Trinity College Cambridge.

To celebrate his remarkable career I present here two simple theoretical investigations on subjects that have interested him. The first is an analytical treatment of orbits in any central potential. The powerful approximation allows an analytical treatment of tidal streams in central potentials. The second discusses mass-segregated isothermal clusters and asks under what circumstances the core will be dominated by a single massive binary as discovered in some of Aarseth's large N-body experiments [1] and earlier in the small clusters studied by Van Albada [23].

Most work on tidal streams has followed the Toomres' pioneering work [22] by treating the whole problem computationally. However, it is often hard to understand exactly what parameters of the model determine the different aspects of the

D. Lynden-Bell (✉)
Institute of Astronomy, The Observatories, Cambridge CB3 0HA, UK
e-mail: dlb@ast.cam.ac.uk

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_20,

outcome. Thus it took the experience and insight of Alar Toomre to explain to me why the model of the Magellanic stream that Doug Lin computed [13] only tore on its third passage past the galaxy. In a V=const logarithmic potential the direction to pericentre moves around the plane of the orbit. For the orbit we had computed it was only at the third pericentric passage that the spin plane of the Magellanic Cloud pointed at the galactic centre. The greater extent of the cloud in its spin plane made it easier to tear then. If instead of having to compute every case anew there were formulae giving the shapes of the tidal streams, one might see how to adjust parameters from a study of the formulae. The new approximation methods I have developed to give analytical orbits can be applied to give the velocities and shapes of tidal streams in such potentials.

2 Analytic Approximation to Orbits

The orbital equations in a central potential are given by energy and angular momentum conservation so,

$$\frac{1}{2}(\dot{r}^2 + h^2/r^2) - \psi(r) = E \tag{1}$$

$$r^2\dot{\phi} = h \tag{2}$$

It is very important to treat E and h as functions of the pericentric and apocentric distances r_p and r_a rather than vice versa. Since $\dot{r}$ is zero at such turning points, we may solve Eq. (1) written first at $r = r_p$ and then at $r = r_a$ to find E and h^2 as follows:

$$E = -\frac{r_a^2\psi_a - r_p^2\psi_p}{r_a^2 - r_p^2} \tag{3}$$

$$h^2 = 2\frac{\psi_p - \psi_a}{r_p^{-2} - r_a^{-2}} \tag{4}$$

Whereas these are simple, the finding of the turning points of an orbit in a given potential for known energy and angular momentum is often awkward. [For Keplerian orbits Eqs. (3) and (4) reduce to $E = -GM/(r_a + r_p)$ and $h^2 = 2GM/(r_p^{-1} + r_a^{-1})$.] For a general central potential the orbital equation arises as follows:

$$d\phi = hr^{-2}dt = hr^{-2}dr/\dot{r} = hr^{-2}dr/(2E + 2\psi - h^2r^{-2})^{1/2}. \tag{5}$$

In the Keplerian case we often use r^{-1} as a new variable but in the general case we shall use $u = r^{-k}$ in place of r. We shall decide later on the most appropriate value of k for our orbit. Our orbital equation (5) now becomes

$$-kd\phi = [S(u)]^{-1/2}du \tag{6}$$

where

$$S(u) = (2E + 2\psi - h^2/r^2)h^{-2}r^2u^2 \tag{7}$$

and of course $r = u^{-1/k}$. If $S(u)$ were quadratic in u [as in the Keplerian case] then the integration of Eq. (6) would be easy. We therefore aim to choose k in such a manner that $S(u)$ can be well approximated by a quadratic in u. Now it is clear from (7) that $S(u)$ is zero at the turning points and from (6) we see that the integral will be sensitive to S at those points. We therefore choose $S_q(u)$, our quadratic approximation to S, to have the same zeros as $S(u)$. Of course we are only interested in making S_q a good approximation to S between those zero points since that is where the orbit lies. Thus $S_q(u)$ has to take the form

$$S_q = q^2(u_p - u)(u - u_a) \tag{8}$$

This is a downward pointing parabola, so its gradients at its zeros are equal and opposite. We now choose k so that this is also true of the exact expression $S(u)$. Since $(2E + 2\psi - h^2/r^2)$ is zero at the end points the only contribution to the gradient of S there comes when that bracket is differentiated, so $S'(u_p) = 2(\psi'_p + h^2/r_p^3)(d\ln r/d\ln u)_p h^{-2}r_p^3u_p = -2(r_p^3\psi'_p + h^2)/(kr_p^k)$ dividing this by a similar expression at apocentre and asking that the ratio be -1 to give opposite gradients, we find on taking logarithms

$$k = \ln\left[-\frac{r_p^3\psi'_p + h^2}{r_a^3\psi'_a + h^2}\right] / \ln[r_p/r_a] \tag{9}$$

where of course h^2 is given by (4). We see now that in general k will depend on the orbit considered whereas for the Kepler case it is one for all orbits and for the simple harmonic case it is two. For gravitational problems k lies between those limits. With k so chosen, plots of $S(u)$ are not far from downward facing parabolae. We set $\bar{u} = \frac{1}{2}(u_p + u_a)$ and define $l = \bar{u}^{-1/k}$ and $e = \frac{u_p - u_a}{u_p + u_a} = \frac{r_a^k - r_p^k}{r_a^k + r_p^k}$ by analogy to the semi-latus rectum and eccentricity of an ellipse. Notice that this eccentricity only agrees with the normal one for Kepler's case, $[k = 1]$. With these definitions we have $u_p = (1 + e)\bar{u}$ and $u_a = (1 - e)\bar{u}$. Furthermore we choose q^2 by requiring that S_q and S have the same value when $u = \bar{u}$. This gives $q^2 = S(\bar{u})/(e^2\bar{u}^2)$. We now return to Eq. (6) and approximate S by S_q. The integration over u then becomes $\int_{u_a}^{u}[\bar{u}^2e^2 - (u - \bar{u})^2]^{-1/2}du/q$

If we now use the substitution $u = \bar{u}(1 + e\cos\eta)$, the expression above is just η/q so the orbit is described by $u = \bar{u}[1 + e\cos(kq\phi)]$, or in terms of r and ϕ,

$$(l/r)^k = 1 + e\cos(kq\phi) \tag{10}$$

which is remarkably close to Kepler's $l/r = 1 + e\cos\phi$. In most cases expression (10) gives a good approximation to the orbit with errors of only 1 or 2%; however, there are times when the plot of $S(u)$ against u is somewhat skew with the maximum displaced from $\bar{u}$. In such cases q is best estimated off the true maximum of S found via Newton–Raphson starting from $\bar{u}$. It is also possible to estimate a new value of k such that the maximum comes at the mean of the new u_a and u_p. We give details of these refinements in the appendix but they are seldom needed for the equation of the orbit. Even when k and q are so refined the equation for the orbit still has the form (10) with revised values of k, q, e, l.

The orbit is not wholly determined by its shape. The radial action is given by J_r where

$$2\pi J_r = \int \dot{r}dr = \int (2E+2\psi-h^2/r^2)^{1/2} r d\ln r = hk^{-1}\int [S(u)]^{1/2}u^{-2}du \quad (11)$$

If we again approximate $S(u)$ by $S_q(u)$ [16] then we can evaluate the action as $2\pi J_r/h = qk^{-1}\int e^2\sin^2\eta(1+e\cos\eta)^{-2}d\eta$, so $J_r = [(1-e^2)^{-1/2}-1]hq/k$. The radial period and the angle between the apses can then be obtained by differentiation. Back in 1963 I defined an adiabatically invariant eccentricity, $e*$, of any central orbit by $1-e*^2 = [1+J_r/h]^{-2}$ [14]. These two eccentricities e and $e*$ become equal for Keplerian orbits as for them the factor k/q, applied to J_r to get e above, is unity. In practice the u^{-2} weighting in the integral for J_r means that our approximation is a little less accurate than it was for the spatial orbit. The time to a given point in the orbit is better estimated directly as in the next section.

3 The Orbital Time to Reach a Given Point

From the energy equation (1) we have, writing $U = r^K$

$$dt = \dot{r}^{-1}dr = [(2E+2\psi-h^2/r^2)r^{-2}]^{-1/2}d\ln r = K^{-1}[\tilde{S}(U)]^{-1/2}dU \quad (12)$$

where $\tilde{S} = (2E+2\psi-h^2/r^2)U^2/r^2$ and $r = U^{1/K}$. Again our aim is to approximate $\tilde{S}$ by a quadratic $S_Q(U)$ with the same zeros, corresponding to the turning points. Such a quadratic must take the form $S_Q(U) = Q^2(U_a-U)(U-U_p)$, which is a downward-pointing parabola. Again we choose K so that $\tilde{S}$ has equal and opposite gradients at the turning points. By a similar argument this leads to

$$4-K = \ln\left[-\frac{r_p^3\psi_p'+h^2}{r_a^3\psi_a'+h^2}\right]/\ln[r_p/r_a] \quad (13)$$

Comparing this with Eq. (9) we see that $K = 4-k$. This leads to $\tilde{S}(U) = (2E+2\psi-h^2/r^2)r^{-2}U^2$ with $r = U^{1/K}$. Remembering that $k = 1$ for all Kepler orbits this suggests $K = 3$ for all such orbits; however, in what follows we show that

$K = 4$ gives significantly closer agreement as demonstrated by the good fit in Fig. 1. Thus the variation of r^4 in any Keplerian ellipse is approximately harmonic in time. In fact that value fits best for orbits with $r_a/r_p = 10,\ k = 3.4$ fits better at $r_a/r_p = 3$, and for smaller eccentricities there is a degeneracy, so almost any such power can be made to fit. We shall write $\bar{U} = (U_a + U_p)/2 = L^K$ and $\tilde{e} = \frac{(r_a^K - r_p^K)}{(r_a^K + r_p^K)}$, then $U_a = (1 + \tilde{e})\bar{U}$ and $U_p = (1 - \tilde{e})\bar{U}$ so $S_Q = Q^2[(\tilde{e}\bar{U})^2 - (U - \bar{U})^2]$. In (12) we approximate $\tilde{S}$ by S_Q and require that they are equal when $U = \bar{U}$, so we find the value of Q from $Q^2\tilde{e}^2 = [2E + 2\psi(L) - h^2/L^2]/L^{-2}$. In Eq. (12) we then set $U = \bar{U}(1 - \tilde{e}\cos\tilde{\eta})$ and find

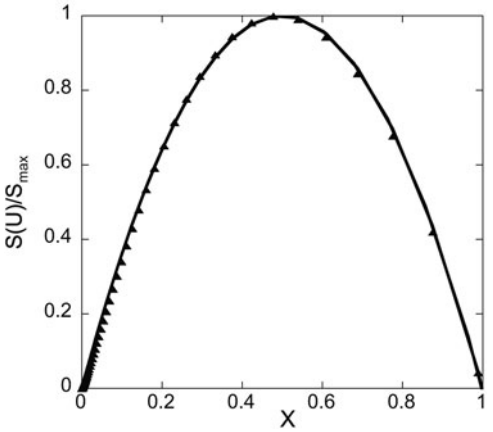


Fig. 1 The function $\tilde{S}(U)$ is calculated for a Kepler orbit with $r_a/r_p = 10$. *Triangles* show $\tilde{S}(U)/\tilde{S}(\bar{U})$, which is well approximated by the continuous parabola. Here $U = r^K$ and $K = 4$. Thus for this Kepler orbit the variation of r^4 is rather accurately sinusoidal in time. X is just U re-zeroed and re-scaled into [0,1]. The orbits of figure 2 in the logarithmic potential have $K = 3.1$ so for them the variation of $r^{3.1}$ is sinusoidal in time

$$QKt = \int_{U_p}^{U} [\tilde{e}^2\bar{U}^2 - (U - \bar{U})^2]^{-1/2} dU = \tilde{\eta} \tag{14}$$

hence

$$U/\bar{U} = (r/L)^K = 1 - \tilde{e}\cos(QKt) \tag{15}$$

This determines the r achieved at a given time t after pericentric passage. $L, K, Q, \tilde{e}$ all depend on r_a and r_p, although for power law potentials K and $\tilde{e}$ depend only on their ratio, and the dependence of K is there so weak that it may be replaced by a suitable average. For the harmonic potential, $K = 2$ exactly for all orbits, so then $K = 4-k$ is also exact. However, in general at high eccentricities this method of determining K from equal and opposite gradients at the turning points gives lopsided graphs of $\tilde{S}(U)$. The reason behind this is that near pericentre its gradient is rising as U increases whereas it falls for a parabola. Thus the graph

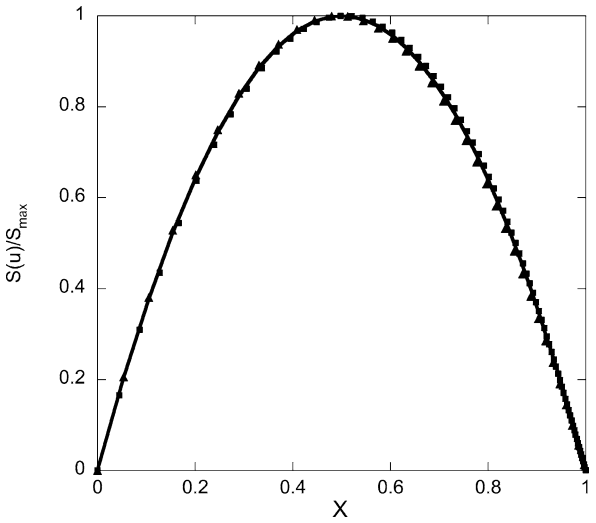


Fig. 2 $S(u)$ is calculated for two orbits in the logarithmic potential. $S(u)/S(\bar{u})$ is plotted as *triangles* for the orbit with $r_a/r_p = 3$ and as *squares* for the orbit with $r_a/r_p = 10$. Both are plotted with $u = r^{-k}$ and $k = 1.44$. The continuous parabola is a good fit to both, showing that a single k works well for both and that Eq. (10) fits the orbits well. X is u re-zeroed and re-scaled for each orbit

rises too fast near pericentre and becomes lopsided towards small radii. In such cases calculation of Q from setting $\tilde{S}$ and S_Q equal at $\bar{U}$ gives a less accurate value than can be obtained from setting their maxima equal. Methods of obtaining higher accuracies for both K and Q are explored in the appendix. Normally K is bigger than $4-k$. Whereas these refinements are not needed for the shape of the orbit, they are needed for the time in highly eccentric orbits. However, we still find that K and k do not vary much, so it is often sufficient to keep one average value for orbits in one power law potential. This was done in [18] where we showed that $k = 2 - \alpha$ is exact for "parabolic" $e = 1$ orbits in potentials of the form $\psi \propto r^{-\alpha}$. For the $V = const$ logarithmic potential suitable average values are $k = 1.44$, $K = 3.1$; see Fig. 2.

4 Application to Tidal Trails

To give a first cut at the problem we shall assume that (1) all the particles streaming from a single tidal encounter are released close to a pericentre but their individual orbits have differing eccentricities and therefore differing periods. These period differences spread them around the sky. Of course we may see the results of disruptions that took place at two or more pericentric passages which will have occurred at different azimuths in the orbital plane as the potential is not Keplerian. (2) The tidally torn body was so small in comparison to r_p that to a good approximation all

the particle orbits have the same pericentre as the victim. [In nearly circular cases some particles would have that as their apocentre.]

Our aim is to describe the positions and velocities of particles at time t after the pericentric passage that released them at some given r_p. Equation (15) gives us the galactocentric radius at which we shall find those particles with any chosen r_a after time t. We take a logarithmic potential and use $K = 3.1, k = 1.44$ so

$$r = L(r_a)\{1 - \tilde{e}(r_a)\cos[Q(r_a)Kt]\}^{1/K} \tag{16}$$

In the above we have emphasised the dependencies on r_a as this is different for each particle of the stream. With r so determined we now find the azimuth ϕ of each particle. In Eq. (10) the angle measured from pericentre is given by

$$\phi = [kq(r_a)]^{-1}\eta = [kq(r_a)]^{-1}\cos^{-1}\{[1/e(r_a)][(l(r_a)/r)^k - 1]\} \tag{17}$$

For each time t Eqs. (16) and (17) are parametric equations for the shape of the tidal trail with parameter r_a, so for each r_a (16) gives r and (17) gives the corresponding ϕ [but read on for details of how the correct branch of $\cos^{-1}$ can be automatically selected.] Notice that the eccentricity of the victim's orbit does not figure in the shape of the trail although in practice particles with eccentricities close to that will no doubt be most numerous in the stream formed. The velocities of every particle of the trail are readily found. The tangential velocity is $\mathbf{h} \times \hat{\mathbf{r}}/r$ and the velocity component away from the galactic centre is found from (10) and (2):

$$\dot{r} = qe(h/l)(r/l)^{k-1}\sin[kq\phi]. \tag{18}$$

The velocity vector is $\mathbf{v} = \dot{r}\hat{\mathbf{r}} + \mathbf{h} \times \hat{\mathbf{r}}/r$ and h is given by (4). In practice, eq. (17) can only be used for small values of t where the principal value of the $\cos^{-1}$ returns the value of η on the correct branch. Once $\pi < \eta < 2\pi$ the principal value will be $2\pi - \eta$ instead, so one gets a spurious value of ϕ corresponding to the wrong branch and as a result those particles of the stream apparently start to move backwards. In practice it is essential to have an automatic method of selecting the correct branch of the $\cos^{-1}$. Luckily KQt, which we know, reaches π just as the particle considered reaches apocentre which is when $\eta = \pi$. Thus if we define $sgn = \sin(KQt)/|\sin(KQt)|$ which changes from 1 to -1 at π, then the expression $(1 - sgn)\pi + sgn.\mathrm{Cos}^{-1}([(l/r)^k - 1]/e)$ returns the value on the correct branch whenever $0 < \eta < 2\pi$. [The capital C as usual denotes the principal value, i.e. the one between 0 and π.] However, that expression never gets larger than 2π whereas η does. To obtain a fully corrected formula that selects the right branch we use the function int(x) which gives the integer part of its argument. When we replace the $\cos^{-1}(\ )$ in Eq. (17) by the expression $\mathrm{int}[(KQt)/(2\pi)].2\pi + (1 - sgn)\pi + sgn.\mathrm{Cos}^{-1}\{1/e[(l/r)^k - 1]\}$ we get the correct value. All this assumes that both t and η are zeroed at the pericentric passage at which the victim was torn apart. Figure 3 shows the development of the tidal streams.

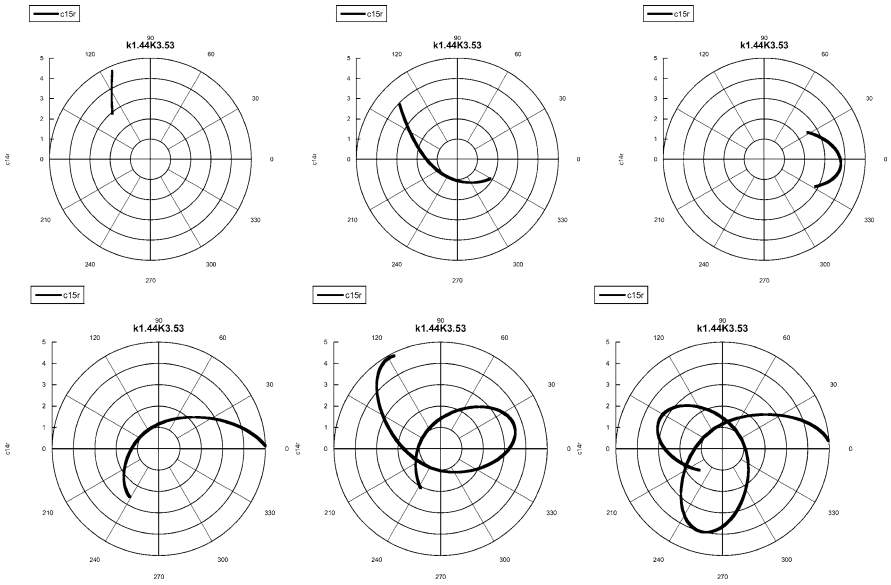


Fig. 3 Tidal trails were formed as a small victim on an orbit in the logarithmic potential with $r_a = 4$ passed its $r_p = 1$ pericentre at an azimuth of 3 o'clock in the figures. The six figures show the development of the tidal stream of particles with apocentres from 3 to 5. The times are 5, 10, 15, 20, 45, and 60 units past the tearing. The azimuthal period of the innermost stream particles is about 15 units

5 Mass Segregation and Cluster Evolution

Henon [12] found that the cores of his cluster models became denser with time. He postulated that this only stopped when the core became sufficiently dense that triple or multiple encounters led to the formation of hard binaries which provided a new source of energy as they became more strongly bound. Antonov [2] showed that even an isothermal sphere bounded by a rigid sphere was unstable to the core contracting to high density and temperature provided that the density contrast between centre and edge was greater than 708. Lynden-Bell and Wood [19] explained this instability in terms of the negative heat capacity of the core; for a good review see Padmanabhan [20]. Work by Hachisu et al. [9] and Cohn [4] led Lynden-Bell and Eggleton [17] and Heggie and Stevenson [11] to study the self-similar runaway of the core and its subsequent re-expansion. This was reviewed in the general context of negative heat capacities in Lynden-Bell [15]. However, all those works concentrated on the ideal case of clusters of equal mass stars. Meanwhile, Spitzer [21] emphasised that the tendency of encounters to equalise the kinetic energies of the interactors would lead to the more massive stars congregating near the centre and that this effect proceeded more rapidly than Antonov's gravothermal instability. Numerical experiments by Aarseth [1] showed that in clusters with a distribution of stellar masses, the core often became dominated by a single massive binary often formed by the two

most massive stars in the cluster. Originally I regarded this as a natural limiting case of Spitzer's mass segregation, but my current investigation in collaboration with Wyn Evans and Sverre Aarseth started when I realised that the precise shape of the mass function might determine whether or when the core became dominated by a single massive binary. Furthermore, the evolution of the whole system might then be changed, as explored by Heggie and Hut [10]. In a cluster with a broad distribution of masses as in the Salpeter function, thermal equilibrium will ensure that the stars of the highest masses will move slowly and be confined to the centre of the core where the potential is quadratic. At thermal equilibrium the Maxwell–Boltzmann distribution of the stars with masses between m and $m + dm$ will be

$$f(m, \mathbf{r}, \mathbf{v})dm/m = A(m)\left(\frac{m\beta}{2\pi}\right)^{3/2} exp[-\beta m(v^2/2 + \psi_0 - \psi)]dm \qquad (19)$$

Here $\psi(r)$ is the gravitational potential [which would be GM/r for a point mass] and ψ_o is its value at the centre. Setting $\psi_0 - \psi = (2/3)\pi G\rho_0 r^2$ for the most massive stars and integrating over all velocities and all radii we find the mass function of the cluster at the high mass end to be

$$N(m)dm = A(m)m^{-3/2}(G\rho_0\beta)^{-3/2}dm \qquad (20)$$

Salpeter's initial mass function for masses above the Sun is $N(m) \propto m^{-2.34}$. To attain this for the very high mass stars we need $A(m) \propto m^{-.84}$. The density at the centre is then $\int\int f d^3v dm = \int mA(m)dm \propto \int m^{0.16}dm$ an integral that diverges at the high-mass end unless we impose a cut-off. Thus at least initially, before stellar evolution has taken its toll of the upper main sequence, the density in the core of such a mass-segregated cluster will be dominated by high-mass stars. If in place of the Salpeter function we set $N(m) \propto m^{-\alpha}$ then the density will only converge at the high-mass end for $\alpha > 3.5$. Thus unless the mass function is considerably steeper than Salpeter's the central density of a mass-segregated cluster with no binaries will be dominated by a few stars of very high mass. A pair of the most massive ones will no doubt form a hard binary on a short timescale and the energy emitted via encounters with it will then prevent core collapse and drive an expansion of the cluster instead. To investigate this matter more fully we decided to set up a model cluster in Maxwell–Boltzmann equilibrium inside a rigid sphere [following Antonov's treatment of the equal mass case].

6 Mass-Segregated Isothermal Clusters

The density corresponding to the distribution function (19) is

$$\rho = \int\int f d^3v dm = \int A(m)exp[-\beta m(\psi_0 - \psi)]dm \qquad (21)$$

At the centre we shall take the distribution of masses to be of the form

$$n_0(m)dm = \rho_0[(\nu-1)!m_c]^{-1}\mu^{\nu-1}e^{-\mu}dm/m \tag{22}$$

where $\mu = m/m_c$ and $(\nu-1)! = \Gamma(\nu)$; ν and m_c are the parameters of the mass distribution, whose central density is ρ_0. $A(\mu) = \rho_0[(\nu-1)!m_c]^{-1}\mu^{\nu-1}e^{-\mu}$. The number density $n(0)$ of the stars at the centre is given by integrating $n_0(m)$ over m but the integral only converges at small m when $\nu > 1$, so for smaller ν we take a low-mass cut-off at m_0. We then find, for $(m_0/m_c) << 1$,

$$n_o = [\rho_0/m_0] \times \begin{cases} (\nu-1)^{-1}, & \nu > 1 \\ \ln(m_c/m_0) - \gamma, & \nu = 1 \\ \Gamma(\nu)(\nu-1)^{-1}(m_c/m_0)^{1-\nu}, & 0 < \nu < 1 \end{cases}$$

The mean mass at the centre is $\bar{m} = \rho_0/n_0$. For $\nu > 1$ we rewrite the distribution in terms of $\bar{m}$,

$$F = \left(\frac{\bar{\beta}m}{2\pi(\nu-1)}\right)^{3/2} \frac{\rho_0(\nu+3/2)!}{(\nu-1)!}\left[1 + \frac{\beta\bar{m}}{(\nu-1)}(v^2/2 + \psi_0 - \psi)\right]^{-(\nu+3/2)} \tag{23}$$

We then see that as ν tends to infinity this becomes an exponential and the mass distribution becomes a δ function about $\bar{m}$. Thus the single mass isothermal sphere is recovered as a limiting case.

Writing Poisson's equation for the density distribution

$$\nabla^2\psi = -4\pi G\rho = -4\pi G\rho_0(1+U)^{-\nu} \tag{24}$$

where $U = \beta m_c(\psi_0 - \psi)$ so in spherical symmetry

$$\bar{r}^{-2}d/d\bar{r}(\bar{r}^2 d\bar{U}/d\bar{r}) = \bar{U}^{-\nu} \tag{25}$$

where $\bar{U} = 1 + U$ and $\bar{r} = (4\pi G\rho_0\beta m_c)^{1/2}r$. Equation (25) is the Lane–Emden equation for a polytrope of negative index $n = -\nu$[3]. For small r,

$$\bar{U} = 1 + \bar{r}^2/6 - \frac{3\nu}{10}(\bar{r}^2/6)^2 + \frac{\nu(\nu+1)}{12}(\bar{r}^2/6)^3 + O(\bar{r}^8) \tag{26}$$

For large $\bar{r}$

$$\bar{U} = W_0\bar{r}^{2/(\nu+1)}; \qquad W_0 = \left[\frac{\nu+1}{\nu+3}\right]^{1/(\nu+1)} \tag{27}$$

The most massive stars confined to the quadratic core have $N(m) \propto \mu^{\nu-7/2}e^{-\mu}$, which for suitable ν can be made to approximate the Salpeter function. However many more stars climb out of the quadratic core of the cluster and inhabit the

$\bar{U} \propto \bar{r}^{2/(\nu+1)}$ halo. At stellar masses for which most stars are in the asymptotic halo but do not reach the bounding sphere we find $N(m) \propto \mu^{-(\nu+7)/2}$ which is much too steep for Salpeter. There is an amusing conundrum here. A larger ν gives rise to a greater preponderance of high-mass stars in the core, see Eq. (22), but a larger ν gives rise to a halo with a greater preponderance of low-mass stars! The resolution of the conundrum lies in the lower power of $\bar{r}$ in the potential when ν is made larger. This gives more room for the halo which is primarily populated by low-mass stars. Such considerations suggest that it may be impossible to set up Maxwell–Boltzmann equilibria when the numbers of stars of high mass are too great as compared with the number of low-mass stars. Direct from the Virial theorem it is easy to prove that in any isolated cluster the mass-weighted root mean square velocity of the stars is only a half of the mass-weighted root mean square velocity of escape. This at once suggests that at equipartition a small number of stars of less than a quarter of the mean mass will soon leave the cluster. When, following Antonov, we confine our cluster within a rigid sphere off which the stars bounce with impunity, the above considerations no longer apply. The stars of still lower mass are mainly confined by the sphere and for them the relevant $N(m) \propto \mu^{\nu-2}$.

7 Critical Radii for Gravothermal Instability

For isothermal clusters of equal mass stars confined within a sphere of radius r_e one may plot the dimensionless variable $B = -E/(GM^2/r_e)$, where E is the total energy which is negative in the interesting cases, against the density contrast ρ_0/ρ_e. For such clusters a maximum occurs at 0.334 at which the density contrast is 708. If one puts a cluster with energy E in a sphere whose radius is larger than $0.334GM^2/(-E)$ a thermal runaway develops with the centre getting denser and hotter. It is already known that this instability develops at smaller density contrasts in clusters made of stars of different masses. Numerical experiments planned with Sverre Aarseth will, we hope, confirm this preliminary indication and should determine under what circumstances the core is dominated by a single binary. Preliminary experiments with clusters that were not initially segregated showed that clusters with Salpeter-like mass functions develop a central hard binary earlier than clusters with greater numbers of low-mass stars which develop cores with many stars in them.

Appendix: Refinements of Orbital Parameters

When we determine q^2 by setting $S_q(\bar{u}) = S(\bar{u})$ we implicitly assumed that the latter has its maximum close to $u = \bar{u}$, so that we got a good estimate of the amplitude of S. However, there are cases when $S(u)$ is somewhat unsymmetrical. In such cases a better estimate of $q^2e^2\bar{u}^2$ is given by the maximum of S which may be found by using first-order Newton–Raphson starting from $\bar{u}$. This gives the estimate

$S_{max} = S(\bar{u}) - [S'(\bar{u})]^2/S''(\bar{u})$. The analogous formula for Q is

$$Q^2\tilde{e}^2\bar{U}^2 = \{\tilde{S}(\bar{U}) - [\tilde{S}(\bar{U})]^2/\tilde{S}''(\bar{U})\} \tag{28}$$

Whereas that improvement in the value of q is sufficient to give good orbital shapes, the timing is not so good at high eccentricity, so it is often necessary to correct the lopsidedness of $\tilde{S}(U)$ by taking a different value of K and thence a different U. These changes leave the form of the orbital equations unchanged but do change the values of the parameters within those forms. We detail the method for determining k such that S does have its maximum at the new $\bar{u}$ and we then quote the more important similar result for K and $\tilde{S}$. In place of basing our values of k and K on the gradients at turning points method, we now aim to base them on the demand that the maxima of S and $\tilde{S}$ must occur at the midpoint between the new u_p and u_a or the new U_p and U_a, respectively. We detail the analysis here to first order in Newton–Raphson; higher orders can be achieved by iteration. Whereas the method gives a more accurate answer when the other fails, its results have none of the beauty of the turning point method. In particular K is no longer $4 - k$, indeed for Kepler $5 - k$ is a better first guess for K and the best approximation ranges from 3.5 at $r_a/r_p = 3$ to 4.0 at $r_a/r_p = 10$, a good compromise being $K = 3.8$. Henceforth we denote the values of k and K given by Eqs. (9) and (13), respectively, as k_0 and K_0. We shall regard them as zero-order approximations to the true k and K that lead to $S(u)$ having its maximum at the midpoint of $u_p = r_p^{-k}$ and $u_a = u_a^{-k}$, and to $\tilde{S}(U)$ having its maximum at the midpoint of r_p^K and r_a^K. We see from Eq. (7) that, however, we may eventually choose k, $S(u)$ is always of the form $S(u) = S_1(r)u^2$. Here $S_1(r) = [2E + 2\psi(r) - h^2/r^2]h^{-2}r^2$ and $S_1' = dS_1/dr$, etc. so

$$dS/du = S_1'[dr/du]u^2 + 2S_1u = -[rS_1'/k - 2S_1]u \qquad A1 \tag{29}$$

Our aim is to choose k so that this is zero at the midpoint of r_p^{-k} and r_a^{-k}. We shall assume that this new k is not far from our old estimate k_0, however, that may have been estimated, so we write $k = k_0 + \delta k$ and work to first order in δk. Now $\delta r_p^{-k} = -\delta k \ln r_p r_p^{-k_0}$ and $\bar{u} = \bar{u}_0 + \delta\bar{u} = (r_p^{-k} + r_a^{-k})/2$, so $\delta\bar{u}/\delta k = -(\ln r_p r_p^{-k_0} + \ln r_a r_a^{-k_0})/2$. Likewise $l^{-k} = \bar{u}$ so $\ln l = -(1/k)\ln\bar{u}$, and so $\delta l/\delta k = -(l_0/k_0)(\ln l_0 + \bar{u}_0^{-1}\delta\bar{u}/\delta k)$. The condition that dS/du is zero at $\bar{u}_0 + \delta\bar{u}$ with $k = k_0 + \delta k$ is found from A1

$$\delta k = \frac{[rS_1'/k_0 - 2S_1]\bar{u}_0}{[\{[2S_1' - (S_1' + rS_1'')/k_0]\delta l/\delta k + rS_1'/k_0^2\}\bar{u}_0 + (2S_1 - rS_1'/k_0)\delta\bar{u}/\delta k]} \tag{30}$$

where all the quantities on the right are evaluated at $r = l_0$. With the new $k = k_0 + \delta k$ now known, e and q are evaluated as in the main body of this chapter with $l = l_0 + \delta l$.

Very similar calculations can be made to find a refined value of K. From (12) $\tilde{S} = S_2(r)U^2$, where $S_2(r) = 2E/r^2 + 2\psi(r)/r^2 - h^2/r^4$, and we then have

$$d\tilde{S}/dU = [rS_2'/K + 2S_2]U \tag{31}$$

Furthermore since $U_p = r_p^K$, etc. we find

$$\delta\bar{U}/\delta K = (\ln r_a r_a^{K_0} + \ln r_p r_p^{K_0})/2 \tag{32}$$

and

$$\delta L/\delta K = -L_0 K_0^{-2}[\ln \bar{U}_0 - K_0(\delta\bar{U}/\delta K/)\bar{U}_0] \tag{33}$$

The condition that $d\tilde{S}/dU$ is zero at the new midpoint then gives

$$\delta K = \frac{-[rS_2'/K_0 + 2S_2]\bar{U}_0}{\{[2S_2' + (S_2' + rS_2'')/K_0]\delta L/\delta K - S_2' r/K_0^2\}\bar{U}_0 + (2S_2 + rS_2'/K_0)\delta\bar{U}/\delta K} \tag{34}$$

where all quantities on the right are evaluated at $r = L_0$. With the new $K = K_0 + \delta K$, L, $\tilde{e}$, and Q are evaluated as in the main body of this chapter. The solution is only changed by the adoption of the new parameters, K, L, $\tilde{e}$, Q. As mentioned earlier typical values of k for logarithmic potentials are 4/3 to 1.5 with 1.44 a value that does well for all orbits. Similarly $K = 3.1$ produces reasonably symmetrical $S(U)$. For the Kepler problem $k = 1$ is of course exact and $K = 3.8$ does well for orbits with $r_a/r_p < 10$. In non-power law potentials such as Yukawa's both k and K vary so they must be chosen orbit by orbit.

References

1. Aarseth, S.V. 1968, Bull Astron, 3(5), 105
2. Antonov, V.A. 1962, Vest. Leningrad University, vol. 7, p. 135, Translation (1995): IAU Symposium 113, 13, p. 525
3. Binney, J.J., Tremaine, S.D. 2008, Galactic Dynamics 2nd ed., Princeton University Press, Princeton, NJ, vol. 23, p. 389
4. Cohn, H. 1980, ApJ, 242(17), 765
5. Freeman, K.C. 1966, MNRAS, 134(3), 1
6. Freeman, K.C. 1966, MNRAS, 134(4), 15
7. Freeman, K.C., Norris, J.E. 1981, ARA&A, 19(1), 319
8. Freeman, K.C. Bland-Hawthorne, J. 2002, ARA&A, 40(2), 487
9. Hachisu, I., Nakada, Y., Nomoto, K., Sugimoto, D. 1978, PThPh, 60(16), 393
10. Heggie, D.C., Hut, P. 2003, The Gravitational Million-Body Problem, Cambridge University Press, Cambridge, p. 22
11. Heggie, D.C., Stevenson, D. 1988, MNRAS, 230(19), 223
12. Henon, M. 1961, Ann Astrophys, 24(12), 369
13. Lin, D.N.C., Lynden-Bell, D. 1982, MNRAS, 198(88), 707
14. Lynden-Bell, D. 1963, Observatory, 83(10), 23
15. Lynden-Bell, D. 1999, Physica A, 263(20), 293
16. Lynden-Bell, D. 2010, MNRAS, 402(9), 1937
17. Lynden-Bell, D., Eggleton, P.P. 1980, MNRAS, 191(18), 483
18. Lynden-Bell, D., Jin, S. 2008, MNRAS, 386(11), 245

19. Lynden-Bell, D., Wood, R. 1968, MNRAS, 138(14), 495
20. Padmanabhan, T. 1990, Phys Rep, 188(15), 285
21. Spitzer, L. 1969, ApJ, 158(21), L139
22. Toomre, A., Toomre, J. 1972, ApJ, 178(7), 623
23. Van Albada, T.S. 1995, Bull Astron Neth, 19(6), 478

Stellar Debris Streams: New Probes of Galactic Structure and Formation

Carl J. Grillmair

Abstract Stellar debris streams provide us with a powerful new means of probing the three-dimensional structure and evolution of the Galaxy. Fourteen "distant" halo streams have been discovered to date, and by their length and continuity, they demonstrate that such streams are long-lived structures. Being relatively nearby, we can accurately measure the positions and motions of individual stars in these streams and use them to construct a history of past accretion events and of the underlying potential field in which they orbit. We briefly discuss these streams and the insights they have given us to date. With the capabilities of new facilities coming on line or in development, we can look forward to an era of continued discovery and a much more detailed understanding of galactic structure and formation.

1 Introduction

Stellar debris streams are among the most powerful and sensitive probes of the Galaxy's potential field on large scales [39, 41]. Streams resulting from weak tidal stripping are particularly useful as the stars in such streams have relatively small random velocities and travel on orbits nearly identical to those of their progenitors. By sampling positions and velocities at many points along such streams, one can measure the exchange of kinetic and potential energy, and thus the shape of the potential field. With a sample of many such streams, one can determine both their orbits and the shape of the Galactic potential in a self-consistent manner.

Debris streams also provide a new window on the formation process of the Galaxy. The streams discovered to date appear to be very long-lived structures and simply counting streams will greatly improve estimates of the number and distribution of dwarf galaxies and star clusters which, through disruption, contributed to the buildup of the Galactic halo [8]. With upcoming and planned surveys, we can look forward to a day when we will be able to lay out a precise, chronological sequence of major events that led to the Galaxy as we see it today.

C.J. Grillmair (✉)
Spitzer Science Center, California Institute of Technology, Pasadena, CA 91125, USA
e-mail: carl@ipac.caltech.edu

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_21,

Very cold debris stream may also be useful for detecting dark matter concentrations in the halo. The collision cross section for something as long as a tidal stream is not negligible, and if there are a multitude of dark matter clumps orbiting in the halo, they should occasionally leave their mark on a tidal stream. Flyby encounters can induce warps and even loops in a stream which may last upward of 10^7 years before ultimately dispersing and leaving gaps. Finding many streams will both enable us to put limits on the prevalence of such dark matter clumps and increase our odds of catching such encounters in flagrante.

Stars do not form in isolation (as far as is known), and through the evaporation and dispersal of stellar clusters, a collisionless system the size of our Galaxy may well contain billions of stellar debris streams. Most of these streams will to a greater or lesser extent be the product of external tidal forces, but not all such streams will be useful probes of the Galactic potential. Most streams will be confined to the disk and, reflecting the diminutive sizes of their progenitors (presumably open clusters and lesser stellar associations), are probably neither long nor populous. Even if we were able to disentangle such streams from the surrounding population, they would likely tell us more about the process of star formation than about the overall structure of the Galaxy. Examples include the Pleides-Hyades, Hercules, and other moving groups that have been known for some time.

A second observational class of streams includes those known as local halo streams. Many such streams have been detected using spectroscopic and proper motion data to isolate groups of stars in the solar neighborhood with relatively low metallicities and occupying a well-defined region of phase space. One of their virtues is that, with complete and/or well-defined samples, we can estimate to what extent the halo population is made up of major and minor streams. This in turn can be used to infer the types of mergers and accretion events the Galaxy has undergone in the past. For example, Helmi et al. [29] estimate that about 10% of metal poor stars in the halo come from a single coherent structure disrupted early in the formation of the Galaxy. Looking at kinematic substructure in the pencil-beam, high-latitude, Spaghetti survey, Starkenburg et al. [67] estimated that 20% of the stars in their sample were in substructures, though their results were also consistent with a halo built entirely of disrupted satellites. Klement et al. [39] used SDSS DR7 and SEGUE spectroscopy to detect five phase space overdensities, of which two were previously undiscovered. Schlaufman et al. [66] find evidence for ten discrete elements of cold halo substructures (ECHOS), seven of them previously undiscovered, and put an upper limit of one-third on the fraction of metal-poor, main sequence turnoff stars that are in ECHOS.

One drawback of local halo streams is that they have thus far been difficult to trace to moderately large distances. A third class of streams therefore includes those that we can discern from a distance, close enough that we can measure relative distances along the streams, but far enough away that they appear coherent and can be traced over an appreciable portion of their orbits around the Galaxy. Like local halo streams, these are streams for which we can resolve individual stars, and they have generally been detected using a combination of wide-field photometric filtering and/or radial velocity information. Since we can in principal measure all six phase

space coordinates for stars that are widely separated along an orbit, these streams provide a relatively new and remarkably sensitive way of constraining the potential field of the Galaxy. We refer to these streams as distant halo streams (DHSs) to distinguish them from local halo streams and these are the focus of this review.

A fourth class of debris streams includes those we can see orbiting or arcing out from other galaxies. These streams are important for reconstructing the merger and accretion histories of galaxies in general and are discussed in a separate contribution.

2 Distant Halo Streams

Thanks in large part to the availability of wide-field digital photometric surveys (primarily the Two-Micron All Sky Survey (2MASS) and the Sloan Digital Sky Survey (SDSS)), recent years have seen a rapid growth in the number of known DHSs. Table 1 lists those streams known at the time of writing, in approximate order of discovery. We list only structures that are highly elongated, longer than 5° on the sky and therefore potentially suitable as orbit tracers and probes of the Galactic potential [50]. Neither the 2MASS nor the SDSS databases have been exhausted, and several new streams are currently being analyzed [28]. The locations of a subset of the streams listed in Table 1 are shown in Fig. 1.

Table 1 Known distant halo streams

Designation	Progenitor	Selected references
Sagittarius	Sagittarius Dwarf Galaxy	[3, 4, 6, 13, 18, 31, 32, 42, 45, 47, 49, 68, 71]
Virgo Stellar Stream	NGC 2419?	[15, 52, 70]
Palomar 5	Palomar 5	[26, 55–57, 64]
Monoceros Ring	Unknown (dwarf galaxy?)	[33, 51, 58, 63, 74]
NGC 5466	NGC 5466	[5, 19, 24]
Orphan Stream	Unknown (dwarf galaxy?)	[7, 19, 20, 54, 65]
GD-1	Unknown (globular cluster?)	[26, 41, 73]
AntiCenter Stream	Unknown (dwarf galaxy?)	[21, 27]
EBS	Unknown (dwarf galaxy?)	[21, 27]
Acheron	Unknown (globular cluster?)	[22]
Cocytos	Unknown (globular cluster?)	Grillmair 2009 [22]
Lethe	Unknown (globular cluster?)	Grillmair 2009 [22]
Styx	Bootes III dwarf?	Grillmair 2009 [22]
Cetus Polar Stream	NGC 5824?	[53]

Note that we do not list the so-called Magellanic stream [48, 72] in Table 1, which appears to be purely a gaseous feature, and whose orientation conflicts with proper motion measurements of the Magellanic clouds [59]. The Magellanic stream is more likely a product gas outflows and ram pressure effects and therefore less useful as a probe of the Galaxy's potential.

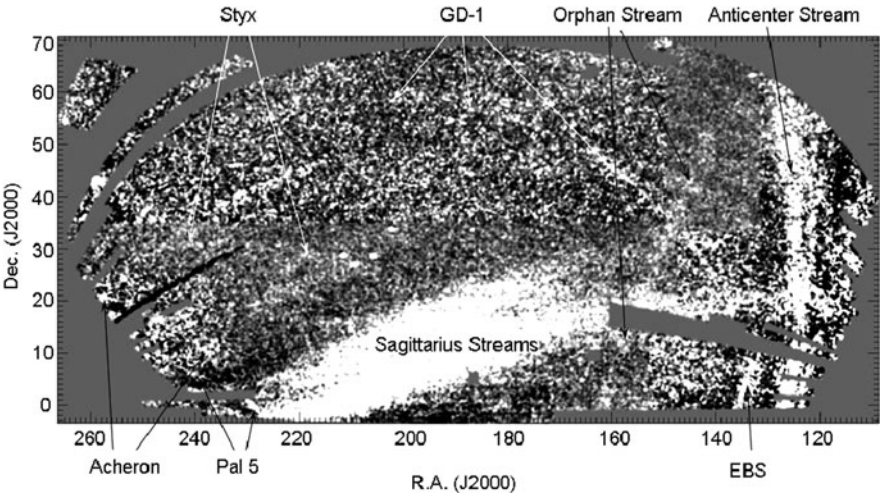


Fig. 1 A composite, matched-filter surface density map of the Sloan Digital Survey area for Data Release 5. A filter based on the color–magnitude distribution of stars in the globular cluster M 13 is used throughout, though optimized for distances ranging from 5 to 45 kpc in different parts of the Survey footprint. Some of the streams listed in Table 1 are indicated

2.1 The Sagittarius Stream

Not long after the discovery of the Sagittarius dwarf galaxy (Sgr) [31], a number of observers began to see signs of a major stellar debris stream extending out from Sgr along a polar orbit around the Galaxy [3, 6, 18, 32, 45, 47, 49, 68, 71]. The Sagittarius stream is now believed to extend completely around the Galaxy at least twice [18], with apo- and perigalactic distances of 60 and 15 kpc, respectively, and it remains the most spectacular debris stream yet discovered. Among the many remarkable aspects of this stream is the presence of a population gradient along the stream [4], in the sense that the metallicity of the population decreases substantially (-0.7 dex) as one moves out from the Sgr to the end of the leading arm [13].

Although the stream is several degrees across over most of the sky, it is surprisingly cold, with a velocity dispersion of $\approx 10\,\mathrm{km}^{-1}$ [46]. Combined with its enormous length and nearly polar orbit, this makes the stream the most sensitive tracer of the Galactic potential yet discovered. Two arms of the Sagittarius stream dominate the SDSS footprint (Fig. 1), and Fellhauer et al. [18] determined that the southern stream is a young leading arm, the northern stream is an old, trailing arm, and that an old leading arm lies behind the young leading arm. The degree of bifurcation apparent in Fig. 1 implies very little orbital precession and requires a halo that is very nearly spherical. To reconcile this lack of precession with the radial velocities of stars in the leading arm (which can only be reproduced with prolate halo models), Law et al. [42] invoke a triaxial halo with a minor/major axis ratio of ~ 0.67. While this model appears to satisfy the current observational constraints, it

must also contend with a potential and longstanding problem that orbits in the plane of the disk become unstable [9].

2.2 The Virgo Stellar Stream

The Virgo Stellar Stream (VSS) was first discovered by Vivas et al. [70] as a small but significant overdensity of RR Lyrae in the QUEST survey, calling the structure the "12.4 hour clump." Subsequent spectroscopy of these and a number of blue horizontal branch stars, as well as SDSS photometry, revealed an overdensity covering at least 100 square degrees, at a heliocentric distance of about 19 kpc and with a mean $V_{gsr} = 100\,\mathrm{km\,s^{-1}}$ and [Fe/H] $= -1.86$ [15]. Based on both position and velocity, Newberg et al. [52] were able to relate their S297+63−20.5 substructure with the VSS. Prior et al. [60] associated RR Lyrae in the southern hemisphere with the stream, more than doubling the length on the sky to about 45° (15 kpc in projection).

It has yet to be established exactly how the VSS and the Virgo Overdensity (VOD) – discovered in SDSS DR7 by Juric et al. [35] – might be related. Although apparently somewhat closer than the VSS, the VOD is both much broader on the sky and more extended along the line of sight.

The VSS itself is highly inclined with respect to our line of sight, and Casetti-Dinescu et al. [12] used the radial velocity, distance, and proper motion of a single RR Lyra in Kapteyn Selected Area 103 (7° west of the nominal center of the VSS) to compute a preliminary orbit. This orbit is inclined 58° to the Galactic plane, with a relatively high eccentricity of $e = 0.8$ and an apogalactic radius in the vicinity of 90 kpc. If borne out by additional detections along this orbit, this makes the VSS very interesting from the standpoint of constraining the total mass of the Galaxy. Although the uncertainties are sill rather sizable, Casetti-Dinescu et al. [12] suggest the remote globular cluster NGC 2419 as a possible progenitor of the stream.

2.3 The Palomar 5 Tidal Tails

The remarkable tidal tails of Palomar 5 were first discovered in SDSS commissioning data by Odenkirchen et al. [55]. At 2.6° long on the sky, these far surpassed in length the tidal tails found in globular clusters up until that time [23, 43] and convincingly demonstrated the promise of wide-field digital sky surveys. As more data became available and as analysis techniques improved, the tails gradually grew in length [56, 64] until reaching their currently known extent of 22° [25], limited at the southern end by the SDSS footprint. The tails are not at all smooth and show fairly regularly spaced clumps and gaps, which have variously been interpreted as a natural consequence of episodic tidal stripping, encounters with Galactic substructure, and even as a possible product of Jeans instability [61]. Koch et al. [40] examined the stellar populations in some of these clumps and found luminosity functions that were

significantly enhanced in the number of faint stars compared with the cluster itself. This is consistent with mass segregation in the cluster through two-body relaxation, followed by tidal stripping of the outermost stars [2, 38].

High-resolution spectroscopy of stars in Pal 5's tails by Odenkirchen et al. [57] showed that the velocity dispersion of tail stars is less than $5\,\mathrm{km\,s^{-1}}$, and perhaps as low as $2\,\mathrm{km\,s^{-1}}$, making them about as cold a tracer of the Galactic potential as one could possibly hope for. Extensive modeling of Pal 5's tails [14] shows that Pal 5, currently near the apogalactic point of its orbit, is unlikely to survive its next passage through the disk.

2.4 The Monoceros Ring

The first indications of a significant overdensity in the Monoceros-Canis Major region came from a photometric analysis of early SDSS data by Newberg et al. [51]. Yanny et al. [74] subsequently used SDSS-measured radial velocities of stars within 30° of the Galactic plane to isolate a large structure moving in a prograde orbit with a circular velocity of $110\,\mathrm{km^{-1}}$ and a velocity dispersion ranging from 22 to $30\,\mathrm{km^{-1}}$. Ibata et al. [33] subsequently traced the structure over 100° of sky, and Rocha-Pinto et al. [63] traced it halfway around the Galaxy using 2MASS M giants. The existence of these M giants implies that the Monoceros Ring contains stars much more metal rich than the value of [Fe/H] = 1.6 measured by Yanny et al. [74].

Penarrubia et al. [58] compiled all the disparate detections of the Monoceros Ring and used them to put constraints on the orbit of the stream. The result was a nearly circular orbit (e=0.1), with an apogalacticon of 23 kpc and inclined 25° to the plane of the Galaxy. Helmi et al. [30] examined numerical models of the disruption of satellite galaxies in a nearly coplanar orbit and showed that the Monoceros Ring could be either a relatively transient tidal arc or a shell like those found in nearby elliptical galaxies. A progenitor for the Monoceros Ring has yet to be identified.

2.5 The NGC 5466 Tidal Tails

The tidal tails of NGC 5466 were almost simultaneously discovered by Belokurov et al. [5] and Grillmair and Johnson [24]. Using SDSS DR4, Belokurov et al. [5] mapped out the strongest, inner 4° of the tails, while Grillmair and Johnson [24] traced a much more tenuous stream over 45° of the sky. Fellhauer et al. [19] used numerical simulations to match the orientation, morphology, and surface density of the tails. Unlike Palomar 5, Fellhauer et al. [19] concluded that, while NGC 5466 may have lost more than 60% of its mass to its tails over time, the cluster is, nevertheless, likely to survive for another Hubble time.

2.6 The Orphan Stream

The Orphan stream was initially discovered by Belokurov et al. [6] in their "Field of Streams," and subsequently analyzed by Grillmair [20] and Belokurov et al. [7]. At between 20 and 50 kpc distant, at least 60° in length, and much narrower than the Sagittarius streams, the Orphan stream immediately attracted interest from modelers. The most recent analysis [54] estimates an orbital apogalacticon at ≈ 90 kpc from the Galactic center, making the Orphan stream perhaps the best candidate yet for constraining the total mass of the Galaxy.

Grillmair [25] noted the existence of a dwarf galaxy (later to become known as Ursa Major II – [76]) close to the projected path of the stream, but deemed it unlikely to be the progenitor based on its small size and significant separation from his modeled orbit of the stream. Fellhauer et al. [19] used numerical simulations to infer that UMa II could indeed be the progenitor, though this has been challenged by more recent simulations by Sales et al. [65]. Subsequent analysis of SEGUE spectroscopy [54] showed that velocities in the stream are inconsistent with an association with either UMa II or Complex A gas cloud. It appears that, for the moment at least, the Orphan stream continues to remain an orphan.

2.7 The GD-1 Stream

GD-1 was first detected by Grillmair and Dionatos [26] in SDSS DR5. The stream subtends about 63° on the sky and is between 8 and 10 kpc distant. Of immediate interest was GD-1's very narrow width, with a FWHM of a quarter of a degree over most of its length. GD-1 appeared to be as cold as the Pal 5 tails, but much longer on the sky. The stream is not limited by the SDSS footprint, but rather appears to fade to invisibility at both the eastern and the western ends. No progenitor for GD-1 has yet been identified.

Willett et al. [73] subsequently used SDSS and SEGUE spectroscopy to characterize GD-1 and refine its orbit. Stars in GD-1 were found to have [Fe/H] $= 2.1 \pm 0.1$ and be in a fairly eccentric (e=0.33), retrograde orbit, with peri- and apogalacticon of 14.4 and 29 kpc, respectively. More recently, Koposov et al. [41] improved on existing proper motions to extract accurate, six-dimensional phase space coordinates at several positions along the stream. The power of stream as a probe of the Galactic potential is demonstrated by the fact that they were able to constrain the circular speed V_c at the Sun's radius to $224 \pm 13\,\mathrm{km}^{-1}$ and a flattening of the potential of $\mathrm{q} = 0.87 + 0.07 - 0.04$ (though the latter is dominated by the disk). This is the best constraint on V_c yet measured and essentially the only strong constraint out at 15 kpc. The photometric distances remain the weakest link in this analysis, and Gaia parallax and proper motion measurements of GD-1 stars will enable still tighter constraints [16].

2.8 The Anticenter Stream

The Anticenter Stream (ACS) was first identified as distinct from the Monoceros Ring by Grillmair [21]. Quite obvious in the "Field of Streams" [5] and in Fig. 1, the SDSS imaging data were not immediately analyzed, presumably due to the initial identification of the feature with the Monoceros Ring. However, Grillmair [21] found that, after removing a smooth rise in star counts a low Galactic latitudes, a distinct ridge remained, which was itself broken into at least three relatively cold "tributaries." Moreover, the ACS lies at least 10° laterally from any extension of the best-fitting Monoceros orbit determined by Penarrubia et al. [58], though this may still be consistent with a single progenitor [30].

Grillmair et al. [27] subsequently measured velocities for stars in the ACS and were able to fit a roughly circular orbit that is qualitatively similar to the orbit of Monoceros, differing primarily in sky position and having a somewhat smaller apogalacticon. Notably, the velocity dispersions measured at two locations in the ACS by Grillmair et al. [27] and Carlin et al. [10] are substantially smaller than those measured by Yanny et al. [74] for the Monoceros Ring. Just as the ACS has been shown to contain distinct components, the ACS as a whole may well constitute just another component (or another wrap) of the Monoceros Ring, all deriving from a sizable progenitor that may have come apart in stages or in kinematically distinct pieces. The issue will hopefully be resolved when wide-field photometry becomes available for the southern hemisphere and we can trace these streams over a much larger arc.

2.9 The EBS Stream

The Eastern Banded Structure or EBS stream was first detected by Grillmair [21]. The EBS is 25° long, about 3° wide, and was found to lie roughly parallel to the Anticenter Stream on the sky, but at a somewhat higher Galactic latitude. Given the multi-component nature of the Anticenter Stream, EBS was at first thought to be associated with the same progenitor. Although radial velocity follow-up has yet to be carried out, the work of Grillmair et al. [27] for the Anticenter Stream suggests that EBS may actually be the next wrap of the ACS.

2.10 Acheron, Cocytos, and Lethe

These tenuous and very narrow streams were found by Grillmair [22] using matched-filtering of SDSS DR5 data. Ranging from 3 to 14 kpc distant, these streams are all on the order of 100 pc wide (FWHM), and Grillmair [22] concluded they were likely to be the remains of globular clusters. No progenitors have yet been identified.

2.11 Styx

The Styx stream was first detected by Grillmair [22] in SDSS DR5. The visible portion of the stream extends over some 53° of sky and is estimated to lie between 38 and 50 kpc distant. With an asymmetric profile and a FWHM of more than 3°, Grillmair [22] concluded that Styx was most likely to be the remains of dwarf galaxy. Interestingly, there is an unusually strong concentration of stars in the stream (though not precisely centered within it) with a very similar distance and turn-off color, and Grillmair [22] speculated that this might actually be the progenitor of Styx. Dubbed Bootes III, the object is extremely sparse and extended for a dwarf galaxy, suggesting that it might be in the final throes of dissolution. However, preliminary analysis of radial velocity measurements [11] seems to indicate that Styx and Bootes III are on very different orbits.

2.12 The Cetus Polar Stream

Using both photometry and spectroscopy from the SEGUE database, Newberg et al. [54] recently detected a fairly broad, tenuous stellar stream in a polar orbit at a heliocentric distance of roughly 34 kpc. The stars have a mean [Fe/H] $\approx$ 2.0 and a velocity dispersion that ranges from 4.5 to 10 km s^{-1}. The latter argues that the stream must have originated in a globular cluster or very diminutive dwarf galaxy, even though the stream is as much as 15° across in places. A preliminary orbit estimate does not put significant constraints on the shape of the halo, though it does suggest the globular cluster NGC 5824 as a possible progenitor.

3 The Next Steps

The strongest, most populous streams in the SDSS have now all been found. The most recent photometric detections [22] have been at a much lower signal-to-noise ratio (between 7 and 13) with stream surface densities down to less than 10 stars per square degree. Lacking a reliable, objective algorithm for matching stars of similar age, distance, and metallicity, the weakest streams have been found and given credence primarily by virtue of their length. While both SDSS and 2MASS show an abundance of "streamlets" of order 5–10° long, such features can also be detected in artificially generated fields of pure noise.

The somewhat subjective techniques for finding streams have so far resulted in a rather strong bias toward streams that are nearly perpendicular to our line of sight. Indeed, the weaker streams found in the SDSS are all oriented within 15° of the plane of the sky. If in the first instance we assume that stream orientations should be isotropically distributed with respect to our line of sight, then we might reasonably expect there to be another six times as many streams out to 30 kpc in the SDSS footprint than are currently known.

Using matched-filtered photometric techniques, where much of the signal comes from blue turn-off and upper main sequence stars, the 2MASS and SDSS surveys are effectively limited to 5 and 50 kpc, respectively. As shown by Rocha-Pinto et al. [63] for the Monoceros Ring, this does not preclude finding streams with bright tracers such as M giants. However, for the less populous streams of globular clusters and small dwarf galaxies, such tracers may be few or non-existent.

Further progress in finding weak streams in the SDSS may result from simultaneously combining photometric filtering with velocity and proper motion measurements. Spectra are now available for $\sim 5 \times 10^5$ stars in the SDSS DR7 and SEGUE data sets. Newberg et al. [53] combined photometric measurements with velocities and metallicities to detect the Cetus Polar Stream. This stream would probably never have been detected by photometric means alone, underscoring the vastness of phase space and the importance of using all available information.

Other upcoming surveys will enable the detection, characterization, and follow-up of many more streams. By virtue of both fainter limiting magnitudes and larger or different sky footprints, both the Pan-STARRS [36] and the SkyMapper [37] projects will greatly increase our inventory of stellar debris streams within 100 kpc. Working in the southern hemisphere, SkyMapper will also permit the tracing of known northern streams over much larger arcs. This in itself will put much tighter constraints on the orbits of the progenitors and on the shape of the Galactic potential.

The Large Synoptic Survey Telescope (LSST) will enable photometry of upper main sequence stars out to 400 kpc [1]. Combined with a larger footprint on the sky, the volume accessible to LSST will be some 2,000 times larger than that sampled by SDSS. While we do not expect to find nearly as many long and populous streams at very large distances (due to both a presumed smaller population of progenitors and the reduced strength of Galactic tidal forces), LSST will enable a much more penetrating survey of the extended halo of the Galaxy. Finding even a few debris streams at such distances will greatly improve our knowledge of the mass, shape, and extent of the dark matter halo.

Of course, detecting debris streams is only a first step. If we want to use them as probes of Galactic structure and formation, substantial resources will need to be dedicated to spectroscopic follow-up. Useful astrophysical constraints will require sampling stellar motions at many points along a stream. Given the very low surface densities of stream stars (10–500 stars per square degree), sample contamination of even well-defined streams by field stars can be upward of 90%, making this an unattractive proposition for typical time allocation committees. With 4,000 fibers over a 5° field of view, the Large Aperture Multi-Object Spectroscopic Telescope (LAMOST, [75]) will be well suited to the task, and a large collaboration is already in place to obtain and analyze spectra of over two million halo stars. The two-degree field facility (2dF, [44]) on the Anglo-Australian Telescope will presumably also find gainful long-term employment following up both SkyMapper and LSST detections.

Distance estimates are currently the largest source of uncertainty in the use of tidal streams as probes of the potential. RR Lyrae have yet to be found in any of the cold, globular cluster streams. Distances have consequently been estimated

photometrically using filter matching or upper main sequence fitting. Unfortunately, in addition to a paucity of stream stars and often severe contamination by field stars, in most cases the ages and metallicities of the stream stars are only poorly known and uncertainties due to inappropriate filter templates are large. Eyre and Binney [17] and Eyre [16] have described and demonstrated an alternative method of finding distances by using proper motions to measure "Galactic parallax." The accuracy of the method is currently on par with photometric distances, limited primarily by the quality of the currently available proper motions.

Finding RR Lyrae in streams would put distance estimates on much firmer footing. This indeed will be a major goal of the Palomar Transient Factory's Galactic Dynamics key project [62], which is expected to detect on the order of 30,000 RR Lyrae over 2π steradians. How many of these can be associated with known streams has yet to be determined, but even one or two per stream will help to pin down portions of each stream, which can then be used as a reference for relative distance estimates along the remainder. Ultimately, we look forward to Gaia, which is expected to be able to measure trigonometric distances to individual giant branch stars out to at least 10 kpc.

By themselves, positions, distance estimates, and line-of-sight velocities are still not enough to tightly constrain the local Galactic potential. As demonstrated by Koposov et al. [41], truly interesting things happen when one has good measurements for all six phase space coordinates at hand. Both the Gaia and the Space Interferometry Mission [69] are expected to make a substantial contributions to the field by improving proper motion measurements by two orders of magnitude over those currently available.

Given the rough and tumble of galaxy formation and the messy and unfinished business of late merging and accretion, the actual distribution of mass in our Galaxy may well be quite lumpy and complex. Detecting, tracing, and characterizing stellar debris streams at all distances and in all quadrants of the Galaxy will be the most important tool in our efforts to map this complexity. With the facilities and surveys coming online or on the horizon, we can look forward to a greatly improved census of the many constituents of our Galaxy and a much more detailed understanding of how they came to be here.

References

1. Abell et al. 2009, LSST Science Book, arXiv:0912.0201
2. Aguilar, L., Hut, P., Ostriker, J.P. 1988, ApJ, 335, 720
3. Alard, C. 1996, ApJ, 458, L17
4. Bellazzini, M., Newberg, H.J., Correnti, M., Ferraro, F.R., Monaco, L. 2006, AA, 457, 21
5. Belokurov, V., et al. 2006, ApJ, 637, L29
6. Belokurov, V., et al. 2006, ApJ, 642, L137
7. Belokurov, V., et al. 2007, ApJ, 658, 337
8. Bullock, J.S., Johnston, K.V. 2005, ApJ, 635, 931
9. Binney, J. 1981, MNRAS, 196, 455

10. Carlin, J.L., Casetti-Dinescu, D.I., Grillmair, C.J., Majewski, S.R., Nidever, D.L., Girard, T.M. 2010, ApJ, submitted
11. Carlin, J.L, Grillmair, C.J., Munoz, R.R., Nidever, D.L., Majewski, S.R. 2009, ApJ, 702, L9
12. Casetti-Dinescu, D.I., Girard, T.M., Majewski, S.R., Vivas, A.K., Wilhelm, R., Carlin, J.L., Beers, T.C., van Altena, W.F. 2009, ApJ, 701, L29
13. Chou, M-Y, et al. 2007, ApJ, 670, 346
14. Dehnen, W., Odenkirchen, M., Grebel, E.K., Rix, H.-W. 2004, AJ, 127, 2753
15. Duffau, S., Zinn, R., Vivas, A.K., Carraro, G, Mendez, R.A., Winnick, R., Gallart, C. 2006, ApJ, 636, L97
16. Eyre, A. 2010, MNRAS, 403, 1999
17. Eyre, A., Binney, J. 2009, MNRAS, 399, 160
18. Fellhauer, M., et al. 2006, ApJ, 651, 167
19. Fellhauer, M., et al. 2007, MNRAS, 375, 1171
20. Grillmair, C.J., 2006, ApJ, 645, L37
21. Grillmair, C.J., 2006, ApJ, 651, L29
22. Grillmair, C.J. 2009, ApJ, 693, 1118
23. Grillmair, C.J., Freeman, K.C., Irwin, M., Quinn, P.J. 1995, AJ, 109, 2553
24. Grillmair, C.J., Johnson, R. 2006, ApJ, 639, L17
25. Grillmair, C.J., Dionatos, O. 2006a, ApJ, 641, L37
26. Grillmair, C.J., Dionatos, O. 2006b, ApJ, 643, L17
27. Grillmair, C.J. Carlin, J.L., Majewski, S.R. 2008, ApJ, 689, L117
28. Grillmair, C.J., Mattingly, S., 2010, in preparation.
29. Helmi, A., White, S.D.M, de Zeeuw, T., Zhao, H. 1999, Nature, 402, 53–55
30. Helmi, A., Navarro, J.F., Meza, A., Steinmetz, M., Eke, V.R. 2003, ApJ, 592, L25
31. Ibata, R.A., Gilmore, G., Irwin, M.J. 1994, Nature, 370, 194
32. Ibata, R., Irwin, M., Lewis, G.F., Stolte, A. 2001, ApJ, 547, 133
33. Ibata, R.A., Irwin, M.J., Lewis, G.F., Ferguson, A.M.N., Tanvir, N. 2003, MNRAS, 340, L21
34. Johnston, K.V., Sigurdsson, S., Hernquest, L. 1999, MNRAS, 302, 771
35. Juric, M., et al. 2008, ApJ, 673, 864
36. Kaiser, N., et al. 2002, SPIE, 4836, 154
37. Keller, S.C., et al. 2007, PASA, 24, 1
38. King, I.R., Sosin, C., Cool, A.M. 1995, ApJ, 452, L33
39. Klement, R., et al. 2009, ApJ, 698, 865
40. Koch, A., Grebel, E.K., Odenkirchen, M., Martinez-Delgado, D., Caldwell, J.A.R. 2004, AJ, 128, 2274
41. Koposov, S., Rix, H.-W., Hogg, D.W. 2010, ApJ, 712, 260
42. Law, D.R., Majewski, S.R., Johnston, K.V. 2009, ApJ, 703, L67
43. Leon, S., Meylan, G., Combes, F. 2000, A&A, 359, 907
44. Lewis, I.J., et al. 2002, MNRAS, 333, 279
45. Majewski, S.R., Skrutskie, M.F., Weinberg, M.D., Ostheimer, J.C. 2003, ApJ 599, 1082
46. Majewski, S.M., et al. 2004, AJ, 128, 245
47. Martinez-Delgado, D., Gomez-Flechoso, M.A., Aparicio, A., Carrera, R. 2004, ApJ, 601, 242
48. Mathewson, D.S., Cleary, M.N., Murray, J.D. 1974, ApJ, 190, 29
49. Mateo, M., Mirabal, N., Udalski, A., Szymanski, M., Kaluzny, J., Kubiak, M., Krezeminski, W., Stanek, K.Z. 1996, ApJ, 458, L13
50. Montuori, M., Capuzzo-Dolcetta, R., Di Matteo, P., Lepinette, A., Miocchi, P. 2007, ApJ, 659, 1212
51. Newberg, H.J., et al. 2002, ApJ, 569, 245
52. Newberg, H.J., Yanny, B., Cole, N., Beers, T.C., Florentin, P.R., Schneider, D.P., Wilhelm, R. 2007, ApJ, 668, 221
53. Newberg, H.J., Yanny, B., Willett, B.A. 2009, ApJ, 700, L61
54. Newberg, H.J., Willett, B.A., Yanny, B., Xu, Y. 2010, ApJ, 711, 32
55. Odenkirchen, M., et al. 2001, ApJ, 548, L165

56. Odenkirchen, M., Grebel, E.K., Dehnen,W., Rix, H.-W., Yanny, B., Newberg, H.J., Rockosi, C.M., Martinez-Delgado, D., Brinkmann, J., Pier, J.R. 2003, AJ, 126, 2385
57. Odenkirchen, M., Grebel, E.K., Kayser, A., Rix, H.-W., Dehnen, W. 2009, AJ, 137, 3378
58. Penarrubia, J., Martinez-Delgado, D., Rix, H.W., Gomez-Flechoso, M.A., Munn, J., Newberg, H., Bell, E.F., Yanny, B., Zucker, D., Grebel, E.K. 2005, ApJ, 626, 128
59. Piatek, S., Pryor, C., Olszewski, E.W. 2008, AJ, 135, 1024
60. Prior, S.L., Da Costa, G.S., Keller, S.C., Murphy, S.J. 2009, ApJ, 691, 306
61. Quillen, A.C., Comparetta, J. 2010, MNRAS, submitted
62. Rau, A., et al. 2009, PASP, 121, 1334
63. Rocha-Pinto, H.J., Majewski, S.R., Skrutskie, M.F., Crane, J.D. 2003, 594, L115
64. Rockosi, C.M., et al. 2002, AJ, 124, 349
65. Sales, L.V., Helmi, A., Starkenburg, E., Morrison, H.L., Engle, E., Harding, P., Mateo, M., Olszewski, E.W., Sivarani, T. 2008, MNRAS, 389, 1391
66. Schlaufman, K.C., et al. 2009, ApJ, 703, 2177
67. Starkenburg et al. 2009, ApJ, 698, 567–579
68. Toten, E.J., Irwin, M.J. 1998, MNRAS, 294, 1
69. Unwin, S.C., et al. 2008, PASP, 120, 38
70. Vivas, A.K., et al. 2001, ApJ, 554, L33
71. Vivas, A.K., Zinn, R., Gallart, C. 2005, AJ, 129, 189
72. Wannier, P., Wrixon, G.T. 1972, ApJ, 173, 119
73. Willett, B.A., Newberg, H.J., Zhang, H., Yanny, B., Beers, T.C. 2009, ApJ, 697, 207
74. Yanny, B., et al. 2003, ApJ, 588, 824
75. Zhao, Y. 1999, Observational Astrophysics in Asia and Its Future, ed. P. Chen, Chinese Academy of Sciences, Beijing, p. 1
76. Zucker, D.B., et al. 2006, ApJ, 650, L41

Chemical Enrichment in Galaxies: Constraints on Nucleogenesis and Galaxy Evolution

Francesca Matteucci

Abstract We will describe the chemical enrichment history in galaxies of different morphological type (spirals, ellipticals, irregulars) as deduced from the abundances and abundance ratios of key chemical elements measured in these galaxies. By means of the comparison between the results of models of galactic chemical evolution, including detailed stellar nucleosynthesis, relaxing the instantaneous recycling approximation and successfully reproducing the local properties of different galaxies, and abundance data, we are able to impose constraints on the nucleogenesis of elements such as α-elements (O, Mg, Ne, Si, S) and Fe. Moreover, the chemical abundances allow us to reconstruct the history of formation and evolution of galaxies by reconstructing their past star formation histories. Our results suggest that: (i) elliptical galaxies formed very quickly and the more massive ones formed stars faster than the less massive ones, (ii) spiral disks formed by slow accretion of cold gas on timescales much longer than those of ellipticals, (iii) dwarf irregular galaxies formed stars mainly through bursts and suffered metal enhanced winds.

1 How to Model the Chemical Enrichment in Galaxies

Galactic chemical evolution studies show how the abundances of the most abundant chemical elements in the interstellar medium evolve in space and time. In order to do that we need to build models containing important prescriptions relative to the most important physical processes acting in galaxies. In particular the main parameters are

- The initial conditions, namely whether the gas which forms a galaxy has a primordial chemical composition or has been pre-enriched by an early stellar

F. Matteucci (✉)
Astronomy Division, Department of Physics, Trieste University, Via G.B. Tiepolo 11, 34134 Trieste, Italy; INAF, Trieste, Via G.B. Tiepolo 11, 34134 Trieste, Italy
e-mail: matteucc@oats.inaf.it

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_22,

generation. Then one should decide whether the system is closed (no exchange of gas with the surrounding) or open (infall, outflow).
- The birthrate function is defined as

$$B(m,t) = \psi(t)\varphi(m) \tag{1}$$

where $\psi(t) = SFR$ is the star formation rate and $\varphi(m) = IMF$ is the initial mass function.
- Stellar evolution and nucleosynthesis which include stellar yields and yields per stellar generation.
- A definition of the rate of gas infall and/or outflow from the system.

The SFR can be parametrized in several ways, such as exponentially decreasing in time with a typical timescale which should be fixed by reproducing the present time properties of the studied object. However, the most common parametrization of the SFR is the Schmidt [42] law where the SFR is proportional to some power of the surface gas density:

$$SFR = \nu\sigma_{gas}^{k} \tag{2}$$

where $k = 1.4 \pm 0.15$ [20] and ν is the efficiency of star formation, namely the inverse of the star formation timescale and it is fixed by reproducing the present time SFR.

More sophisticated parametrizations include also a dependence upon the total surface mass density, such as the one suggested by Dopita and Ryder [12]:

$$SFR = \nu\sigma_{tot}^{k_1}\sigma_{gas}^{k_2} \tag{3}$$

with $1.5 < (k_1 + k_2) < 2.5$.

The IMF is a probability function and is normally described by a power law with one or more slopes:

$$\varphi(M) = aM^{-(1+x)} \tag{4}$$

and it represents the number of stars with masses in the interval M, M+dM.

The IMF has been derived only for the stars in the solar vicinity and we recall the one of Salpeter [40] with only one slope $x = 1.35$ for stars $> 10\,M_{\odot}$ and the more recent ones, extended to low-mass stars with more than one slope [23, 41]. We do not have any information about the IMF in other galaxies than the Milky Way.

The stellar yields represent the mass ejected in the form of newly formed elements by each star and depend on the stellar mass and its chemical composition. Stars contribute differently to the galactic chemical enrichment according to their mass:

- Low and intermediate mass stars ($0.8 \leq M/\mathrm{M}_{\odot} \leq 8.0$) produce ^{4}He, C, N (primary and secondary) and s-process ($A > 90$) elements. An element is called *secondary* if it needs some metal seed, already present in the stars at its birth, to be produced. A typical example is N which is produced by means of the C–N–O cycle at expenses of C and O. If C and O are produced in situ by the star then N should be considered to be a primary element (deriving directly from H and He).
- Massive stars ($M \geq 10\,\mathrm{M}_{\odot}$) die as SNe II and SNe Ib/c. They produce mainly α-elements (O, Ne, Mg, Si, S, Ca), some Fe-peak elements, s-process elements ($A < 90$), and r-process elements during the SN explosion.
- White dwarfs of C and O (therefore originating in the low and intermediate mass range) in binary systems can give rise to Type Ia SNe which produces mainly Fe and Fe-peak elements.
- Very massive objects (VMO) ($M > 100\,\mathrm{M}_{\odot}$, popIII), if they ever existed, could form only up to a metallicity of $10^{-4} Z_{\odot}$ (i.e. [43]) and therefore their effect on the element production is negligible [3].

The computation of the SN (II,Ia,b,c) rates is fundamental for studying the galactic chemical evolution. While the Type II,Ib,c rates are relatively simple to compute since one assumes that these SNe originate from single massive stars, the calculation of the Type Ia SN rate is rather more complex if one wants to take into account the binary nature of their progenitors (e.g. [14, 15, 33]). We recall here the main scenarios for the progenitors of Type Ia SNe: (i) the single degenerate scenario [50], namely C-deflagration in a C–O WD reaching the Chandrasekhar mass (M_{Ch}) after accreting material from a red giant companion. (ii) The double degenerate scenario, namely the merging of two C–O WDs, due to gravitational wave radiation, which explodes by C-deflagration when M_{Ch} is reached [18]. An empirical derivation of Type Ia SN progenitors [29, 30] suggests that $\sim 50\%$ of all SNe Ia should originate in binary systems (with a C–O white dwarf) with lifetimes $< 10^8$ year, whereas the rest originates from binary systems exploding on a much broader distribution of stellar lifetimes.

The galactic outflows in galaxies can be parametrized assuming that the wind rate is proportional to the SFR:

$$WR = \lambda SFR \tag{5}$$

where λ is a free parameter indicating the efficiency of the galactic wind.

Infall of primordial or slightly enriched gas is also an important process in galaxy evolution. One of the most common parametrizations of the infall rate is

$$IR = Ae^{-t/\tau} \tag{6}$$

where A is a constant derived by reproducing the present time total surface mass density and τ is a free parameter representing the timescale for the gas accretion.

2 The Chemical Evolution of Spirals: The Milky Way

Here we will describe a particular model for the chemical evolution of the Milky Way taken as representative of spiral galaxies. In particular, we show the results of the two-infall model originally suggested by Chiappini et al. [9]. In this model the Milky Way is assumed to form by means of two main gas infall episodes: the first occurring on a relatively short timescale (≤ 2 Gyr) and forming the inner halo, the Bulge and the thick disk, the second occurring on much longer timescales and giving rise to the thin disk. The timescale found for the formation of the solar vicinity is $\tau = 7$ Gyr, whereas the timescale for the formation of the inner disk is shorter and the timescale for the formation of the outer disk is longer. This mechanism produces an *inside-out* formation of the galactic disk. This kind of model is generally the most widely adopted to describe the formation of the Milky Way from a chemical evolution point of view (see also [1] and [8]). The first to suggest the inside-out model were Chiosi and Matteucci [10] followed by Matteucci and François [34]. In Fig. 1 we show the star formation rate, as predicted by the two-infall models, by adopting the SFR parametrization of Eq. (3) with $k_2 = 1.4$ and $k_1 = 0.4$ and a threshold density for the SFR of $7\,M_{\odot}\,pc^{-2}$. The IMF is that of Scalo [41].

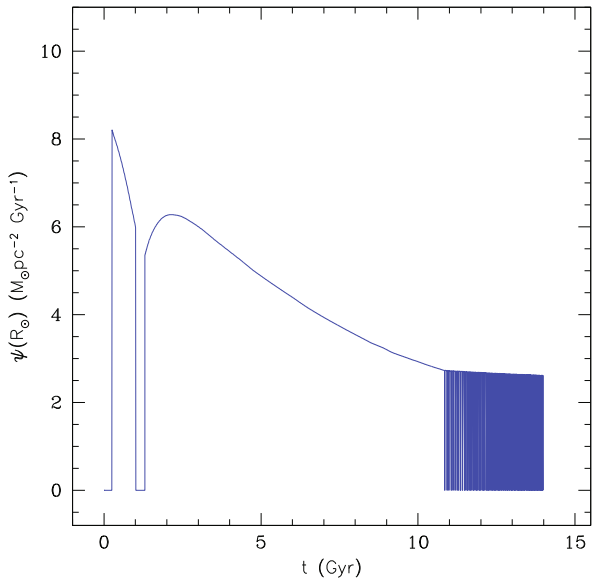


Fig. 1 The SFR in the halo and in the disk at the solar position as a function of time, as predicted by the two-infall model. The oscillations at late times are due to the fact that the surface gas density is very close to the threshold value. The efficiency of star formation is $\nu = 2\,Gyr^{-1}$ in the halo and $\nu = 1\,Gyr^{-1}$ in the disk

2.1 The Time-Delay Model

In Fig. 2 we show some recent calculations for the [X/Fe] vs. [Fe/H] relations [39]. Two model results are differing only for the assumed stellar yields: (1) yields from low and intermediate mass stars by van den Hoeck and Groenewegen [48] and yields from massive stars by Woosley and Weaver [51]; (2) yields from low and intermediate mass stars by Karakas [19] and yields from massive stars by the Geneva group for He and CNO and from Kobayashi et al. [21] for the other elements. Comparing the results of the two models with the observational data allows us to choose the best set of stellar yields, which results to be that of model 2. So, this is an example of how to constrain the nucleogenesis. The behaviours of the relative abundance ratios [X/Fe] vs. [Fe/H], shown in Fig. 2, are generally interpreted as due to the time-delay in the production of Fe by Type Ia SNe relative to the production of α-elements by Type II SNe, under the assumption of a constant IMF. This was first suggested by Tinsley [46] and then shown quantitatively by Matteucci and Greggio [33]. From the plots of Fig. 2 one can infer also the timescale for the halo formation, by assuming that the halo maximum metallicity is [Fe/H]=−1.0 dex, and this timescale is the time at which this metallicity is attained by the interstellar medium (ISM). There

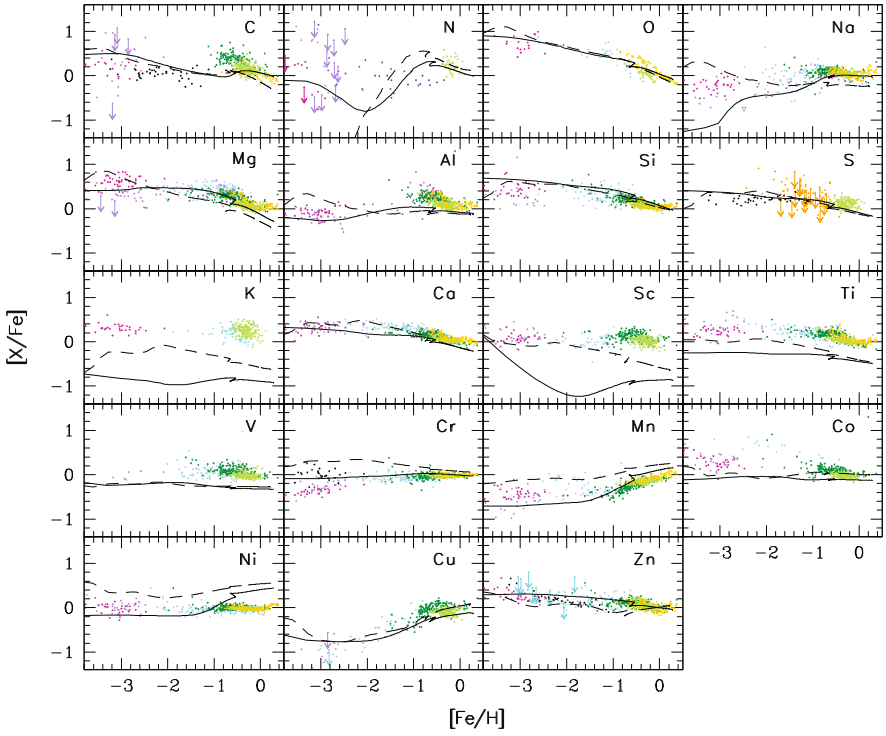


Fig. 2 Predicted and observed [X/Fe] ratios. The results of model 1 (see text) are represented by a *dashed line* whereas those of model 2 by a continuous line. Figure from Romano et al. [39]

are some elements like oxygen for which both models well reproduce the observed [O/Fe] vs. [Fe/H]. From this diagram the timescale for halo formation can be derived and it corresponds to $\tau_h \sim$ 1.5-2.0 Gyr [9, 13, 33, 34].

In Fig. 3 we show the predictions of the solar abundance ratios of each element relative to Fe, namely the abundance ratios at the time of the birth of the solar system (4.5 Gyr ago) compared with the observed solar abundances [2]. Also in this case, the yields of model 2 seem to be the best ones.

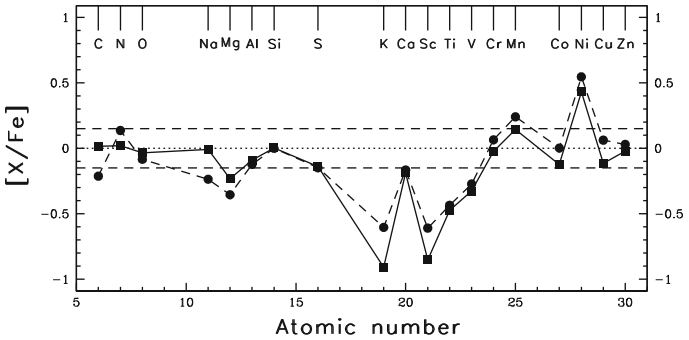


Fig. 3 Predicted and observed solar abundances. The observed abundance ratios are indicated by the *dotted line*. The *dashed lines* represent 0.15 dex of error. The *dashed line* refers to model 1 whereas the continuous line to model 2. Figure from Romano et al. [39]

3 Time-Delay Model and Star Formation Histories

The time-delay model, when applied to systems with different star formation histories, produces different behaviours than the one observed in the solar neighbourhood. In Fig. 4 we show the predicted [α/Fe] ratios as functions of [Fe/H] for spheroids such as the galactic bulge or elliptical galaxies and for irregular galaxies compared with the one in the Milky Way. The main differences are due to the different age–[Fe/H] relations in the three different cases.

In Fig. 5 we show the predictions of a model specifically developed for the Bulge by Ballero et al. [4]. In particular, the predicted [Mg/Fe] is compared with the most recent and accurate data and the agreement is quite good. These data confirm that the majority of the Bulge stars should show an overabundance of α-elements over a large range of [Fe/H]. The main assumption of the model is a strong SFR with efficiency ten times higher than that in the solar vicinity ($\nu_{Bulge} = 20\, Gyr^{-1}$ and a timescale for the Bulge formation is no longer than 0.1 Gyr). We predict a very similar evolution also for elliptical galaxies of the same mass as the Bulge (see [38]).

Another example of how chemical evolution models can constrain the nucleogenesis is given by the diagram [O/Mg] vs. [O/H] (Fig. 6) which shows an unexpected steep decrease for [O/H]$>$ 0 in the Bulge stars. This behaviour can be easily explained only if O yields from massive stars with mass loss are assumed. In fact, in

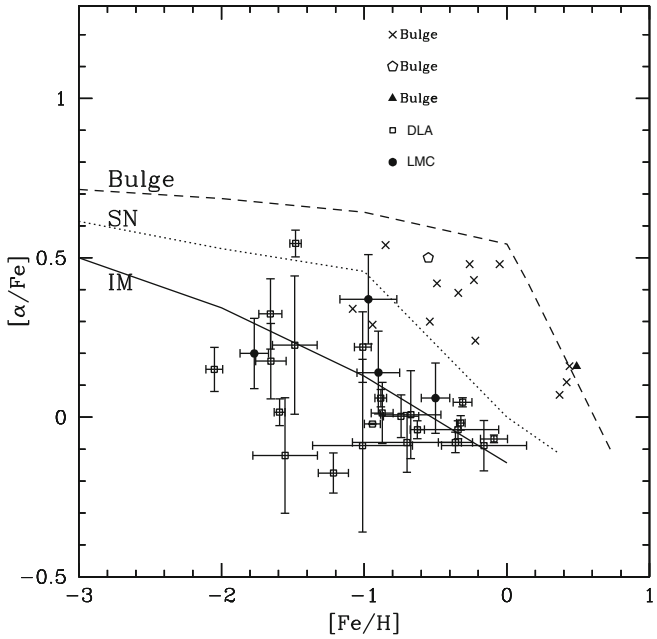


Fig. 4 The effect of the time-delay model coupled with different star formation histories. The *top line* represents the predictions for a system where the star formation proceeded very fast such as in the Bulge or in elliptical galaxies. The strong SFR induces a fast growth of [Fe/H] so that when Type Ia SNe start occurring, the metallicity of the gas is already above the solar one. The *middle line* represents the situation for the solar neighbourhood with a moderate star formation and the *bottom line* indicates a model for a typical irregular galaxy such as the Magellanic Clouds: here the SFR is assumed to be quite slow so that when Type Ia SNe start contributing Fe, the ISM metallicity is still quite low. Data for the LMC and some Damped-Lyman-α (DLA) systems are shown for comparison, indicating that these latter could well be irregular galaxies at high redshift [5, 6, 17, 35, 49]

this case the O production from massive stars is strongly depressed from stars with metallicity larger than solar. This is because mass loss subtracts large quantities of C and He from further processing. Thus, the net effect is an increase of C and He production and a depression of O production. This process is active for metallicities larger than solar because of the dependence of stellar mass loss on metallicity.

4 Abundance Ratios in Dwarf Spheroidals

Dwarf spheroidal galaxies (dSphs) are small systems surrounding the Milky Way which are often indicated as its building blocks in the framework of a hierarchical growth of structures. It is interesting to compare the abundance ratios, recently measured in an accurate way also in these systems, with those of the Milky Way. In Fig. 7 we show such a comparison, indicating that apart from a small overlap of the dSph data with the galactic data at relatively low metallicities, the dSph data have a

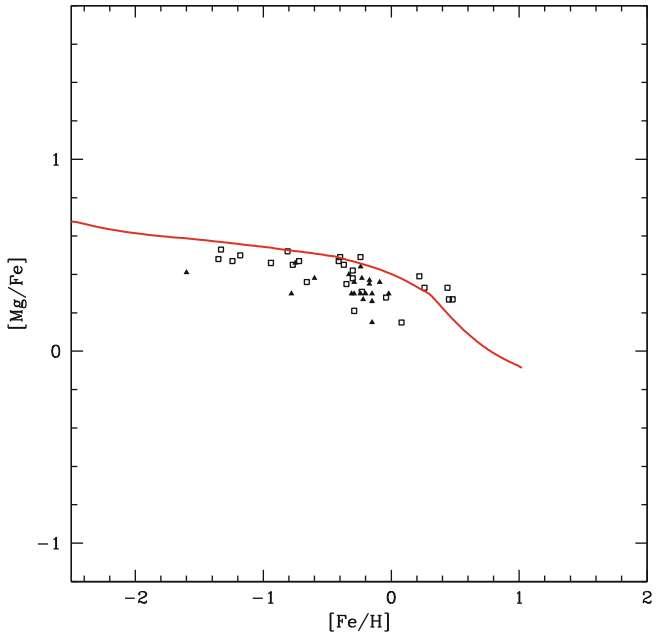


Fig. 5 Predicted and observed [Mg/Fe] vs. [Fe/H] for the Bulge. The model results are from Ballero et al. [4] where references to the data can be found: the main assumptions of the model are a strong SFR with an efficiency 10 times higher than that in the solar vicinity ($\nu_{Bulge} = 20 Gyr^{-1}$) and a timescale for the Bulge formation no longer than 0.1 Gyr

quite different behaviour with low [α/Fe] ratios at low metallicities. This behaviour resembles the one predicted for systems with slow star formation, as shown in Fig. 4. Certainly the different behaviour of the [α/Fe] ratios in dSphs has something to do with their star formation histories. To check this point Lanfranchi and Matteucci [24, 25] have computed models for dSphs by adopting the SF histories suggested by the colour–magnitude diagrams of these systems. Their models include also powerful galactic winds, assumed to be proportional to the SFR, which devoid the dSphs of all the residual gas after star formation. As one can see in Fig. 7 (right panel), the models can reproduce the observations, in particular the fast decline of the [α/Fe] ratios in dSphs. This is due to the combination of a slow SFR and a strong wind which subtracts gas from the system thus further lowering the SFR.

5 Elliptical Galaxies

Monolithic models assume that ellipticals suffer a strong star formation and quickly produce galactic winds when the energy from SNe injected into the ISM equates the potential energy of the gas. Star formation is assumed to halt after the development of a galactic wind and the galaxy is assumed to evolve passively afterwards. The original model of Larson [27] suggested that galactic winds should occur later in more massive objects due to the assumption of a constant efficiency of star forma-

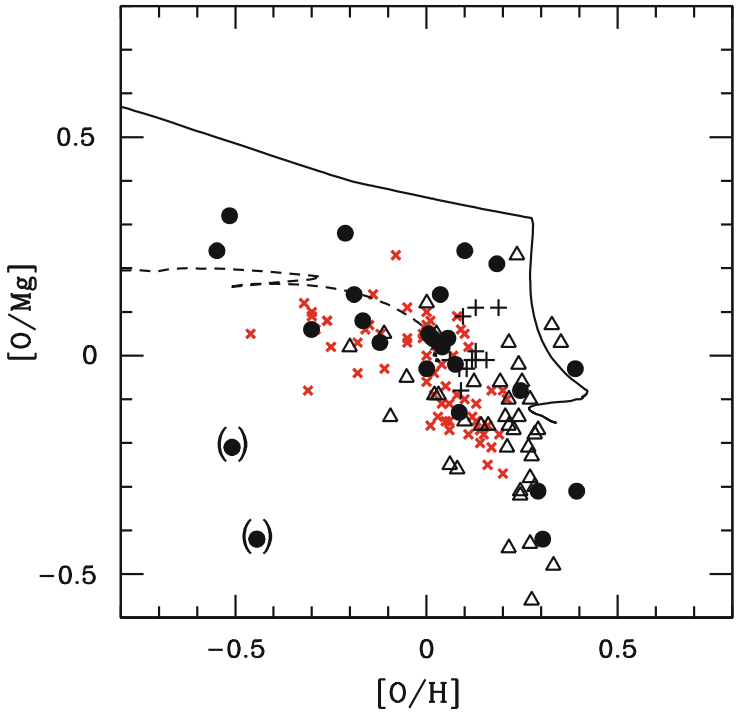


Fig. 6 Figure from McWilliam et al. [36]. The models are for the Bulge (*upper line*) and the disk (*lower line*). In both models the yields from massive stars with mass loss by Maeder [28] have been adopted. As one can see the mass loss becoming important only for $Z > Z_{\odot}$ can explain the steep decline of the [O/Mg] ratio, both in the Bulge and in the disk

tion in ellipticals of different mass and to the increasing depth of the potential well in more massive ellipticals. Unfortunately, this prediction is at variance with the observations indicating that the average $<$ [Mg/Fe] $>_*$ ratio in the dominant stellar population increases with galactic mass. This trend instead suggests a shorter period of star formation for larger galaxies. This was first suggested by Trager et al. [47], Worthey et al. [52], and Matteucci [31] who also computed models for ellipticals by assuming a shorter period of star formation in big ellipticals. In order to obtain that, an increasing efficiency of star formation with galactic mass was assumed with the consequence of obtaining a galactic wind occurring earlier in the massive than in the small galaxies. She called this process "inverse wind" and showed that such a model is able to reproduce the increase of $<$ [Mg/Fe] $>_*$ with galactic mass.

More recently, Pipino and Matteucci [38] presented a revised monolithic model which allows for the formation of ellipticals by a fast merger of gas lumps at high redshift. The model is multizone and predicts that each elliptical forms "outside-in" (star formation stops in the outer before the inner regions owing to a galactic wind). In other words, the galactic wind develops outside-in. Following the original suggestion by Matteucci [31], they assumed an increasing efficiency of star formation with the galactic mass. They also suggested a shorter timescale τ for the gas assembly

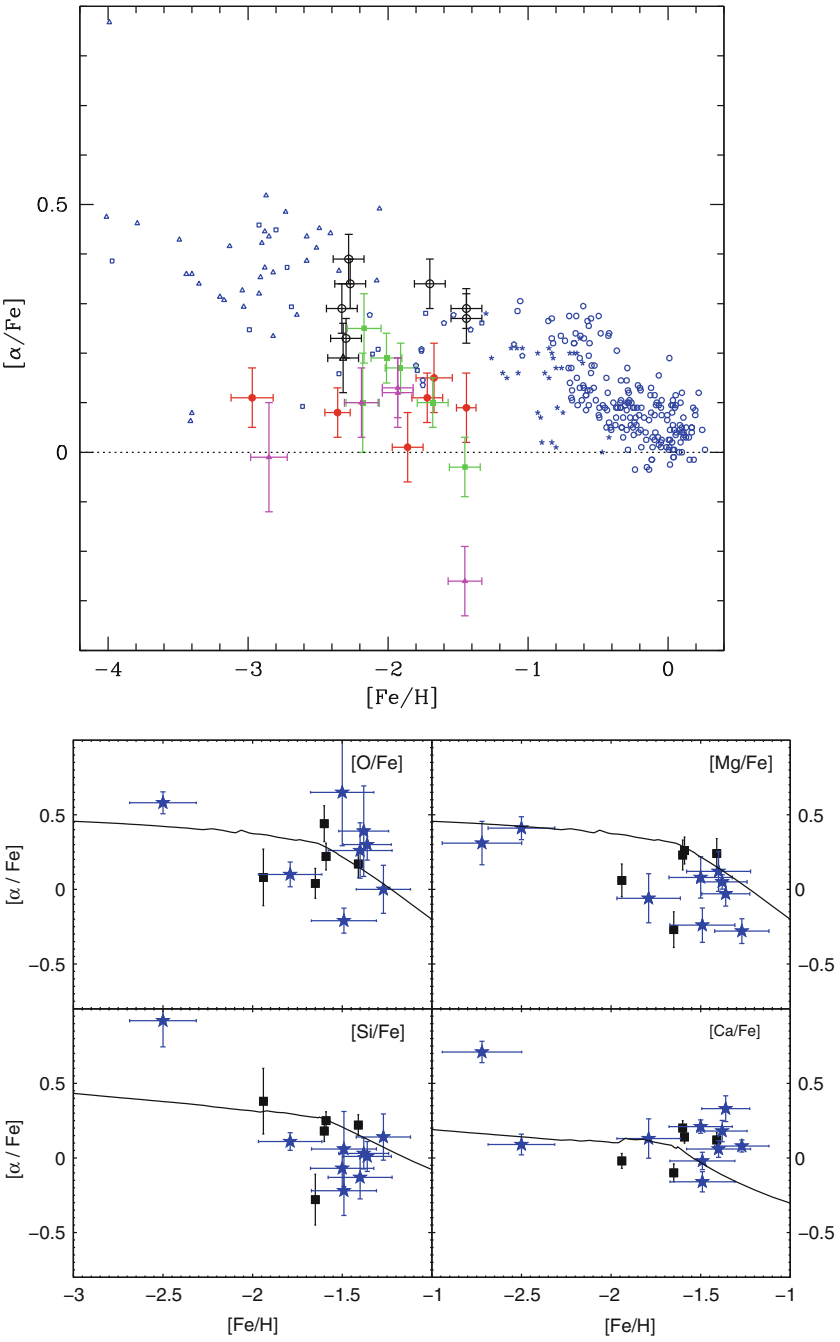


Fig. 7 *Left panel*: figure from Shetrone et al. [44]. The *blue dots* represent galactic stars whereas the others are the dSphs. *Right panel*: predicted and observed [α/Fe] ratios in the galaxy Carina. The model (*continuous line*) is from Lanfranchi et al. [26] and the data are from Koch et al. [22]. The figure is from Koch et al. [22]

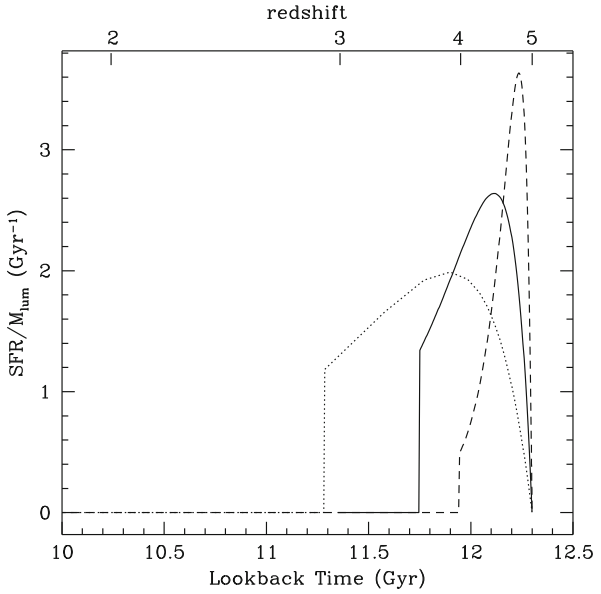


Fig. 8 The predicted star formation histories (star formation rate per unit stellar mass) for galaxies of 10^{12} (*dashed line*), 10^{11} (*continuous line*) and 10^{10} $M_{\odot}$ (*dotted line*). Such a behaviour is obtained by assuming that the efficiency of star formation is increasing with galactic mass whereas the timescale for the assembly of the gaseous lumps giving rise to the galaxies is a decreasing function of mass (downsizing both in star formation and mass assembly). In these models [38] the galactic wind occurs first in the more massive galaxies than in less massive ones

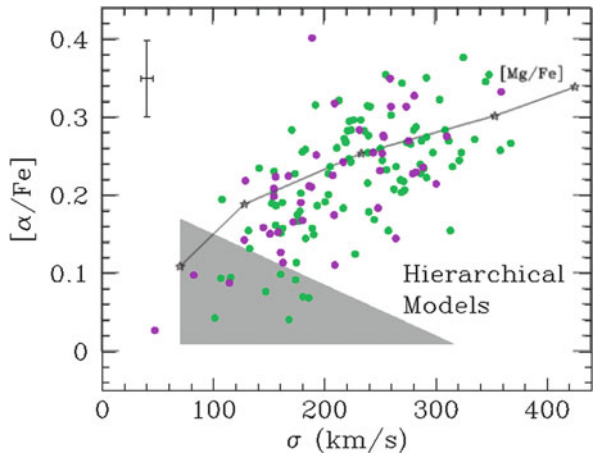


Fig. 9 The [Mg/Fe] vs. σ (velocity dispersion) for ellipticals (*continuous line*) as predicted by Pipino and Matteucci [38] compared with observations (*dots*) and with the predictions of classical hierarchical semi-analytical models (*shaded area*). Figure adapted from Thomas et al. [45]

with increasing galactic mass. In Fig. 8 we show the predicted histories of star formation in the "inverse wind scenario". As one can see, the most massive ellipticals show a shorter and more intense episode of star formation than the less massive ones.

Pipino and Matteucci [38], recomputed the relation $< [Mg/Fe] >_*$ versus mass (velocity dispersion) and compared it with the data by Thomas et al. [45], who showed how classical hierarchical semi-analytical models (but see [11]) cannot reproduce the observed $< [Mg/Fe] >_*$ vs. velocity dispersion trend, since in this scenario massive ellipticals have longer periods of star formation than smaller ones. In Fig. 9, we have plotted the predictions of the revised monolithic model (continuous line) compared with data and hierarchical clustering predictions.

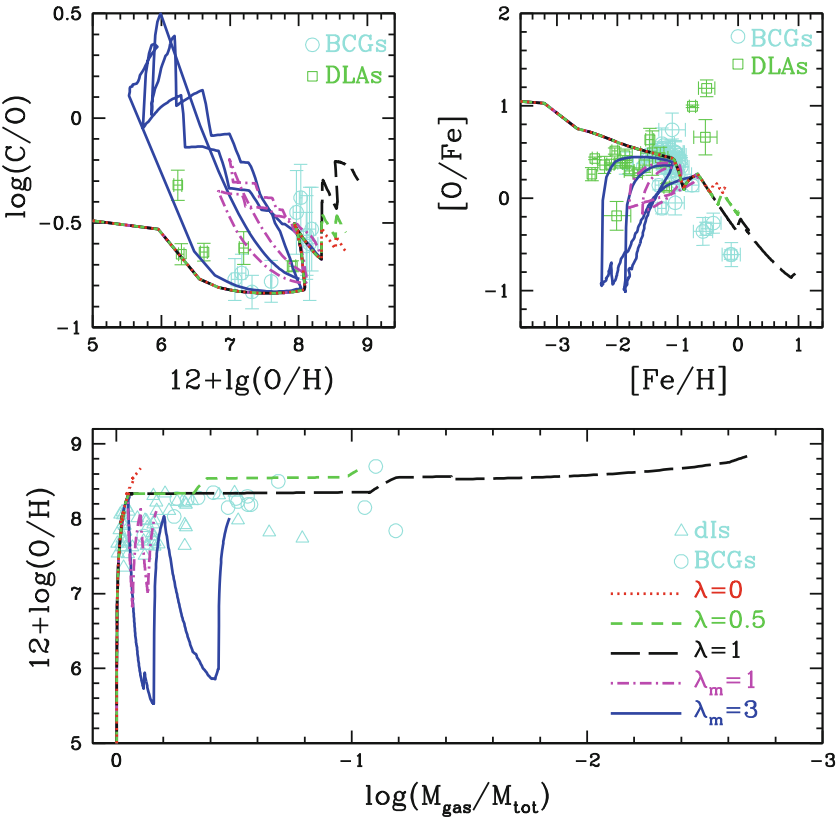


Fig. 10 Comparison between model results and data for DIGs and BCGs. The models all assume four SF bursts. Some models are computed with normal wind whose efficiency is represented by λ whereas other models with metal-enhanced winds whose efficiency is λ_m. Data and models from Yin and Matteucci [53]

6 Dwarf Irregulars

Among local star-forming galaxies, sometimes referred to as HII galaxies, most are dwarfs. Dwarf irregular galaxies can be divided into two categories: dwarf irregular (DIG) and blue compact galaxies (BCG). These latter have very blue colours due to active star formation at the present time. Various models have been proposed in the last years (e.g. [7, 16, 32, 37]) for these galaxies and many authors agree in concluding that in order to explain the low metallicity and the high gas content in these objects one has to assume either a bursting star formation or a low and continuous star formation regime coupled with metal-enhanced galactic winds which subtract mostly metals, leaving a large fraction of gas inside these galaxies. As an example we show the results of models with bursts and metal-enhanced winds in Fig. 10. The presence of bursts of SF followed by quiescent periods together with metal-enhanced winds creates the oscillating behaviour seen in Fig. 10. This is due to the fact that some elements stop to be produced after the burst (e.g. α-elements) whereas other elements continue to be produced (e.g. Fe and N). Different metal-enhanced wind efficiencies in different galaxies can nicely explain the spread observed in the data of dwarf irregular galaxies. On the other hand, normal winds, where H and He are lost at the same rate as metals, subtract too much gas to these galaxies and do not explain the whole spread in the abundances even if the wind efficiency varies from galaxy to galaxy, as it is shown in Fig. 10. Therefore, we conclude that metal-enhanced winds are required to explain dwarf irregulars.

Acknowledgments First of all I would like to thank the organizers of this exciting conference, and in particular Ken Freeman, for inviting me. I acknowledge support from the Italian Ministry of Research (MIUR) to the Project PRIN2007, Prot.2007JJC53X-001.

References

1. Alibés, A., Labay, J., Canal, R. 2001, A&A, 370, 1103
2. Asplund, M. 2005, ARA&A, 43, 481
3. Ballero, S., Matteucci, F., Chiappini, C. 2006, New Astron, 11, 306
4. Ballero, S., Matteucci, F., Origlia, L., Rich, M.R. 2007, A&A, 467, 123
5. Barbuy, B., Grenon, M. 1990. In: Bulges of Galaxies, eds. B.J. Jarvis, D.M. Terndrup, ESO/CTO Workshop, La Serena, Chile, p. 83
6. Barbuy, B., Ortolani, S., Bica, E. 1998, A&AS, 132, 333
7. Bradamante, F., Matteucci, F., D'Ercole, A. 1998, A&A, 337, 338
8. Chang, R.X., Hou, J.L., Shu, C.G., Fu, C.Q. 1999, A&A, 350, 38
9. Chiappini, C., Matteucci, F., Gratton, R. 1997, ApJ, 477, 765
10. Chiosi, C., Matteucci, F. 1980, In: Star Clusters, IAU Symp. 85, Reidel, Dordrecht, p. 109
11. De Lucia, G., Springel, V., White, S.D.M., Croton, D., Kauffmann, G. 2006, MNRAS, 366, 499
12. Dopita, M.A., Ryder, S.D. 1994, ApJ, 430, 163
13. François, P., Matteucci, F. Cayrel, R., Spite, M., Spite, F., Chiappini, C. 2004, A&A, 421, 613
14. Greggio, L. 2005, A&A, 441, 1055

15. Greggio, L., Renzini, A. 1983, A&A, 118, 217
16. Henry, R.B.C., Edmunds, M.G., Koeppen, J. 2000, ApJ, 660, 674
17. Hill, V., François, P., Spite, M., Primas, F., Spite, F. 2000, A&A, 364, L19
18. Iben, I. Jr., Tutukov, A. 1984, ApJ, 284, 719
19. Karakas, A. 2010, astro-ph/09122142
20. Kennicutt, R.C. 1998, ApJ, 498, 541
21. Kobayashi, C., Umeda, H., Nomoto, K., Tominaga, N., Ohkubo, T., 2006, ApJ, 653, 1145
22. Koch, A., et al. 2008, AJ, 135, 1580
23. Kroupa, P., Tout, C.A., Gilmore, G. 1993, MNRAS, 262, 545
24. Lanfranchi, G., Matteucci, F. 2003, MNRAS, 345, 71
25. Lanfranchi, G., Matteucci, F. 2004, MNRAS, 351, 1338
26. Lanfranchi, G., Matteucci, F., Cescutti, G. 2006, A&A, 453, 67
27. Larson, R.B. 1975, MNRAS, 173, 67
28. Maeder, A. 1992, A&A, 264, 105
29. Mannucci, F., Della Valle, M., Panagia, N., Cappellaro, E., Cresci, G.,Maiolino, R., Petrosian, A., Turatto, M. 2005, A&A, 433, 807
30. Mannucci, F., Della Valle, M., Panagia, N. 2006, MNRAS, 370, 773
31. Matteucci, F. 1994, A&A, 288, 57
32. Matteucci, F., Chiosi, C. 1983, A&A, 123, 121
33. Matteucci, F., Greggio, L. 1986, A&A, 154, 279
34. Matteucci, F., François, P. 1989, MNRAS, 239, 885
35. McWilliam, A., Rich, R.M. 1994, ApJS, 91, 7
36. McWilliam, A., Matteucci, F., Ballero, S., Rich, R.M., Fulbright, J.P., Cescutti, G. 2008, AJ, 136, 367
37. Pilyugin, L.S. 1993, A&A, 277, 42
38. Pipino, A., Matteucci, F. 2004, MNRAS, 347, 968
39. Romano, D., Karakas, A., Tosi, M., Matteucci, F. 2010, A&A in press
40. Salpeter, E.E. 1955, ApJ, 121, 161
41. Scalo, J.M. 1986, Fund Cosmic Phys, 11, 1
42. Schmidt, M. 1959, ApJ, 129, 243
43. Schneider, R., Omukai, K., Inoue, A.K., Ferrara, A. 2006, MNRAS, 369, 1437
44. Shetrone, M.D., Coté, P., Sargent, W.L.W. 2001, ApJ, 548, 592
45. Thomas, D., Maraston, C., Bender, R. 2002, In: R.E. Schielicke, ed. Reviews in Modern Astronomy, Wiley-VCH, Berlin, vol. 15, p. 219
46. Tinsley, B.M. 1979, ApJ, 229, 1046
47. Trager, S.C., Faber, S.M., Gonzalez, J.J., Worthey, G. 1993, AAS, 183, 4205
48. van den Hoeck, L.B., Groenewegen, M.A.T. 1997, A&AS, 123, 305
49. Vladilo, G. 2002, A&A, 391, 407
50. Whelan, J., Iben, I. Jr. 1973, ApJ, 186, 1007
51. Woosley, S.E., Weaver, T.A. 1995, ApJS, 101, 181
52. Worthey, G, Trager, S.C., Faber, S.M. 1995, ASP Conf. Ser., 86, ASP, San Francisco, 203
53. Yin, J., Matteucci, F., Vladilo, G. 2010, A&A, submitted

Chemodynamical Simulations of Galaxies

Chiaki Kobayashi

Abstract We predict the frequency distribution of elemental abundance ratios from carbon to zinc as a functions of time and location, which can be directly compared with the next generation of the galactic archeology project such as the HERMES. We perform the chemodynamical simulations of a Milky Way-type galaxy from a CDM initial condition, using a self-consistent hydrodynamical code with supernova feedback and chemical enrichment. In the simulated galaxy, the kinematical and chemical properties of the bulge, disk, and halo are consistent with the observations. The bulges have formed from the assembly of subgalaxies at $z \gtrsim 2$ and have higher [α/Fe] ratios because of the lack of contribution of Type Ia Supernovae. The disks have formed with a constant star formation over 13 Gyr and show a decreasing trend of [α/Fe] and increasing trends of [(Na,Al,Cu,Mn)/Fe]. However, the thick disk stars tend to have higher [α/Fe] and lower [Mn/Fe] than thin disk stars. About 60% of the thick disk stars have formed in the satellite galaxies before they accrete on the disk in this CDM-based simulation.

1 Introduction

While the evolution of the dark matter is reasonably well understood, the evolution of the baryonic component is much less certain because of the complexity of the relevant physical processes, such as star formation and feedback. With the commonly employed, schematic star formation criteria alone, the predicted star formation rates (SFRs) are higher than what is compatible with the observed luminosity density. Thus feedback mechanisms are in general invoked to reheat gas and suppress star formation. We include the feedback from stellar winds, core-collapse supernovae (normal Type II Supernovae SNe II and hypernovae HNe), and Type Ia Supernovae (SNe Ia) in our hydrodynamical simulations. Supernovae inject not only thermal

C. Kobayashi (✉)
Research School of Astronomy and Astrophysics, The Australian National University; Mt. Stromlo Observatory, Weston ACT 2611, Australia
e-mail: chiaki@mso.anu.edu.au

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_23,

energy but also heavy elements into the interstellar medium (ISM), which can enhance star formation. Chemical enrichment must be solved as well as energy feedback. Supernova feedback is also important for solving the angular momentum problem and the missing satellite problem, and for explaining the existence of heavy elements in intracluster medium and intergalactic medium, and possibly the mass-metallicity relation of galaxies ([20], hereafter K07).

In the next decade, high-resolution multi-object spectroscopy (HERMES) and space astrometry mission (GAIA) will provide kinematics and chemical abundances of a million stars in the local group. Since different heavy elements are produced from different supernovae with different timescales, elemental abundance ratios can provide independent information on "age." Therefore, stars in a galaxy are fossils to untangle the history of the galaxy formation. The galactic archeology technique can be used to study the galaxy formation and evolution in general. Metallicities are measured in various objects with different galaxy mass scale and as a function of redshift/time. The internal structure of galaxies has been observed with integral field spectrographs (e.g., the SAURON project, SINFONI on VLT). In order to untangle the formation and evolution history of the galaxy from observational data, a "realistic" model that includes star formation and chemical enrichment is required.

2 Chemical Enrichment Sources

Hypernovae (HNe) – The explosion mechanism of core-collapse supernovae is still uncertain, although a few groups have succeeded in exploding core-collapse supernovae. However, the ejected explosion energy and ^{56}Ni mass (which decays to ^{56}Fe) can be directly estimated from the observations, i.e., the light curve and spectra fitting of individual supernova. As a result, it is found that many core-collapse supernovae ($M \geq 20\,\mathrm{M}_\odot$) have more than 10 times larger explosion energy ($E_{51} \gtrsim 10$) and produce a significant amount of iron. We calculate the nucleosynthesis yields for wide ranges of metallicity ($Z = 0 - Z_\odot$) and the explosion energy (normal SNe II and HNe). Assuming that a large fraction of supernovae with $M \geq 20\,\mathrm{M}_\odot$ is HNe, the evolution of the elemental abundance ratios from oxygen to zinc is in excellent agreement with observations in the solar neighborhood, bulge, halo, and thick disk ([19], hereafter K06).

Figure 1 shows the evolution of heavy element abundance ratios [X/Fe] against [Fe/H] with our new yields (short-dashed lines) and with only SNe II (long-dashed lines, Nomoto et al. [24] yields adopted). The star formation history and the total number of SNe Ia are determined to meet the observed metallicity distribution function. In the early stage of galaxy formation, only SNe II explodes, and [α/Fe] stays constant. Because of the delayed Fe production by SNe Ia, [α/Fe] decreases toward 0. α-elements, O, Mg, Si, S, and Ca, show the plateau at [Fe/H] $\lesssim -1$. Ti is underabundant overall, which will be solved with the 2D calculation of nucleosynthesis. The observed decrease in the odd-Z elements (Na, Al, and Cu) toward

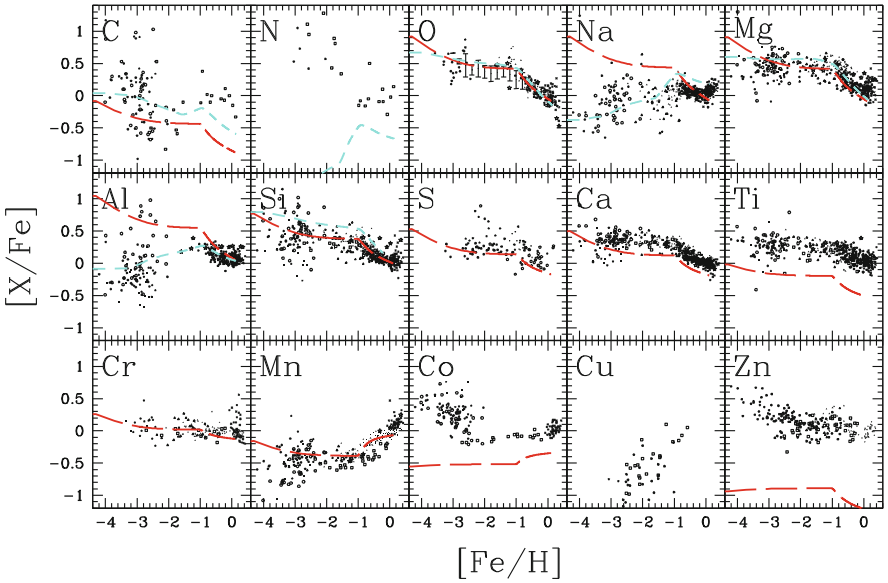


Fig. 1 Evolutions of heavy element abundance ratios [X/Fe] against [Fe/H] for one-zone models with only SNe II (*long-dashed line*), with our new yields and SN Ia model (*short-dashed line*), with the double-degenerate scenario of SNe Ia (*dotted line*), and with AGB stars (*solid line*). The *dots* are observational data (see K06 for the references)

low [Fe/H] is reproduced by the metallicity effect on nucleosynthesis. The iron-peak elements (Cr, Mn, Co, and Ni) are consistent with the observed mean values at $-2.5 \lesssim$ [Fe/H] $\lesssim -1$, and the observed trend at the lower metallicity can be explained by the energy effect under the assumption of inhomogeneous enrichment. Note that Cr II observations are plotted.

The most important improvement is in Zn. The observed abundance of Zn ([Zn/Fe] ~ 0) can be explained only by such large contribution of HNe. Since the observed [Zn/Fe] shows an increase toward lower metallicity, the HNe fraction may be larger in the earlier stage of galaxy formation. At high metallicity, since neutron-rich isotopes $^{66-70}$Zn are produced, the HNe fraction can be as small as 1%. In the following chemodynamical simulations, we adopt $\epsilon_{\rm HNe} = 0.5, 0.5, 0.4, 0.01, 0.01$ for $Z = 0, 0.001, 0.004, 0.02, 0.05$, which gives better agreement with the observed present HNe rate. Pair-instability supernovae, which produce much more Fe, more [S/Fe], and less [Zn/Fe] should not contribute in the galactic chemical evolution.

Type Ia supernovae (SNe Ia) – The progenitors of the majority of SNe Ia are most likely the Chandrasekhar (Ch) mass white dwarfs (WDs). For the evolution of accreting C+O WDs toward the Ch mass, two scenarios have been proposed; one is the double-degenerate scenario, i.e., merging of double C+O WDs with a combined mass surpassing the Ch mass limit. However, it has been theoretically

suggested that it leads to accretion-induced collapse rather than SNe Ia. The other is our single-degenerate (SD) scenario, i.e., the WD mass grows by accretion of hydrogen-rich matter via mass transfer from a binary companion.

We construct a new model of SNe Ia, based on the SD scenario, taking account of the metallicity dependences of the WD wind [17] and the mass-stripping effect on the binary companion star [15]. Our model naturally predicts that the SNe Ia lifetime distribution spans a range of $0.1 - 20$ Gyr with the double peaks; the main-sequence+WD systems with the timescale of $\sim$ 0.1–1 Gyr are dominant in star-forming galaxies, while the red-giants+WD systems with $\sim 1 - 20$ Gyr timescales are dominant in early-type galaxies.

Because of the metallicity effect, i.e., because of the lack of winds from WDs in the binary systems, the SNe Ia rate in the systems with [Fe/H] $\lesssim -1$, e.g., high-z spiral galaxies, is supposed to be very small. We succeed in reproducing the galactic supernova rates with their dependence on the morphological type of galaxies, and the cosmic SNe Ia rate history with a peak at $z \sim 1$. At $z \gtrsim 1$, the predicted SNe Ia rate decreases toward higher redshifts and SNe Ia will be observed only in the systems that have evolved with a short timescale of chemical enrichment. This suggests that the evolution effect in the supernova cosmology can be small.

From [Fe/H] ~ -1, SNe Ia start to occur producing more Fe than α-elements, and thus [α/Fe] decreases toward the solar abundance. The decreasing [Fe/H] depends on the SNe Ia progenitor model. Our SNe Ia model can give better reproduction of the [(α, Mn, Zn)/Fe]–[Fe/H] relations in the solar neighborhood than other models such as the double-degenerate (DD) scenario (dotted lines). With the DD scenario, the typical lifetimes of SNe Ia are $\sim$ 0.1 Gyr, which results in the too early decrease in [α/Fe] at [Fe/H] ~ -2. Even with our SD model, if we do not include the metallicity effect, [α/Fe] decreases too early because of the shortest lifetime, $\sim$ 0.1 Gyr. In other words, the metallicity effect is more strongly required in the presence of the young population of SNe Ia.

For SNe Ia, we take the nucleosynthesis yields from Nomoto et al. [25] where the metallicity dependence is not included. Ni is overproduced at [Fe/H] $\gtrsim -1$, which will be solved by tuning the propagation speed of the burning front and the central density of the white dwarf.

Asymptotic giant branch (AGB) stars – Stars with initial masses between about 0.8–8 $M_\odot$ (depending on metallicity) produce light elements such as C and N, while the contribution of heavier elements is negligible in the galactic chemical evolution (solid lines). The nucleosynthesis yields of AGB stars involves uncertainties due to convection and mass loss. We introduce the new calculation by Karakas [8], which is based on Karakas and Lattanzio [9]. The Na overproduction problem has been solved with the updated reaction rates.

Needless to say, the envelope mass and pre-existing heavy elements are returned by stellar winds from all stars. Surprisingly, in many hydrodynamical simulations, this contribution is not included, which should affect the metallicity and star formation history. Rotating massive stars ($\gtrsim 40\,M_\odot$) could also produce C and N, but are not included in our models.

3 Chemodynamical Code

The details of our GRAPE-SPH code are described in Kobayashi [11] and can be summarized as follows:

(i) The *gravity* is computed with the tree method using the special purpose computer GRAPE (GRAvity PipE) system at the National Astronomical Observatory of Japan. For hydrodynamics, the smoothed particle hydrodynamical (SPH) method is adopted, and the SPH formulation is almost same as Navarro and White [23]. The GRAPE-SPH code is highly adaptive in space and time by means of individual smoothing lengths and individual timestep, and it has very high performance (a week for one simulation with $N \sim 50,000$).
(ii) *Radiative cooling* is computed with the metal-dependent cooling functions, which are generated with the MAPPINGS III software [32] as a function of [Fe/H]. [O/Fe] is fixed with the observed [O/Fe]–[Fe/H] relation in the solar neighborhood.
(iii) Our *star formation* criteria are the same as Katz [10]; (1) convergent, (2) cooling, and (3) Jeans unstable. The SFR is determined from the Schmidt law; the star formation timescale is proportional to the dynamical timescale ($t_{\rm sf} \equiv \frac{1}{c} t_{\rm dyn}$), where $c = 0.1$ is chosen from the size-luminosity relation of elliptical galaxies [12]. We also adopt the probability criterion [10]. In our simulations, the SFR depends both on the local density and metallicity, which is different from other simplified models such as one-zone models.
(iv) If a gas particle satisfies the star formation criteria, a part of the mass of the gas particle turns into a star particle. Since one star particle has the mass of $10^{5-6}\,{\rm M}_\odot$, the star particle is not a single star but an association of many stars. The masses of individual stars span according to the *initial mass function* (IMF). We adopt the Salpeter IMF that is invariant to time and metallicity with a slope $x = 1.35$ for $0.07 - 120\,{\rm M}_\odot$.
(v) For the *feedback* of energy and heavy elements, we do not adopt the instantaneous recycling approximation. Via stellar winds, SNe II/HNe, and SNe Ia, thermal energy and heavy elements are ejected from an evolved star particle as a function of time and distributed to the surrounding gas particles within 1 kpc.
(vi) The *photometric evolution* of one star particle is identical to the evolution of a simple stellar population (SSP). SSP spectra are taken from Kodama and Arimoto [21] as a function of age and metallicity.

4 Cosmological Simulations

Hypernova feedback play an essential role not only on chemical enrichment but also on galactic dynamics. Since the resolution is not enough to describe the small-scale physics, the modeling of feedback involves a parameter. The model has to be checked with a different scale of observational constraints. The most stringent one is

the mass–metallicity relation of galaxies. We simulate the evolution of dark matter, gas, and stellar systems from the cosmological initial condition of a $10h^{-1}$ Mpc cubic box with periodic boundary, $N_{\rm DM} = N_{\rm gas} = 128^3$ particles, $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_{\rm b} = 0.04$, $n = 1$, and $\sigma_8 = 0.9$ (see K07 for the details). Figure 2 shows the time evolution of the densities of gas, stellar V-band luminosity, and gas metallicity. Star formation takes place in a distributed fashion at high redshifts and becomes most active around $z \sim 4$. The distribution of stars appears smooth at high redshifts, but concentrated at lower redshifts. Because of the feedback, massive galaxies are surrounded by hot gas and heavy elements are also distributed in the intergalactic medium. Both in the gas phase and stars, metallicity gradients are generated in high-density regions.

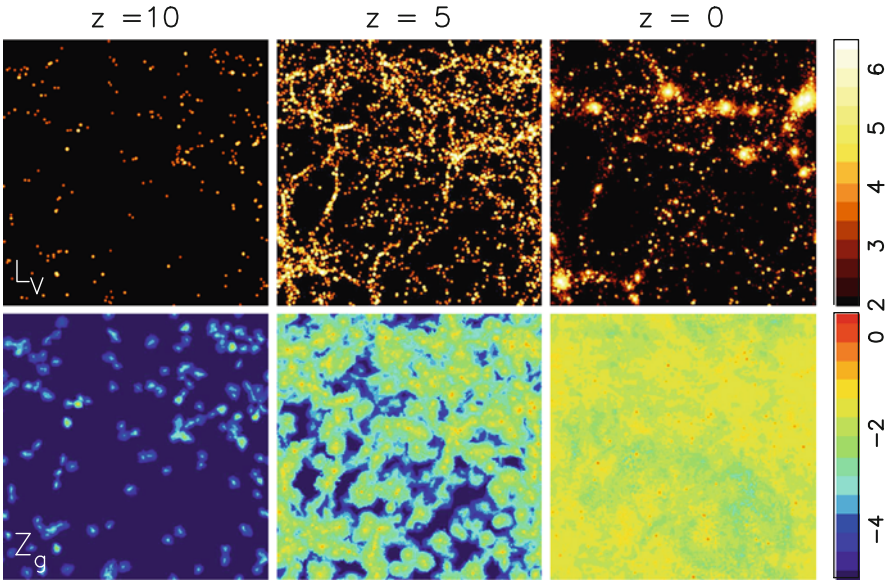


Fig. 2 The time evolution of our cosmological simulation in a periodic box $10\,h^{-1}$ Mpc on a side. We show the projected stellar V-luminosity (*upper panels*) and gas metallicity $\log Z_{\rm g}/Z_\odot$ (*lower panels*)

With the larger energy ejection by HNe, the SFR starts to be suppressed from $z \sim 6$ onward and is overall reduced by a factor of 3 at $0 \lesssim z \lesssim 3$. We then succeed in reproducing both the observed cosmic SFRs and the stellar density evolution (K07). The present stellar fraction is less than 10% being consistent with the recent observational estimate [5], while 25% of baryons turn into stars without feedback.

The metal enrichment timescale depends on the environment. In large galaxies, enrichment takes place so quickly that [O/H] reaches ~ -1 at $z \sim 7$, which is consistent with the sub-solar metallicities of the Lyman break galaxies [26]. The low metallicities of DLA systems [27] are also consistent with our galaxies, provided these systems are dwarf galaxies or the outskirts of massive galaxies. The low [C/H] of the IGM [30] can be explained if the IGM is enriched only by SNe II and HNe.

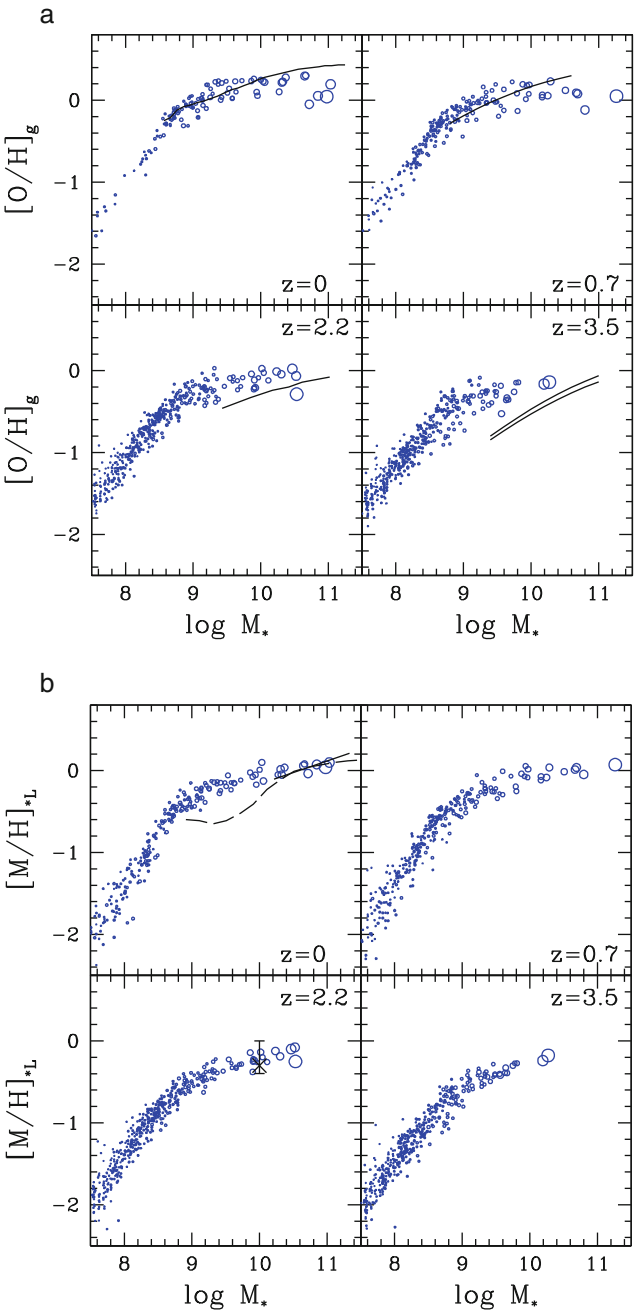


Fig. 3 **a** Mean metallicities of cold gas ($T < 10^4$ K) within 10 kpc, plotted against the total stellar mass. The *lines* are observational data: Tremonti et al. ([33], $z = 0$), Savaglio et al. ([29], $z = 0.7$), Erb et al. ([4], $z = 2.2$), and Maiolino et al. ([22], $z = 3.5$). **b** Mean stellar metallicities within 10 kpc, V-band luminosity-weighted. Kobayashi and Arimoto ([14], *solid line*) from metallicity gradients, Gallazzi et al. ([6], *dashed line*), and Quider et al. ([28], $z = 2 - 3$) for lensed galaxy

The average metallicity of the Universe reaches [O/H] ~ -2 and [Fe/H] ~ -2.5 at $z \sim 4$, but reaches the same values at $z \sim 3$ in the IGM.

How are heavy elements ejected from galaxies to the IGM? In the simulation, we can trace the orbit of gas particles over time. Exploiting this, we define as wind particles those that are not in galaxies now, but have been in galaxies before (Fig. 15 in K07). In this simulation, $\sim 10\%$ of baryons turn into stars, $\sim 10\%$ of the gas stays in galaxies ($\sim 8\%$ is hot), and $\sim 20\%$ is ejected as galactic winds. The rest, half of the baryons, never accretes onto galaxies. Tracing the orbits of gas particles, we can also examine from which galaxies the wind gas particles are ejected and measured the ejected wind mass from each galaxy (Fig. 16 in K07). Winds are efficiently ejected from small galaxies, with $\sim 80\%$ of accreted baryons being ejected from $M_{\rm tot} \sim 10^{11}{\rm M}_{\odot}$ galaxies. A similar relation is also found for the ejected metal fraction, i.e., the ratio between the wind metal mass to the total metal mass. It is interesting that the wind fraction and the ejected metal fraction correlate well with the stellar metallicity. Based on this finding, we conclude that the origin of the mass–metallicity relation is the mass-dependent galactic winds.

Figure 3 shows the mass–metallicity relations of simulated galaxies (dots) and observations (lines). For the gas-phase metallicity, there are uncertainties such as inhomogeneity and aperture effect and a large scatter is seen. The simulation is comparable to observations up to $z \sim 2$, but there is an offset at $z = 3.5$. This may be a problem, but the normalization of observations needs to be checked. For the stellar metallicity, a tight relation is present since $z = 5$ in the simulation. At high redshifts, observations are available for only two lensed galaxies, which is remarkably consistent with the simulation.

5 Elliptical Galaxies

For the formation of elliptical galaxies, two scenarios, monolithic collapse vs. major merger, have been debated. In Kobayashi [11, 12], we simulate the formation and chemodynamical evolution of 128 ellipticals from the CDM initial fluctuations, using the GRAPE-SPH code that includes star formation, supernovae feedback, and chemical enrichment. In our CDM-based scenario, galaxies form through the successive merging of subgalaxies with various masses.

We found that a half of ellipticals should form without major merger to explain the observed radial metallicity gradients [11]. In our chemodynamical simulations, metallicity gradients are destroyed by mergers and are not enough regenerated by induced star formation. In observations, there is a significant scatter (e.g., [14]) and both the average and the scatter are well reproduced with our simulations. The scatter stems from the difference in the merging histories. Galaxies that form monolithically have steeper gradients, while galaxies that undergo major mergers have shallower gradients. Spolaor et al. [31] found a tight relation between gradients and galaxy mass, which favors monolithic-like collapse, although some additional process is required to explain the flat gradient of dwarf ellipticals. The SAURON

observations also showed significant rotation in early-type galaxies and a tight relationship between the local Mg_2 index and velocity dispersion, which will also put constraint on the formation process.

In the case of major mergers, the surface brightness profile is affected and the effective radius becomes slightly larger than the case without major mergers. The scaling relations of elliptical galaxies are basically preserved, but an intrinsic scatter along the fundamental plane is originated from the difference in merging histories [12]. The compact galaxies at high redshifts [3] can be easily understood with this dynamical effect. Mass-to-light ratios and baryon fractions are also important constraints on the origin of elliptical galaxies [13].

6 The Milky Way Galaxy

In Kobayashi et al. [18] and Kobayashi and Nakasato [16], we simulate the chemodynamical evolution of the Milky Way-type galaxy from the CDM initial fluctuation (see [16], hereafter KN10, for the details). The initial condition is similar to those in Kobayashi [11, 12], but with the initial angular momentum of $\lambda \sim 0.1$, the total mass of $\sim 10^{12}\,M_\odot$, and $\sim 1,20,000$ particles. We choose an initial condition where the galaxy does not undergo major mergers, otherwise no disk galaxy can form. The cosmological parameters are set to be $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$.

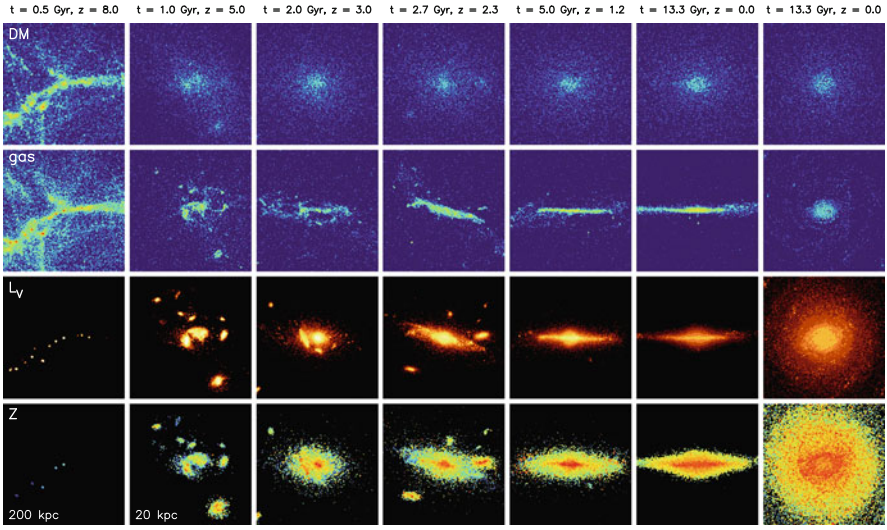


Fig. 4 Time evolution of the projected density of dark matter, gas, V-band luminosity and the luminosity-weighted stellar metallicity in the range of [M/H]$= -1$ to 0.1 at $z = 8, 5, 3, 2.3, 1.2$, and 0. The *leftmost panels* are 200 kpc and the *other panels* are 20 kpc on a side. The *rightmost* panels are for the face-on views at $z = 0$

Figure 4 shows the time evolution of the projected density of dark matter, gas, V-band luminosity, and luminosity-weighted stellar metallicity. After the start of the simulation, the system expands according to the Hubble flow. The CDM initial fluctuations grow into the structures of nodes and filaments, and small collapsed halos are realized both in dark matter and in gas. In the halos, the gas is allowed to cool radiatively and star formation takes place since $z \sim 15$. According to the hierarchical clustering of dark halos, subgalaxies merge to form large galaxies, which induces the initial starburst. Under the CDM picture, any galaxy forms through the successive merging of subgalaxies with various masses. In this simulated galaxy, the bulge is formed by the initial starburst that is induced by the assembly of gas-rich sub-galaxies with stellar masses of $\sim 5 - 10 \times 10^9\,M_\odot$ and gas fractions of $0.2-0.4$ at $z \gtrsim 3$. Because of the angular momentum, the gas accretes onto the plane forming a rotationally supported disk that grows from inside out. In the disk, star formation takes place with a longer timescale, which is maintained not by the slow gas accretion, but by the self-regulation due to supernova feedback. Many satellite galaxies successively come in and disrupt, but there is no major merger event after $z \sim 2$, which is necessary to retain the disk structure. Metallicity gradients, increasing toward higher density regions, are generated both in the gas phase and in the stars from $z \sim 5$ onward. The galaxy center and the disk plane are determined from the center of gravity and the aliment of the angular momentum vector, respectively.

The bulge has a de Vaucouleurs surface brightness profile with an effective radius of ~ 1.5 kpc and the disk has an exponential profile with a scale length of ~ 5 kpc. In the surface brightness profile, an excess is seen at $r \sim 12$ kpc, which corresponds to the remnant of satellite galaxies. The bulge in the simulated galaxy seems to be larger than the Galactic bulge and these properties may be more consistent with M31. The total masses inside 1, 10, and 100 kpc are $9.4 \times 10^9\,M_\odot$, $2.2 \times 10^{11}\,M_\odot$, and $1.1 \times 10^{12}\,M_\odot$, respectively. The baryon fraction is ~ 0.65 at $r < 2$ kpc and then decreases to ~ 0.2 at $r > 8$ kpc. The rotation velocity of the disk rapidly increases to reach $\sim 230\,\mathrm{km\,s^{-1}}$ at $r \sim 3$ kpc, stays constant until $r \lesssim 20$ kpc, and then gradually decreases until $r \sim 40$ kpc. This is roughly consistent with the observed rotation curve of the Milky Way Galaxy, which shows a plateau at the local standard of rest ($180 - 250\,\mathrm{km\,s^{-1}}$) for $3 - 6 \lesssim r \lesssim 20 - 60$ kpc.

In the following, we define the three major components simply from the location of the stars at the present-day: the radius of $7.5 \le r \le 8.5$ kpc and the height of $|z| \le 0.5$ kpc for the solar neighborhood, $r \le 1$ kpc for the bulge and $5 \le r \le 10$ kpc for the halo. The thick disk stars are defined from the kinematics: the ratio between rotation velocity and velocity dispersion $v/\sigma < 1$ in the solar neighborhood.

The resultant star formation histories are different for different components. In the bulge, 80% of stars are older than 10 Gyr, and 60% have [O/Fe] > 0.3. In the disk, 50% of solar neighborhood stars are younger than 8 Gyr and 80% have [O/Fe] < 0.3. The thick disk stars tend to be older and have higher [α/Fe] than the thin disk stars. The formation timescale of the thick disk is 4 Gyr in this simulation. In the halo, most of stars are as old as the bulge, although the numerical resolution of this

simulation is not enough to discuss the halo in detail. However, there does not seem to be very much difference in the SFR between the inner and the outer halos.

6.1 Age–Metallicity Relations

The chemical enrichment timescale is also different for the different components. Figure 5 shows the age–metallicity relations. In the solar neighborhood, the average metallicity reaches [Fe/H] ~ 0 at $t \sim 2$ Gyr and does not show strong evolution for

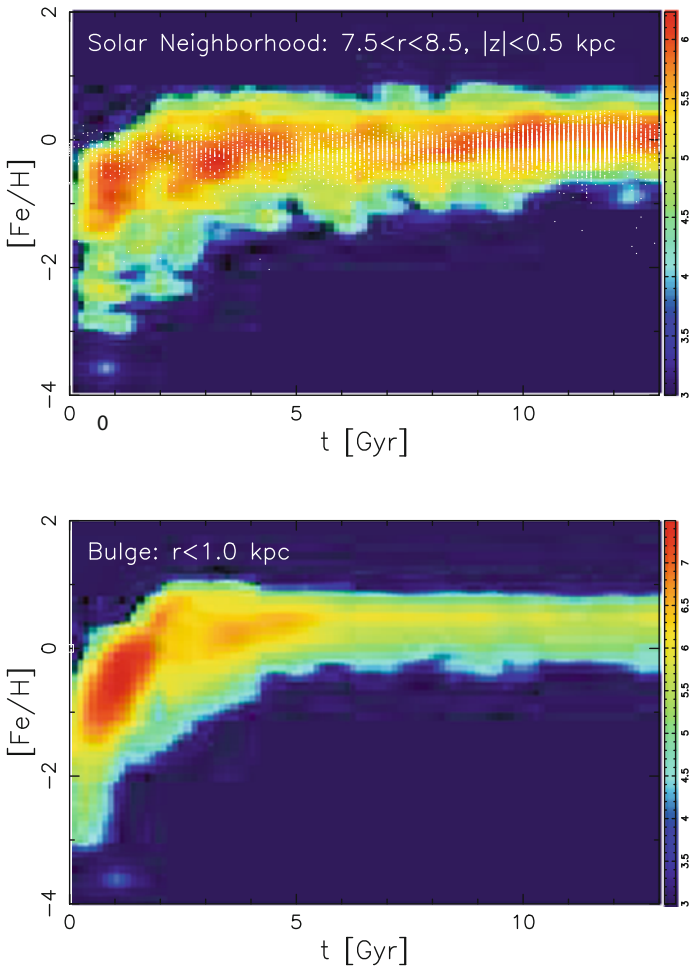


Fig. 5 Age−metallicity relations in the solar neighborhood (*top panel*) and bulge (*bottom panel*). The *contours* show the frequency distribution of stars in the simulated galaxy and *red* is for the highest frequency. The *white dots* show the observations in the solar neighborhood [7]

$t \gtrsim 2\,\mathrm{Gyr}$. The scatter in metallicity at given time is caused by the inhomogeneity of chemical enrichment in our chemodynamical model; there is a local variation in star formation, metal production by supernovae, and metal flow by the inflow and outflow of the ISM. As a result, both the average and the scatter are in excellent agreement with the observations (dots) in spite of the uncertainties in the observational estimates of the ages.

In the bulge, star formation takes place more quickly and thus the chemical enrichment timescale is much shorter than in the disk. The age–metallicity relation shows a more rapid increase than in the disk. The maximum metallicity reaches super solar ([Fe/H] $\sim$ 1) at $t \sim 2\,\mathrm{Gyr}$. Although the SFR becomes small after $\sim 5\,\mathrm{Gyr}$, a few stars form at $\gtrsim 5\,\mathrm{Gyr}$. These have super-solar metallicity in general and the average metallicity do not show time evolution.

6.2 [α/Fe] – [Fe/H] Relations

The difference in the chemical enrichment timescales results in a difference in the elemental abundance ratios. The best-known clock is the α-elements to iron ratio [α/Fe] since SNe Ia produce more iron than α-elements with longer timescales than SNe II. Nevertheless, it should be noted that low-mass SNe II ($10 - 13\,\mathrm{M}_\odot$) also provide relatively low [α/Fe] because of their smaller envelope mass compared to more massive SNe II. This mass dependence is also important in dwarf spheroidal galaxies.

Figure 6 shows the [O/Fe]–[Fe/H] relations, and the other α-elements show the same trends. In the solar neighborhood, at the beginning, only SNe II contributes, and [α/Fe] shows a plateau ([α/Fe]$\sim$ 0.5). Around [Fe/H] ~ -1, SNe Ia starts to occur, which produces more iron than α-elements. This delayed enrichment of SNe Ia causes the decreasing trend in [α/Fe] for higher [Fe/H]. This trend is in great agreement with the observations (dots). A significant scatter is seen, which is caused by the inhomogeneity of chemical enrichment in our chemodynamical model. At [Fe/H] $\gtrsim -1$, the majority of [O/Fe] is lower than 0.2 and the peak [O/Fe] is -0.15. Around [Fe/H] ~ -1, the scatter looks a bit larger than observed (Fig. 6). This may be because the mixing of heavy elements among gas particles is not included in our model. At [Fe/H] $\lesssim -1$, the scatter is caused from SNe Ia and/or low-mass SNe II in the inhomogeneous enrichment in our chemodynamical models. The [α/Fe] scatter in the simulation could be larger, since there is also a variation depending on the explosion energy and the remnant mass (neutron star and black-hole) for SNe II and HNe (faint SNe). In other words, statistical comparison with the observed scatter could provide constraints on the unsolved physics of supernova explosion.

Using stochastic chemical evolution model, Argast et al. [1] claimed that the scatter of [α/Fe] caused by supernova yields is too large compared to observations and an efficient mixing process is required. This is because the nucleosynthesis yields adopted in that paper [24] were different from currently accepted values. The low [α/Fe] values resulted from the larger amount of iron assumed for low-mass

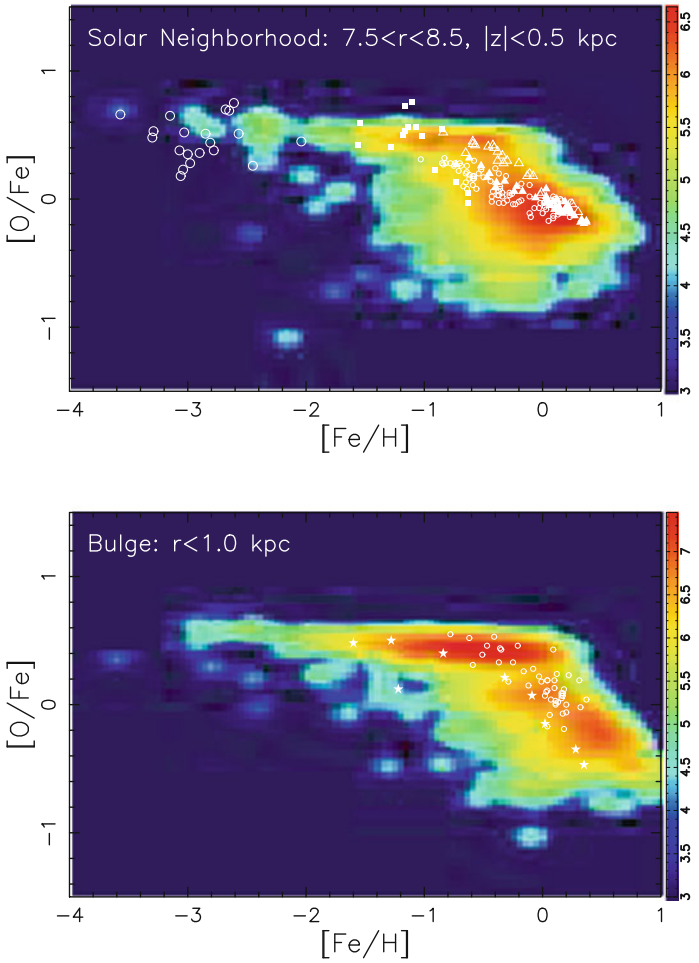


Fig. 6 [O/Fe]−[Fe/H] relations in the solar neighborhood (*top panel*) and bulge (*bottom panel*). The *contours* show our simulation and the *dots* are observational data (see [15] for the references)

SNe II (0.15 $M_\odot$ for 13–15 $M_\odot$). Now the iron mass is estimated from the light curve and spectral fitting and is 0.07 $M_\odot$ for SNe II with $< 20\,M_\odot$. The high [α/Fe] values are reduced by including hypernovae. With hypernovae, the scatter around [O/Fe] ~ 0.5 is as small as the observations [2].

In the bulge, the [α/Fe]−[Fe/H] relation is very much different. The chemical enrichment timescale is so short that the metallicity reaches super solar before SNe Ia contributes. Thus, the [α/Fe] plateau continues to [Fe/H] $\sim +0.3$. This is roughly consistent with the observations. In this simulated galaxy, some new stars are still forming in the bulge, which in general have super solar metallicity and low [α/Fe] because of the large contribution from SNe Ia. Although a small fraction of stars with the age of ~ 3 Gyr ($t \sim 10.3$ Gyr) do have high [α/Fe] ([O/Fe] $= 0.3$),

which is caused by the inhomogeneous enrichment, stars younger than 1 Gyr have $-0.5 \leq$ [O/Fe] ≤ 0 and $0 \leq$ [Fe/H] ≤ 0.8.

6.3 [X/Fe]–[Fe/H] Diagrams

Finally, using our chemodynamical simulation, we predict the frequency distributions of the elements from O to Zn as a function of time and location. Figure 7 shows the mass density of stars in the [X/Fe]–[Fe/H] diagrams for the solar neighborhood (see [15] for the bulge and thick disk) at present. Because of the delayed enrichment of SNe Ia, α-elements (O, Mg, Si, S, and Ca) show a plateau at [Fe/H] ~ -1, and then the decreasing trend against [Fe/H], where [Mn/Fe] also shows the increasing trend. Odd-Z elements (Na, Al, and Cu) show the increasing trend at [Fe/H] $\lesssim -1$ because of the metallicity dependence of nucleosynthesis yields. These are in excellent agreement with the available observations.

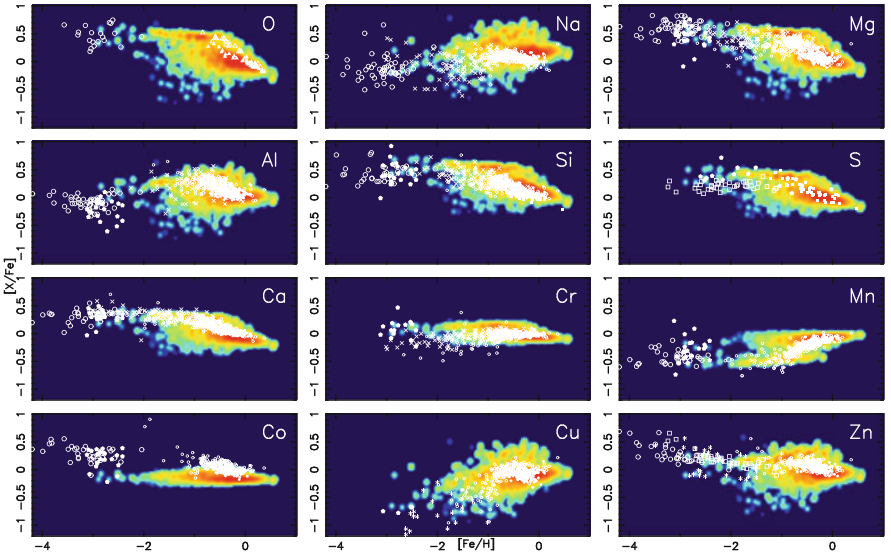


Fig. 7 [X/Fe]−[Fe/H] relations in the solar neighborhood. The *contours* show our simulation and the *dots* are observational data (see [15] for the references)

In the bulge, the star formation timescale is so short that the [α/Fe] plateau continues to [Fe/H] $\sim +0.3$. Because of the smaller contribution from SNe Ia, the majority of stars show high [α/Fe] and low [Mn/Fe]. [(Na, Al, Cu, Zn)/Fe] are also high because of the high metallicity in the bulge.

The stellar population of the thick disk is neither disk-like nor bulge-like as summarized in Table 1. In the thick disk, [α/Fe] is as high, and [Mn/Fe] is as low, as in the bulge because of the short formation timescale. However, [(Na, Al, Cu, Zn)/Fe]

Table 1 Stellar populations in each component

	Age	[(O, Mg, Si, S, Ca)/Fe]	[(Na, Al, Cu, Zn)/Fe]	[Mn/Fe]
Solar neighborhood	Young	Low	~ 0	~ 0
Bulge	Old	High	High	Low
Thick disk	Old	High	~ 0	Low

are not as high as in the bulge because of the lower chemical enrichment efficiency. This is because half of the thick disk stars have already formed in satellite galaxies before they accrete onto the disk and the metals have been ejected from the satellite galaxies by the galactic winds.

6.4 [α/Fe]−[Mn/Fe] Diagram

In Fig. 8, for the solar neighborhood, [Mn/Fe] is plotted against [α/Fe]=([O/Fe] + [Mg/Fe])/2, which clearly shows the sequence of the SNe Ia contribution. With SNe II and HNe only, [α/Fe] is as high as ~ 0.5 and [Mn/Fe] is as low as ~ -0.5. With more SNe Ia, [α/Fe] decreases, while [Mn/Fe] increases. The three populations of the observed stars are along this trend; (i) the EMP stars (large open circles and filled pentagons) are found in the left-bottom region with high [α/Fe] and low [Mn/Fe]. (ii) The thick disk stars (small open circles) populate the following region, [α/Fe] $\sim$ 0.2–0.4 and [Mn/Fe] ~ -0.4 to -0.2. (iii) The thin disks stars (small closed circles) occur at [α/Fe] ~ 0.1 and [Mn/Fe] ~ -0.1, which are formed from the ISM largely enriched by SNe Ia. In other words, it is possible to select thick disk stars only from the elemental abundance ratios.

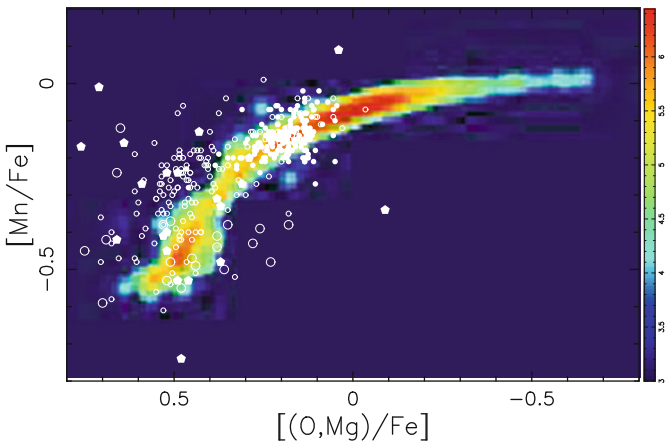


Fig. 8 The [α/Fe]−[Mn/Fe] relation in the solar neighborhood. The *contours* show our simulation and the *dots* are observational data (see [15] for the references)

References

1. Argast, D., Samland, M., Thielemann, F.-K., Gerhard, O.E. 2002, A&A, 388, 842
2. Cayrel, R., et al. 2004, A&A, 416, 1117
3. Damjanov, I., et al. 2009, ApJ, 695, 101
4. Erb, D.K., Shapley, A.E., Pettini, M., Steidel, C.C., Reddy, N.A., Adelberger, K.L. 2006, ApJ, 644, 813
5. Fukugita, M., Peebles, P.J.E. 2004, ApJ, 616, 643
6. Gallazzi, A., Charlot, S., Brinchmann, J., White, S.D.M., Tremonti, C.A. 2005, MNRAS, 362, 41
7. Holmberg, J., Nordström, B., Andersen, J. 2007, A&A, 475, 519
8. Karakas, A. 2010, MNRAS, 403, 1413
9. Karakas, A., Lattanzio, J.C. 2007, PASA, 24, 103
10. Katz, N. 1992, ApJ, 391, 502
11. Kobayashi, C. 2004, MNRAS, 347, 740
12. Kobayashi, C. 2005, MNRAS, 361, 1216
13. Kobayashi, C. 2009, in IAU JD 8, eds. D. Kim et al., in press
14. Kobayashi, C., Arimoto, N. 1999, ApJ, 527, 573
15. Kobayashi, C., Nomoto, K. 2009, ApJ, 707, 1466
16. Kobayashi, C. Nakasato, N. 2010, ApJ, submitted (KN10)
17. Kobayashi, C., Tsujimoto, T., Nomoto, K., Hachisu, I., Kato, M. 1998, ApJ, 503, L155
18. Kobayashi, C., Nakasato, N., Nomoto, K. 2003, In: IAU Symp. 208, eds. J. Makino, P. Hut, Tokyo, Japan, p. 419
19. Kobayashi, C., Umeda, H., Nomoto, K., Tominaga, N., Ohkubo, T. 2006, ApJ, 653, 1145 (K06)
20. Kobayashi, C. Springel, V., White, S.D.M. 2007, MNRAS, 376, 1465 (K07)
21. Kodama, T., Arimoto, N. 1997, A&A, 320, 41
22. Maiolino, R., et al. 2008, A&A, 488, 463
23. Navarro, J.F., White, S.D.M. 1993, MNRAS, 265, 271
24. Nomoto, K., et al. 1997, Nuclear Physics, A616, 79c
25. Nomoto, K., et al. 1997, Nuclear Physics, A621, 467c
26. Pettini, M., et al. 2001, ApJ, 554, 981
27. Prochaska, J., Naumov, S.O., Carney, B.W., McWilliam, A., Wolfe, A.M. 2000, AJ, 120, 2513
28. Quider, A.M., Pettini, M., Shapley, A.E., Steidel, C.C. 2009, MNRAS, 398, 1263
29. Savaglio, S., et al. 2005, ApJ, 635, 260
30. Schaye, J., et al. 2003, ApJ, 596, 768
31. Spolaor, M., Kobayashi, C., Forbes, D.A., Couch, W.J., Hau, G.K., 2009, MNRAS, in press
32. Sutherland, R.S., Dopita, M.A. 1993, ApJS, 88, 235
33. Tremonti, C.A., et al. 2004, ApJ, 613, 898

Elemental Abundance Patterns of Disk Substructure

Gayandhi De Silva

Abstract The billions of stars that reside in the Milky Way contain within themselves fossil information about the processes that shaped our Galaxy. As the Galactic disk underwent a dissipative formation and a dynamical evolution, much of the stellar dynamical information has been lost. The chemical information in the form of elemental abundance patterns remains preserved. Recovering such fossil stellar substructures is the goal of the expanding field of Galactic Archaeology. However, the existence of truly relic stellar groups within the Milky Way disk has been a topic of controversy. It is unclear if co-moving groups of stars are the remnants of earlier epochs of star formation, accretion events, or results of disk dynamics. I will present some recent high-resolution spectroscopic studies of several stellar groups in the Galactic disk. I will compare their chemical properties against disk open clusters, which are the likely leftovers of earlier star formation events, and against solar neighborhood stars to examine the true origins of these disk substructures.

1 Introduction

The Milky Way contains billions and billions of stars. Over 90% of the Galaxy's stellar content is found in its disk, scattered into the field as a result of the disk's dissipative formation and subsequent dynamical evolution. Within the global model of hierarchical structure formation, we do not have a complete understanding of how large disk galaxies like ours came to existence. Observations indicate that stellar and gaseous disks might not be uncommon at higher redshifts, while theoretical models within the ΛCDM paradigm have problems explaining the presence of unperturbed large disks at the same redshifts. Since a stellar disk is a defining component of our Galaxy and all other spiral galaxies, understanding disk formation is fundamental to understanding galaxy formation.

G. De Silva (✉)
Anglo Australian Observatory, Sydney, NSW 2122, Australia
e-mail: gdesilva@aao.gov.au

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_24,

Many disk galaxies including our own Galaxy show a system of two disks, a thin and thick disk. However, the thick disk component is not always observed and cannot be considered an essential part of galaxy formation. Both dynamically and morphologically the thin disk is by far the most important. Thus it is the thin disk that is defining the diskyness of any disk galaxy. This chapter will focus on the study of the Milky Way's thin disk.

To obtain a detailed physical understanding of the events that shaped the Galactic disk, we examine its individual stars. The chemical abundances as measured from stellar photospheres can be indicative of the conditions of a star's birth site. The incidence of internal and external pollution within stellar clusters is found to be negligible among the long-lived low-mass stars [7]. Their photospheric abundances should match the elemental abundance ratios of the proto-stellar gas cloud from which they formed. Different chemical elements are synthesized during different stages of stellar evolution and distributed to the interstellar medium via various process such as supernova explosions (SNe) or stellar winds. By studying the abundances of these elements we can estimate the frequency and importance of each synthesis process. Therefore, the chemical abundance patterns within the disk stars can be regarded as a fossil imprint and hold key information of the events that took place during the disk's formation and evolutionary history.

Studies by Edvardsson et al. [12] and other such works (e.g. [1, 4, 33]) on the chemodynamical evolution of the Galactic disk show how trends in various chemical elements can be used to resolve disk structure. Detailed chemical abundance patterns of the disk stars offer the possibility to reconstruct components of the protogalactic disk and so improve our basic understanding of the disk formation process. The long-term goal of Galactic Archaeology is to reconstruct the original star-forming events of the Galaxy. Employing the technique of chemical tagging [27] to disk field stars, one can identify now dispersed ancient stellar aggregates, the building blocks of the Galactic disk.

1.1 Disk Substructure

Ubiquitous throughout the disk, woven through the background field stars, are star clusters which represent the sites of star formation events. These disk clusters, referred to as open clusters, are considered the ideal stellar laboratories for investigating Galactic chemical evolution [28]. The newly formed young clusters provide information about the current star formation conditions, while studies of the intermediate-aged clusters together with the older open clusters provide us with a timeline of change. However, it has long been recognized that open clusters lead a life of uncertainty between survival and disruption. It is expected that most open cluster stars will dissolve into the general disk population within the first billion years. Therefore, old (> 1 Gyr) open clusters, which are the most informative about the early stages of disk formation, are hard to come by.

Out of the old open cluster population of the disk, a small number of extremely old clusters with ages of $8 - 10$ Gyr also exist (e.g., M67, Collinder 261, NGC

6791). Their survival as bound clusters suggests that they had high star formation efficiencies, which helped to keep the cluster stars bound, or are the leftover cores of originally larger clusters, where the outer stars have been dispersed into the field. These clusters are as old as the disk itself and elemental abundances of these rare objects provide a wealth of information about the conditions of the disk during its assembly, including possible merger events.

Other structures exist in the disk which are also of significant importance. These are the remnants of dispersed star-forming events, known as moving groups and superclusters, where the stellar members are unbound but share a common motion around the Galaxy. Analogous to stellar streams in the Galaxy halo which are remnants of merged satellites, old moving groups, and superclusters represent the ancient stellar building blocks of the disk. In his pioneering studies, Olin Eggen identified several moving groups and superclusters [13–24]. With the availability of the Hipparcos catalog of parallaxes and proper motions it became possible to study these substructures in the Milky Way disk in unprecedented detail. Several studies have analyzed the Hipparcos catalog and re-identified the Eggen groups as well as found new ones. Essentially, all of these studies have focused purely on the dynamical/kinematical nature of the substructures in the disk.

The moving groups and superclusters thus provide substructures within the stellar disk that can be identified through kinematics. However, it is not yet clear if all, or indeed the majority, of these substructures actually originate from star-forming clusters that have or are in the process of being dispersed. Many studies indicate that such substructures can equally well be the result of the dynamical interaction between the stars in the field and the bar in the central parts of the Milky Way (see e.g. [5]). Therefore not all kinematically defined disk substructures represent a prior star formation event. In this chapter we highlight the spectroscopically defined substructures of the disk. In Section 2 we discuss the commonly studied chemical elements and the general methodology for deriving their abundances, as well as the uncertainties involved. In the subsequent sections we review the currently known abundance patterns of old open clusters, the abundance patterns of moving groups and superclusters, including new results on the Hyades supercluster and the Argus association.

2 Elemental Abundances

Unlike in the halo where stellar streams are easily identifiable via their dynamics, much of the dynamical information of the disk stars has been *scrambled* during the dissipative formation of the disk. The dynamical information of disk stars is limited to the stars last scattering event, i.e., the last event that modified the stars original kinematics (interactions with molecular clouds, scattering due to spiral arm resonances, etc.). The information imprinted in the stellar elemental abundances has survived the disk's dissipative history. Thus it should be possible to establish the uniqueness of disk substructure by studying the internal spread in elemental abundances and comparing their abundance patterns to the patterns in the field disk stars.

The chemical elements of interest can be divided into several different groups. This includes the light elements (e.g., C, N), the odd-Z elements (e.g., Na, Al), alpha-elements (e.g., Ca, Si), Fe-peak elements (e.g., Ni, Co), the slow neutron capture (s-process) elements (e.g., Zr, Ba), and the rapid neutron capture (r-process) elements (e.g., Eu, Gd). Within these broad groups, there are observed abundance oddities, such as Mg which does not always follow the other alpha-element abundance ratios, Ti which shows both alpha- and Fe-peak element trends, and subgroups such as light and heavy s-process elements. Several dominant mechanisms are responsible for the synthesis of the different element groups. The lighter elements are synthesized in stellar interiors via the p–p chain and subsequent CNO cycle. The alpha-elements are predominantly produced from Type II SNe, while the Fe-peak elements are produced in Type Ia SNe. Of the heavy elements the s-process is mostly produced from the He-burning phase of intermediate to low-mass AGB stars, while the likely site of r-process synthesis is Type II SNe [3, 37].

However, there may not always be one process involved in the production of individual elements, rather it is likely that there are several contributors in the synthesis of the elements. Supernova models produce different yields depending on the mass of the progenitor, metallicity, assumed mass cut, and other details of the detonation process. Models of stellar interiors show that other mixing processes in giant stars may result in extra production of lighter elements. The effects of extra mixing in giant stars can result in changes to their photospheric abundances. In this case the measured values would not represent conditions of the stars original birth. However, such mixing processes are not expected in the low-mass dwarf stars as discussed in Section 1. The imprints of the processes undergone in the massive stars would be present in any subsequent generations of low-mass stars that formed after the massive stars have enriched the ISM either via stellar winds or SNe.

2.1 High-Resolution Spectroscopy

To obtain detailed elemental abundances for a large range of chemical elements, high-resolution spectroscopy is a must. For many alpha- and Fe-peak elements a resolution of ∼30,000 will be sufficient to derive abundances to within 0.1 dex internal accuracy. Many neutron capture elements have weak lines with width <10 mÅ. For such lines a resolving power of ∼60,000 with a signal-to-noise ratio of over 100 per resolution element is necessary to reach target accuracy.

With the required quality data, the general method of determining abundances involves the measurement of the equivalent widths (EWs) of the individual lines. Stellar parameters are estimated either via photometry or spectroscopically. Often photometric stellar parameters are used as a starting point for the iterative process of spectroscopic determinations. The most common approach is by requiring excitation equilibrium of Fe I lines to derive the effective temperature (T_{eff}) and ionization equilibrium of Fe I and Fe II lines to derive surface gravity (log g). The microturbulence ζ is determined from the condition that Fe I lines show no trend versus

EW. The resulting T_{eff}, log g, ζ, and metallicity are then adopted to generate model stellar atmospheres.

Note that the spectroscopic derivation of the stellar parameters is affected by the atomic data of the lines chosen as well as the number of lines chosen to achieve excitation and ionization balance. Also, the parameters derived spectroscopically can differ from the photometric values by varying amount for the different types of stars. Therefore, when comparing the abundance results of different studies, the difference in deriving the stellar parameters is likely to give rise to systematic offsets.

The abundances of other elements are derived using the selected model atmospheres and the measured EWs of the lines. Most studies are able to measure several lines for a given element, especially for the numerous alpha- and Fe-peak elements. However, for some elements, only one line maybe present in the available wavelength windows. In cases of blended line features, EW measurements may not be possible, the abundances are derived via spectral synthesis. There are several sources of systematic effects, including the choice model atmospheres, the origin of the line atomic data, the adopted solar or reference value, and if departure for LTE effects have been considered. When comparing literature abundances it is important to account for such systematic effects before analyzing the results.

3 Open Clusters

Stars are born in clusters and all stars in an open cluster have the same age and chemical composition [7, 9]. The Milky Way's open cluster population span from the newborn clusters as observed in the current star-forming regions to old clusters with ages comparable to the age of the disk. Despite the large number of known open clusters ($\sim$ 1700; [11]), only a small fraction has been subject to high-resolution spectroscopic studies, although several recent studies and surveys (such as the BOCCE survey) are now targeting more open clusters.

No age–metallicity relation has been observed among the open cluster population, where the oldest clusters can be as metal rich as the younger counterparts [28]. This is contrary to what is observed among the disk field stars and cannot be explained with chemodynamical models. De Silva et al. [10] compiled high-resolution literature studies of old open clusters and showed that large cluster-to-cluster abundance variations exist. While some of the scatter was due to systematic effects, some elements showed large variations which could not be explained by systematics alone. Further, various element-to-element abundance patterns were seen among their sample, indicating the existence of chemical signatures unique to the clusters based on their site of formation. This is likely due to localized inhomogeneities at the site of formation and highlights the different conditions during the formation of the individual clusters.

A radial metallicity gradient as traced by open clusters shows a general decrease in metallicity with increasing galactocentric radius, which eventually flattens out after about 12 kpc [28, 29, 38]. However, [36] claimed evidence for a stepped radial

metallicity distribution with the first plateaux at solar level metallicity until about 10 kpc and the second at [Fe/H] ~ -0.3 in the outer disk. Recently Pancino et al. [31] obtained high-resolution abundances for five old open clusters and combined their results with other high resolution literature studies. Consistent with earlier studies they find a decreasing radial metallicity gradient until $R_{GC} \sim 12$ kpc, after which it flattens, which is also consistent with disk cepheids and B stars.

The presence of alpha-element enhancements with Galactocentric radius among the open clusters is also of interest. Work by Yong et al. [38] indicated an increased alpha-enhancement at larger radii, indication of a period of rapid star formation in the outer disk possibly linking to a merger event. However, there have not been enough data to confirm the claims and require further study of the outer disk clusters.

4 Moving Groups and Superclusters

The concept of moving groups and superclusters was first advocated by Olin Eggen in the 1960s. Basically the stars form from a common progenitor gas cloud. As the cluster orbits around the Galaxy, it disperses into a tube-like structure around the Galaxy plane, and after several galactic orbits, will dissolve into the Galaxy background. The tube-like unbound groups of stars occupying extended regions of the Galaxy were defined by Eggen as superclusters. If the Sun happens to be inside this tube structure, the group members will appear to be spread over the sky, but may be identified as a group through their common space velocities. These group stars located within the solar neighborhood were defined as a moving group, and believed to be a subset of larger systems known as stellar superclusters.

In a long series of papers, Eggen identified a number of moving groups and superclusters; the most well known are the Hyades, Pleiades, Sirius, Arcturus, and HR1614 moving groups and superclusters. His groups were not well received by the astronomical community at the time due to the lack of accurate distance measurements and Eggen's inconsistent membership lists. With the availability of accurate parallaxes and proper motions from the Hipparcos satellite, several authors demonstrated the existence of dynamically defined stellar groups, some coinciding with those advocated by Eggen.

4.1 HR1614 Moving Group

The HR1614 moving group is the only published moving group that has been identified as homogeneous in elemental abundances and age [8, 26]. Eggen [17] initially identified the group by studying stars with a V-velocity within $\pm$ 10 km/s of the star HR1614. Feltzing and Holmberg [26] using Hipparcos parallaxes recovered a population of metal-rich stars at U = 10 km/s, V = -60 km/s, and tilted in the UV

plane, which they associated with the HR1614 moving group. This group is quite distinctive due to its age of about 2 Giga year and high metallicity of [Fe/H] = +0.25. Its stellar motions suggest that these stars formed in the inner disk [6].

De Silva et al. [8] using high-resolution spectra demonstrated that there moving group stars are chemically homogeneous. Their differential abundance study found that internal star-to-star scatter does not exceed 0.01 dex in [Fe/H]. A similar level of abundance homogeneity was seen over a range of other elements including odd-Z, alpha-, Fe-peak, and s- and r-process elements. This level of chemical homogeneity is expected for stars which formed from a common site, as is the case in open cluster stars. The fact that dispersed, unbound group of stars also show such a high degree of chemical homogeneity implies that these stars originally formed together in a cluster from a common parent gas cloud. Therefore, HR1614 is indeed a true relic of an earlier star-forming event.

Not all stars that lie in the same UV velocity plane showed this level of homogeneity. In the sample studied by De Silva et al. [8], four stars were found to show solar-level abundances and deviated greatly from the majority of their sample. They argued that these stars were non-members of the moving group due to their positions on the CMD, variability, and that their abundances matched the disk field abundances. Bobylev et al. [2] showed that there is a large range in age within the kinematically defined stellar groups. Therefore, it is important to note that the true moving group, as identified via chemical homogeneity, is a subset masked behind the larger kinematically defined structures.

4.2 Hyades Supercluster

The existence of a kinematically defined group of stars associated with the Hyades open cluster has been suggested in the 1970s by Olin Eggen and established more recently by Galactic field surveys. It is supported by dynamical simulations (e.g., [5]), where a concentration of stars are found at U $\sim$ 40 km/s and V $\sim$ -20 km/s in the $U - V$ velocity plane. However, Famaey et al. [25] have suggested that the kinematically defined Hyades supercluster consists of a mixture of field stars which have been kinematically swept up by a passing spiral density wave and is not part of a dispersed stellar aggregated.

Eggen identified the stellar members of the Hyades supercluster as part of a larger star-forming event, where the outer cluster stars have dispersed into a supercluster, leaving only its core members in the present-day Hyades open cluster. With kinematics alone it is unclear if the Hyades supercluster is the debris of an earlier star-forming event, or an artifact of dynamical origin. As no high resolution abundances study had targeted the Hyades supercluster in the literature, we initiated such a preliminary abundance study.

We selected putative members of the Hyades supercluster by exploring the stars along the line of sight of the likely dispersion orbits of the larger Hyades star-forming aggregate. In this work, the adopted velocity of Hyades supercluster stars

near the Sun was (U = 40, V = −17) km/s, consistent with the currently accepted motion of the Hyades open cluster and the overdensity regions seen in Galactic disk surveys such as the Geneva-Copenhagen survey. High-resolution (R ∼ 60,000), high signal-to-noise (S/N ∼ 100) spectra were collected using UVES at Paranal Observatory for a sample of stars in the line of sight of the dispersion orbits which had known metallicities of [Fe/H] > −0.2.

Abundances were measured for a range of elements. We find that seven stars have metallicities comparable to the Hyades open cluster, while the remaining stars have solar-level abundances. Figure 1 shows [Fe/H] versus the stellar effective temperatures for the sample of stars. The dotted line marks the solar level while the dashed line marks the Hyades metallicity of [Fe/H] = 0.13 [32]. Unlike in the case of HR1614 moving group we do not find a high level of homogeneity in the Hyades supercluster sample. Most of the stars lie about the solar metallicity and are likely to be field stars. The fact that seven stars have metallicities comparable to the Hyades open cluster may be used to support the existence of dispersed stars of Hyades origin within the kinematical group. However, since the disk field also hosts similarly metal-rich stars, it is not possible to clearly differentiate the supercluster stars as dispersing Hyades cluster stars or those from the general field. Certainly further study into extended samples of the supercluster is necessary to fully address the origin the Hyades supercluster. Larger sample showing little or no star-to-star scatter about the Hyades open cluster metallicity would provide stronger evidence of the existence of dispersing stars of Hyades cluster origin.

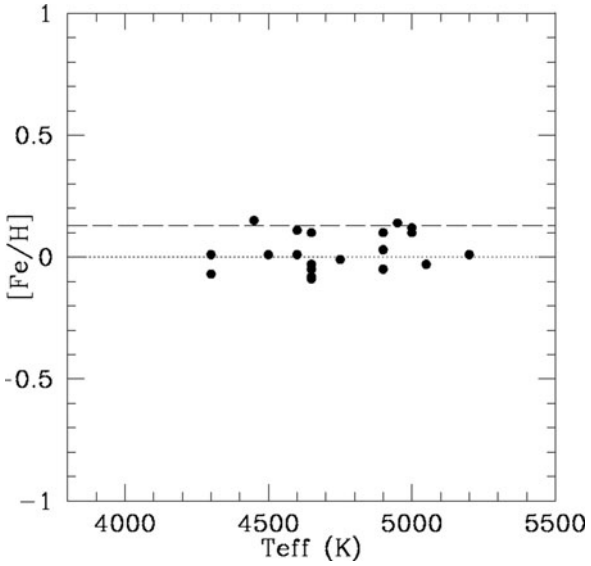


Fig. 1 [Fe/H] versus effective temperature T_{eff} for the putative members of the Hyades supercluster. The *dotted line* marks the solar abundance and the *dashed line* marks the Hyades open cluster metallicity at [Fe/H] = 0.13. Typical measurement uncertainty is 0.05 dex

4.3 The Argus Association

The Argus association was discovered in the SACY survey [34]. It is a sparse young stellar moving group with an age of 30 Myr located in the solar neighborhood. It has been suggested that the open cluster IC 2391 is related to this moving group [30]. The Argus association is well defined in UVW velocity space with U = −22 and V = −14 which overlaps directly with open cluster IC 2391 [35]. The Argus association is not the same as the IC2391 supercluster mentioned by Eggen [20]. Many of Eggen's supercluster members are early-type stars and are therefore not included in the SACY database.

Selected members of the Argus association and IC 2391 were observed using UVES at Paranal Observatory. The resulting high-resolution and high S/N spectra were analyzed and preliminary abundances of several elements are plotted in Fig. 2. The filled circles are IC 2391 stars and the crosses are the unbound association stars. The measurement error is on average 0.1 dex. As seen in Fig. 2, the cluster stars and unbound members have matching abundances within the current measurement uncertainty. This is indeed a positive sign indicating that the two structures share a common origin and that the unbound association stars are the dispersing members of the cluster, where originally they were part of the bound cluster system.

5 Conclusions

We have reviewed the abundance patterns of bound and unbound substructures within the Galactic disk. The internal homogeneity observed within open clusters, and the various abundance trends seen across the open cluster population, plus the fact that dispersed moving groups such as the HR1614 moving group is chemically identifiable, make a solid step forward for the viability of large-scale chemical tagging in order to identify the fossil stellar aggregates that led to the Galactic disk.

Recent and on-going high-resolution abundance studies are aiming to chemically identify other dispersed stellar aggregates of the disk which still retain some form of kinematical identity. The Hyades supercluster study find several stars with abundances matching that of the Hyades open cluster, although much of the sample stars are likely to be field stars which are also co-moving with the Hyades supercluster. The reality of the Hyades supercluster is still unclear and further extended samples of stars are needed to verify its identity. However, more positive results are detected for the young nearby Argus association. The abundances of the unbound association stars match the abundances of the open cluster IC2391 within the measurement uncertainty. This abundance match, together with their common kinematics, is a direct indication that the two structures have a common origin, and that the unbound association stars were originally part of the open cluster system.

The ultimate goal of large-scale chemical tagging is to reconstruct stellar aggregates that no longer retain any kinematical identify. High-resolution multi-object spectrographs such as HERMES (Freeman, this issue) on the AAT will be capable

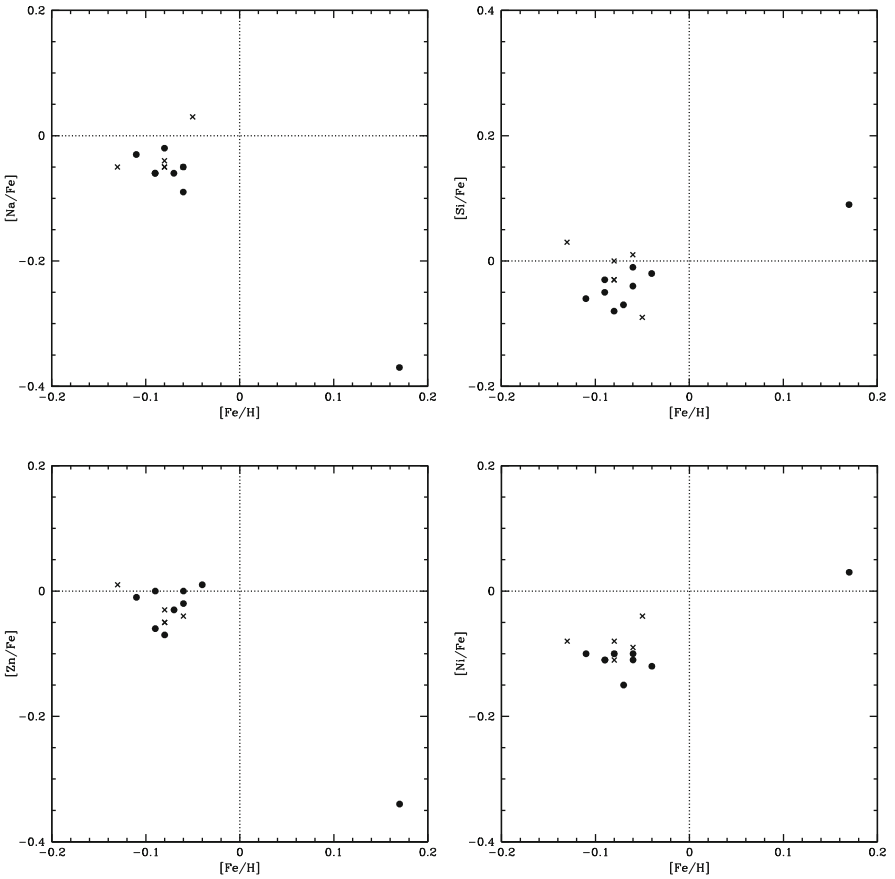


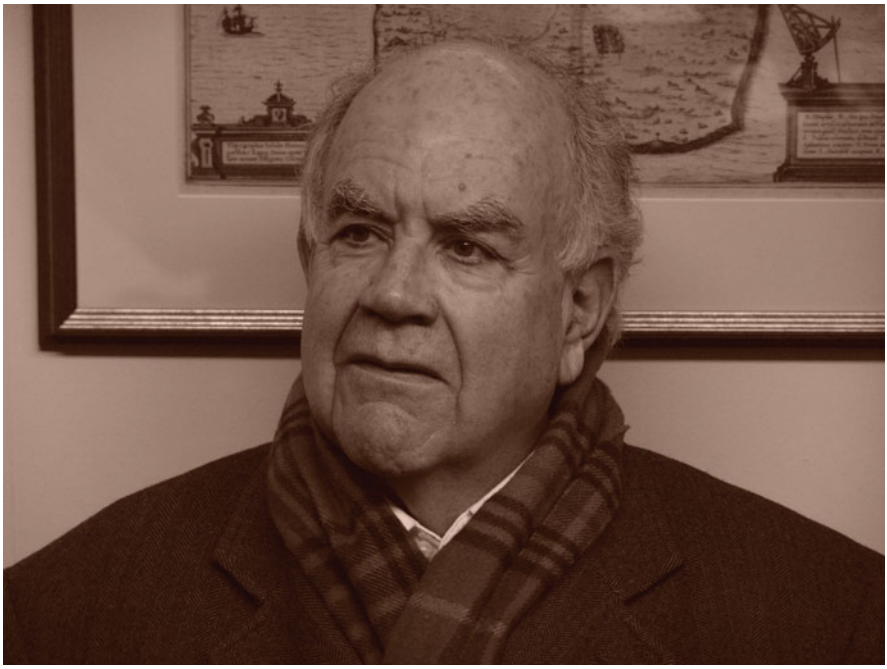
Fig. 2 Elemental abundances of the stellar members of open cluster IC 2391 (*filled circles*) and Argus association (*crosses*). The *dotted lines* mark the solar abundance levels. The estimated measurement uncertainty is 0.1 dex. One cluster star deviates greatly from the rest of the sample. Although the cause is yet unknown, it is most likely due to an error in the abundance analysis

of obtaining the data needed for large-scale chemical tagging of our Galactic disk. With such large data samples, it will be possible to reconstruct the original building blocks of the Disk.

References

1. Bensby, T., Feltzing, S., Lundström, I. 2003, A&A, 410, 527
2. Bobylev, V.V., Bajkova, A.T., Mylläri, A.A. 2010, AstL, 36, 27
3. Burbidge, E.M., Burbidge, G.R., Fowler, W.A., Hoyle, F. 1957, RvMP, 29, 547
4. Chen, Y.Q., Nissen, P.E., Zhao, G., Zhang, H.W., Benoni, T. 2000, A&AS, 141, 491
5. Dehnen, W. 1998, AJ, 115, 2384
6. Dehnen, W. 1999, ApJ, 524, L35

7. De Silva, G.M., Sneden, C., Paulson, D.B., Asplund, M., Bland-Hawthorn, J., Bessel, M.S., Freeman, K.C. 2006, AJ 131, 455
8. De Silva, G.M., Freeman, K.C., Bland-Hawthorn, J., Asplund, M., Bessel, M.S. 2007a, AJ 133, 694
9. De Silva, G.M., Freeman, K.C., Asplund, M., Bland-Hawthorn, J., Bessell, M.S., Collet, R. 2007b, AJ 133, 1161
10. De Silva, G.M., Freeman, K.C., Bland-Hawthorn, J. 2009, PASA 26, 11
11. Dias, W.S., Alessi, B.S., Moitinho, A., Lépine, J.R.D. 2002, A&A, 389, 871
12. Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D.L., Nissen, P.E., Tomkin, J. 1993, A&A, 275, 101
13. Eggen, O.J. 1970, PASP, 82, 99
14. Eggen, O.J. 1971, PASP, 83, 271
15. Eggen, O.J. 1974, PASP, 86, 162
16. Eggen, O.J. 1977, ApJ, 215, 812
17. Eggen, O.J. 1978a, ApJ, 222, 191
18. Eggen, O.J. 1978b, ApJ, 222, 203
19. Eggen, O.J. 1983, AJ, 88, 813
20. Eggen, O.J. 1991, AJ, 102, 2028
21. Eggen, O.J. 1992a, AJ, 104, 1482
22. Eggen, O.J. 1992b, AJ, 104, 1906
23. Eggen, O.J. 1996, AJ, 111, 1615
24. Eggen, O.J. 1998, AJ, 115, 2453
25. Famaey, B., Jorissen, A., Luri, X., Mayor, M., Udry, S., Dejonghe, H., Turon, C. 2005, A&A, 430, 165
26. Feltzing S., Holmberg, J. 2000, A&A 357, 153
27. Freeman, K., Bland-Hawthorn, J. 2002, ARAA, 40, 487
28. Friel, E.D. 1995, ARAA, 33, 381
29. Friel, E.D., Janes, K.A., Tavarez, M., Scott, J., Katsanis, R., Lotz, J., Hong, L., Miller, N. 2002, AJ, 124, 2693
30. Makarov, V.V., Urban, S. 2000, MNRAS, 317, 289
31. Pancino, E., Carrera, R., Rossetti, E., Gallart, C. 2010, A&A, 511, 56
32. Paulson, D.B., et al. 2003, AJ, 125, 3185
33. Reddy, B.E., Tomkin, J., Lambert, D.L., Allende Prieto, C. 2003, MNRAS, 340, 304
34. Torres, C.A.O., Quast, G.R., da Silva, L., de La Reza, R., Melo, C.H.F., Sterzik, M. 2003, ASSL, 299, 83
35. Torres, C.A.O., Quast, G.R., da Silva, L., de La Reza, R., Melo, C.H.F., Sterzik, M. 2006, A&A, 460, 695
36. Twarog, B.A., et al. 1997, AJ, 114, 2556
37. Wallerstein, G., et al. 1997, RvMP, 69, 995
38. Yong, D., Carney, B.W., de Almeida, M.L.T. 2005, AJ, 130, 597



Searching for Structures and Streams in the Extended Solar Neighbourhood with RAVE

Mary Williams and The RAVE Collaboration

Abstract RAVE (RAdial Velocity Experiment) is an ambitious program to conduct a 25,000 deg^2 survey to measure radial velocities, metallicities and abundance ratios of up to one million stars. Among many possibilities, RAVE provides means to study the remnants of past accretion events, local resonances and the general velocity distribution of the extended solar neighbourhood. RAVE data can be used either on its own or in combination with additional data. We detail one application of the first approach, showing the detection of a new stream. Two applications of the second approach are then shown with red clump stars in RAVE, where we examine the Hercules moving group and the tilt of the velocity ellipsoid in the extended solar neighbourhood.

1 Introduction

Kinematic surveys of stars in the solar neighbourhood are ideally suited to detect substructures in the nearby regions of the galaxy as well as establish overall properties of the stellar kinematics. Radial velocity surveys are particularly useful as they obtain the missing motion component for many stars and high-accuracy results can be achieved for a large sample of stars.

A prime example of such surveys is RAVE (RAdial Velocity Experiment), an ambitious program to conduct a 25,000 deg^2 survey to measure radial velocities (RVs), stellar parameters, metallicities and abundance ratios of up to one million stars. RAVE utilizes the wide field (30 deg^2) multi-object spectrograph 6dF instrument on the 1.2-m UK Schmidt Telescope of the Anglo-Australian Observatory (AAO). The observations are in the Ca-triplet spectral region at 840–875 nm with an effective resolution of $R = 7,500$. Starting in April 2003, RAVE had collected more than 4,00,000 spectra at the end of 2009 selected between $9 < I < 13$.

M. Williams (✉)
Astrophysikalisches Institut Potsdam, Potsdam, Germany
e-mail: mary@aip.de

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_25,

RAVE's radial velocities are accurate to $1.8\,\mathrm{km\,s^{-1}}$ when compared to external measurements [13]. This precision means that as a stand-alone resource, the radial velocities (RVs) are the best way to investigate fine velocity substructure. In Section 2 we describe an application of using RVs to detect a new halo stream in the solar neighbourhood. The RVs can also be combined with additional data to derive the full 6D phase space information for the stars. This requires deriving photometric distances to the stars as parallaxes are not available for the majority of RAVE targets in this pre-GAIA era. In Section 3 we describe our distance derivation for red clump giants in the RAVE sample to derive their space velocities. Two applications of these results are then described: following the variation of the Hercules moving group's strength in the extended solar neighbourhood and the variation of the tilt of the velocity ellipsoid with distance from the Galactic plane.

2 Radial Velocities

RAVE is a magnitude-limited spectroscopic survey, thus avoiding intractable kinematic bias. Furthermore, RAVE's RVs are accurate to $\sim 1.8\,\mathrm{km\,s^{-1}}$ and so fine substructures are best detected using radial velocities alone: combining with proper motions and distances means a significant loss of accuracy. To detect such substructure with RVs, however, requires a comparison model to establish the statistical significance of any substructures detected. To this end we utilize the Besançon Galaxy model [11] to create a mock sample of RAVE stars. A detailed analysis of comparing the Besançon Galaxy model to RAVE will be presented in an upcoming paper by A. Ritter; however, we concentrate here on the comparison in a region where we have identified a new stream.

We utilize the internal release of RAVE from January 2010 that contains 3,32,747 RVs of 2,52,790 individual stars. We use only those observations for which the $SNR > 13$ and the Tonry and Davis cross-correlation coefficient, $R > 5$ to remove potentially erroneous observations. For multiple observations of single stars the RVs were averaged as were the stellar parameters when available.

The mock sample created from the Besançon Galaxy model is made to mimic the I-band distribution for RAVE stars in $l \times b = 50^\circ \times 20^\circ$ regions and the number of stars in $l \times b = 25^\circ \times 10^\circ$ sub-regions. We further cut the RAVE and mock sample to $-70 < b < -50$, $10.2 < J$. The cut in Galactic latitude is made to zero-in on the region of interest. The magnitude cut reflects that the structure is in the more distant giants in the sample and particularly removing foreground giants enhances its visibility.

The observed structure was identified in the Galactic longitude-RV plane for $-70 < b < -50$, $10.2 < J$. To compare the Besançon sample to the RAVE sample we establish cells of size $\Delta l \times \Delta V_{\mathrm{rad}}$ and compare for each cell the number of stars from RAVE and the mock Besançon sample. Following a procedure similar to [8] we identify the over-dense region by varying the cell sizes with longitude slices ranging from $\Delta l = 25,\ 35,\ 50,\ 70$ and radial velocity bins

ranging from $\Delta V_{\rm rad} = 20,\ 25,\ 35, \ldots, 100$. The region around $35^\circ < l < 75^\circ$, $-250 < V_{\rm rad} < -150\,{\rm km\,s^{-1}}$ is consistently identified as being significant: for 100% of the cells it is identified as 3σ significant and 63% have a 6σ detection where Fig. 1 displays the 6σ detections in the l-RV plane. The maximum signal is 10σ for the $50^\circ < l < 75^\circ$, $-260 < V_{\rm rad} < -180\,{\rm km\,s^{-1}}$ cell. We thus conclude that the over-density is statistically significant. Adjacent latitude cuts of $-50^\circ < b < -30^\circ$ and $-90^\circ < b < -70^\circ$ do not exhibit an overdensity: the substructure is highly localized. As it lies in the direction of the constellation of Aquarius we name it the Aquarius stream. The stream does not appear to be associated with any previously known structures, such as the Sagittarius dwarf galaxy debris, but rather is a new stream in the nearby halo of the galaxy.

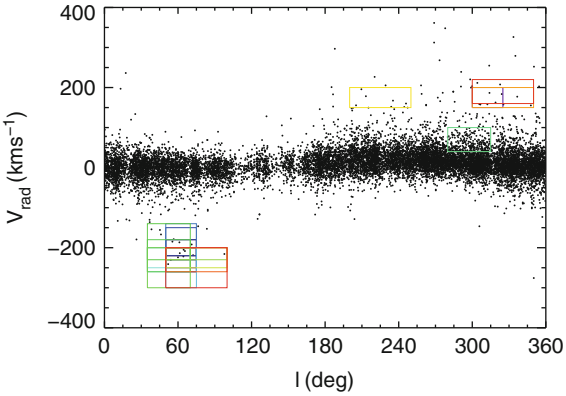


Fig. 1 The RAVE data in the range $-70 < b < -50$, $J > 10.2$. The new Aquarius stream is identified multiple times as a 6σ over-density for different cell sizes in the region $V_{\rm rad} < -150\,{\rm km\,s^{-1}}$, $30^\circ < l < 110^\circ$, where the 6σ detections are marked by the boxes

3 The Red Clump

The He-burning red clump forms a sizeable (20%) population of stars in the RAVE data set. They can be selected using the $\log g$ and $T_{\rm eff}$ plane, derived from the RAVE spectra. The red clump can be used as a reasonable standard candle, and hence can provide photometric distances for stars in the RAVE data covering a large range of distances from the Sun. We can thus use this range in distance to examine features of the overall population, or large sub-populations, over the extended neighbourhood of our Sun.

We use the calibration of [1] for $M_K({\rm RC}) = -1.61$. To minimize contamination by first-ascent giants we use the limits $0.4 < H - K < 0.65$, $1.5 < \log g < 2.7$ to select 48,000 red clump giants from the RAVE data set (as above, averaged for repeats and quality controlled). Combining the corresponding distances with RAVE radial velocities and proper motions, sourced by RAVE from Starnet2, Tycho2 and UCAC2, we calculated the space velocities $V_R,\ V_\phi,\ V_Z$ for these stars.

We use the Dehnen [5] values for the solar peculiar velocity of $(U,\ V,\ W) = 10,\ 5.3,\ 7.2\,\mathrm{km\,s^{-1}}$ with respect to the LSR, which we set at a rotation velocity of $220\,\mathrm{km\,s^{-1}}$. Note that in this analysis we use cylindrical coordinates rather than the cartesian UVW system as we are examining motions at large distances from our Sun. The magnitude limits of the RAVE survey translate into a distance range of $500 < d < 3{,}500$ pc. However, at larger distances transverse velocity errors become prohibitively large and so we limit our analysis here to $d < 1{,}500$ pc.

Errors in the space velocities, and any other feature, can be propagated through using a Monte Carlo procedure. This method consists of generating 1,000 random realizations each for the errors in $\mu_\alpha,\ \mu_\delta$, radial velocity and distance, assuming a Gaussian distribution for each in turn. The space velocities are then calculated for these realizations and standard deviations computed as well as the covariance terms. We exclude those stars with $\sigma_{V_\phi},\ \sigma_{V_R},\ \sigma_{V_Z} < 50\ \mathrm{km\,s^{-1}}$ and $|\mathrm{Cov}(V_\phi,\ V_R)|,\ |\mathrm{Cov}(V_\phi,\ V_Z)|,\ |\mathrm{Cov}(V_R,\ V_Z)| < 1{,}000\ (\mathrm{km\,s^{-1}})^2$ to clean our sample of those stars with the largest errors and covariances. Following [12] we also trim our sample of contaminating dwarfs whose distances are likely overestimated by using the cuts $\sqrt{V_R^2 + (V_\phi - 220)^2 + V_Z^2} < 200\ \mathrm{km\,s^{-1}}$ and $(V_\phi - 220) < 100\ \mathrm{km\,s^{-1}}$.

3.1 The Hercules Moving Group

The Hercules moving group, also known as the U-anomaly, is an over-density of stars at positive V_R with $150 < V_\phi < 190\ \mathrm{km\,s^{-1}}$. It has been shown to most probably arise as a result of a resonance caused by the influence of the Galactic bar [3, 4, 7, 10]. There is, however, some debate over the mechanisms involved. In one scenario, the group is a result of the solar position being near the Outer Lindblad Resonance (OLR) with the bar [4]. This resonance is for orbits whose epicyclic frequency, κ, is related to the bar's rotation frequency, Ω_b, and the circular orbit frequency, $\Omega(R)$, by $(\Omega(R) - \Omega_b) = -\kappa/2$. Thus, at the OLR stars see the end of the bar swinging by at the top of each epicycle. Kalnajs [9] showed that the placement of the Sun near the OLR gives a splitting of the velocity distribution into two modes, which he interpreted as giving rise to the Hyades and Sirius streams. In Dehnen's interpretation, the large mode is the LSR mode around $(U, V) = (0, 0)$, while the second mode is the Hercules group, a view that is supported by [3]. However, an alternative view is that the stream consists of stars on "hot", chaotic orbits that have come from within corotation [7, 10]. Bensby et al. [2] showed that the stars within the Hercules group have abundances and ages consistent with being a typical mixture of thin and thick disk stars, which supports the resonance explanation for the group. To date, however, the group has only been investigated in the immediate solar neighbourhood. With the RAVE red clump giants we examine the variation of the group's strength with radius, R, and height above the plane, Z, over the extended solar neighbourhood. In doing so we provide new constraints for models of the group's origin.

We measure the Hercules over-density via a symmetry argument in that Hercules is seen in positive V_R in a certain V_ϕ range. We begin by computing the fraction of stars between $150 < V_\phi < 190\,\mathrm{km\,s^{-1}}$ over all stars with positive V_R, i.e.

$$f_{+\mathrm{V_R}} = \frac{N_{150<V_\phi<190,\ 0<V_R}}{N_{0<V_R}}. \tag{1}$$

The fraction in the Hercules V_ϕ velocity range for negative V_R is likewise computed and denoted as $f_{-\mathrm{V_R}}$. We then estimate the strength of the Hercules group signal as the ratio of these two values:

$$S_{\mathrm{Hercules}} = \frac{f_{+\mathrm{V_R}}}{f_{-\mathrm{V_R}}} \tag{2}$$

This value is independent of the number of stars in the sample and is also stabler than the direct fraction $N_{150<V_\phi<190,\ 0<V_R}/N_{150<V_\phi<190,\ V_R<0}$. It can be seen as a convolution of the number of stars in the Hercules group with any movement of the group's position in V_ϕ. When there are more stars in the region around Hercules compared to the other side $S_{\mathrm{Hercules}} > 1$. If the two sides are equal and Hercules is thus not present $S_{\mathrm{Hercules}} = 1$.

We examine the variation of S_{Hercules} with planar distance from the Galactic centre, R, and height above the Galactic plane, Z, by cutting up the disk into rings $\Delta R = 250$ pc thick in radius and $\Delta Z = 500$ pc thick in height. We then calculate S_{Hercules} for each bin where, to minimize Poisson noise, we only use those slices with the number of stars in the box $N > 1,000$.

Figure 2 displays the variation of the Hercules over-density as a function of the middle values for each slice of R and Z. The over-density peaks just below the solar circle at $R = 7,750$ pc for both the $Z = -250$ pc and the $Z = +250$ pc bins. Furthermore, the over-density is less marked on average for the $Z = -750$ pc slice, further away from the disk. Both of these features are consistent with a resonance within the Galactic disk: Hercules is strongest closer to the plane than further away and as you move through radius the resonance emerges and disappears. Additionally, the kinematic modelling papers of Dehnen [4] and Fux [7] showed that the phase space structure associated with the Hercules group depends strongly on radius and to a lesser extent bar angle, enabling these studies to constrain the OLR location and bar angle. Our results could therefore be a further constraint particularly on the location of the OLR and thus the pattern speed of the bar. Interestingly, the $Z = -750$ pc slice also tends to peak at a lower R than the $Z = -250,\ +250$ pc slices (unfortunately there are insufficient stars in most $Z = +750$ pc slices). This could further indicate a dependence on the shape of the Galactic and bar potentials.

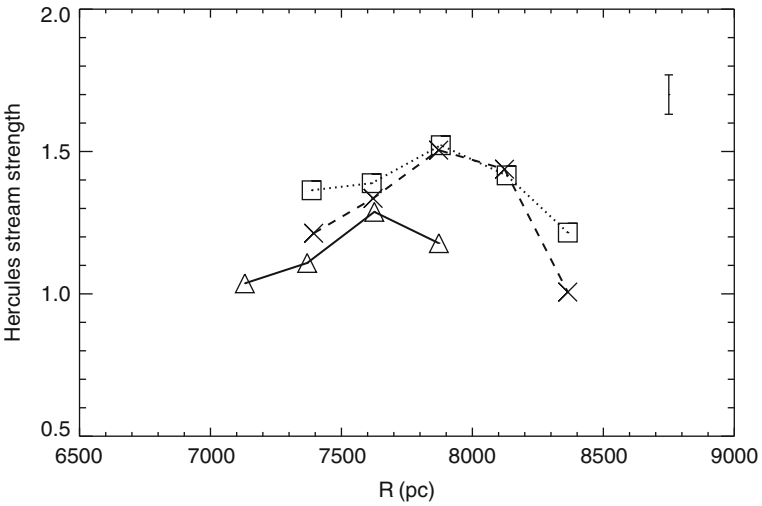


Fig. 2 Variation of the Hercules over-density as a function of the middle R and Z for each bin. The heights above the plane are given as $Z = -750$ pc (*solid line*), $Z = -250$ pc (*dotted line*) and $Z = +250$ pc (*dashed line*). The over-density peaks just below the solar circle and also is strongest closer towards the Galactic plane. This is the first measurement of the variation of the Hercules group's strength away from the immediate solar neighbourhood and the variations could be seen to support a resonant origin for the group

3.2 Tilt of the Velocity Ellipsoid

The tilt of the velocity ellipsoid is related to the mass distribution and so to the potential of the galaxy and its measurement beyond the immediate solar neighbourhood constrains this fundamental property. The RAVE red clump stars can be utilized to examine the tilt of the velocity ellipsoid as a function of distance from the Galactic midplane. Previous measurements using red clump stars in RAVE by Siebert et al. [12] found a tilt at 1 kpc above the Galactic plane of $7.3° \pm 1.8°$, consistent with a short scale length of the stellar disc ($R_d \sim 2$ kpc) if the dark halo is oblate or with a long scale length ($R_d \sim 3$ kpc) if the dark halo is prolate. Fuchs et al. [6] investigated the kinematics of two million M-dwarfs from the Sloan Digital Sky Survey (SDSS) in a kiloparsec-sized region around the Sun. Their results gave a strong trend with Z for the velocity ellipsoid tilt of $\alpha = -20°$ at $Z = 750$ pc. Here we repeat the measurement of Siebert et al. [12] but this time mimicking the selection criteria of the Fuchs results for comparative purposes.

We apply the limits on variances and covariances as above, but this time we divide our sample into slices of $\Delta Z = 100$ pc to follow Fuchs analysis. Each slice is only retained if it has a number of stars $N > 800$, once again to minimize the effects of Poisson noise. Following the distribution of the SDSS M-dwarfs we limit our sample to $7,000 < R < 9,000$ pc. The velocity ellipsoid tilt is then calculated via the relation

$$\alpha_{\rm tilt} = -\frac{1}{2}\arctan\frac{2\sigma^2_{V_R V_Z}}{\sigma^2_{V_R} - \sigma^2_{V_Z}}. \quad (3)$$

Figure 3 shows the tilt as a function of distance from the midplane $Z = 0$. The tilt changes below and above the Galactic plane in a symmetrical fashion with the velocity ellipsoid tilting towards the centre at larger values of $|Z|$. The values of the tilt are much more modest value than Fuchs values, with a tilt of only $\alpha_{\rm tilt} \sim -7°$ at $Z = +750$ pc. While the different studies examine different populations of stars the underlying potential is expected to be the same and so this discrepancy requires further investigation. Reassuringly, this result agrees with that found by Siebert et al. [12]. Also plotted are the relations expected from a spherical halo for the radii $R = 7,\ 8$, and 9 kpc. Comparing our values to these relations we see that near the plane ($|Z| < 300$ pc – which is of the order of the thin disk scale height) the tilt is minimal, as could be expected for a planar potential. Beyond this point the tilt becomes quite marked and is indeed sometimes larger than that expected from the spherical potential. Clearly further modelling is required to understand the implications of these results for the shape of the Galactic potential.

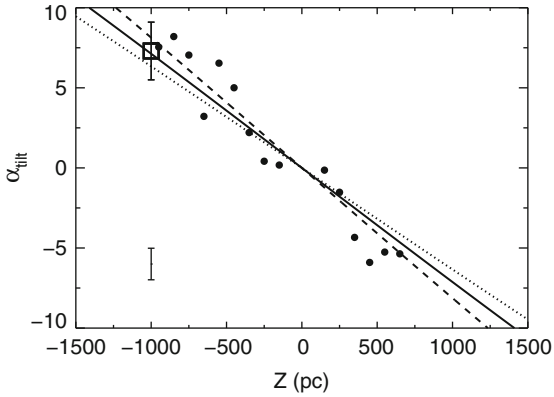


Fig. 3 Tilt of the velocity ellipsoid as a function of distance from the Galactic plane, Z. The *square point* is the result from Siebert et al. [12], while the *solid line* is the tilt expected from a spherical halo potential, with $\alpha_{\rm tilt} = \tan(Z/8,000\ {\rm pc})$, where 8,000 pc is the middle of the range of R for the stars. Also plotted are $\alpha_{\rm tilt} = \tan(Z/7,000\ {\rm pc})$ (*dashed*) and $\alpha_{\rm tilt} = \tan(Z/9,000\ {\rm pc})$ (*dotted*)

4 Conclusion

RAVE will pass half a million stars within the next year. We have shown three applications of this substantial kinematic data set, one involving the primary data product – the radial velocities – alone and two using supplementary data to derive the full 6D phase space information for red clump stars. These results show how the RAVE data set can be used to unmask hitherto unknown streams and to study the

properties of known kinematic structures as well as the underlying potential of the extended solar neighbourhood.

References

1. Alves, D.R., et al. 2000, ApJ, 539, L732
2. Bensby, T., Oey, M.S., Feltzing, S., Gustafsson, B. 2007, ApJ, 655, L89
3. Chakrabarty, D. 2007, A&A, 467, 145
4. Dehnen, W. 1999, ApJ, 524, L35
5. Dehnen, W., Binney, J.J. 1998, MNRAS, 298, 387
6. Fuchs, B., et al. 2009, AJ, 137, 4149
7. Fux, R. 2001, A&A, 373, 511
8. Helmi, A., White, S.D.M., de Zeeuw, P.T., Zhao, H. 1999, Nature, 402, 53
9. Kalnajs, A.J. 1991, In: Dynamics of Disc Galaxies, ed. B. Sundelius, Göteborgs University, Gothenburg, p. 323
10. Raboud, D., Grenon, M., Martinet, L., Fux, R., Udry, S. 1998, A&A, 335, L61
11. Robin, A.C., Reylé, C., Derrière, S., Picaud, S. 2003, A&A, 409, 523
12. Siebert, A., et al. 2008, MNRAS, 391, 793
13. Zwitter, T., et al. 2008, AJ, 136, 421

On the Age–Metallicity–Velocity Relation in the Nearby Disk Using the RAVE Survey

Borja Anguiano, Kenneth C. Freeman, Matthias Steinmetz, and Elizabeth Wylie de Boer

Abstract The age–metallicity–velocity relation (AMVR) in the nearby disk is a fundamental issue in our understanding of the evolution of the Milky Way. However, there are still major differences between various versions of this relation (e.g., Twarog 1980, ApJ, 242, 242, Meusinger et al. 1991, A&A, 245, 57, Edvardsson et al. 1993, A&A, 275, 101, Rocha-Pinto et al. 2000, A &A, 358, 850, Quillen & Garnett 2001, Galaxy Disks and Disk Galaxies, ASP, San Francisco, CA, vol. 230, p. 87, Feltzing et al. 2001, A&A, 377, 911, and Holmberg et al. 2007, A&A, 475, 519). Also, we have found considerable differences between different methods for age-dating stellar populations and have shown how different methods introduce bias in the AMVR. We are therefore working in a new derivation of these possible relations in the nearby disk, using a carefully selected sample of subgiants stars from RAVE and Geneva–Copenhagen surveys. RAVE provides accurate kinematics and a first estimate of metallicity, temperature, and gravity. Follow-up observations allow us to obtain accurate fundamental parameters via spectrophotometry and high-resolution spectroscopy in order to derive reliable ages and generate an AMVR for stars of the Galactic disk.

1 The Age–Metallicity and Age–Velocity Relation in the Galactic disk

The progressive chemical enrichment of the interstellar medium suggests that the young generation of stars is more enriched in heavy elements. According to this simple chemical picture we may expect to find an age–metallicity relation (AMR) in which the old stars in the disk are metal poor and the young population metal rich. From observational evidence it has been known for a long time that old, metal-rich stars permeate the Galaxy [5]. Moreover the shape of the AMR is still controversial, the metallicity dispersion changes from 0.12 [11, 16, 20] to up to 0.20 dex [4]

B. Anguiano (✉)
Astrophysikalisches Institut Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany
e-mail: baj@aip.de

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_26,

suggesting there is no AMR for field stars. Feltzing et al. [6] found an AMR only for stars younger than 2 Gyr and recently the Geneva–Copenhagen survey has found an AMR with a metallicity scatter of 0.20 dex [9]. For the cluster population in the Galactic disk the old open clusters are as metal rich as the young ones [8].

The kinematical evidences about the evolution of the disk are also controversial. While Freeman [7], Edvardsson et al. [4], Quillen and Garnett [14], and Soubiran et al. [17] indicate that the disk heating occurs for the first 2 Gyr and then *saturates* when $\sigma_z \sim 20$ km/s, Wielen [21], Dehnen and Binney [3], Nördstrom et al. [13], and Holmberg et al. [9] show a *steady increase* in the velocity dispersion with age. Apart from the observational disagreement the heating mechanisms, e.g., transient spiral arms, accretion of satellites, giant molecular clouds scattering are not yet well understood.

2 The Age–Metallicity–Velocity Relation Project

The new generation of spectroscopy surveys brings new opportunities to understand the formation and evolution of the Galaxy through their stellar components. One of the most important surveys in our days is the RAdial Velocity Experiment (RAVE), the survey work in the Ca-triplet region and provides accurate radial velocities, cross-matched proper motions, stellar atmosphere parameters, and distances for hundred of thousand stars [18, 22]. We are working in a new derivation of the AMR and AVR using a carefully selected sample of subgiants stars from RAVE and the Geneva–Copenhagen survey.

2.1 Sample Definition

Subgiants are suitable stars for dating the Galaxy as the isochrones separate well in the M_V/log g − log T_{eff} plane, the ages of the individual stars are probably the main source of error together with bias in the selection sample in the discrepancies commented in Section 1. From the RAVE survey we have selected around 2,000 subgiants candidates using a first estimate of stellar parameters and 404 no binaries subgiants from the Geneva–Copenhagen survey in the M_V–T_{eff} plane.

2.2 Follow-Up Observations and Data Analysis

Further observations are needed to improve the stellar parameters in order to clean the sample and get accurate ages. We have collected 1,253 low-resolution spectra from the double beam spectrograph (DBS) in the ANU 2.3-m telescope in Siding Spring Observatory (Australia). The resolving power is 400, with a spectral range from 3,150 to 6,400 Å, the dispersion is 1.96 Å/px, and the width of the slit 5″. All the observed stars present very high S/N (up to 150 at 5,500 Å). In order to get an

accurate flux calibration we have observed standard stars at different airmasses for every night.

Spectroscopy observations and data reduction techniques can provide accurate values for gravity (around 0.1 dex) through comparison of low-resolution spectra with recent grids of synthetic flux spectra [2]. We are currently working on the derivation of the stellar parameters via chi-squared statistic using synthetic model atmospheres [12]. This technique allows us to derive a widely applicable method of deriving stellar parameters from low-resolution spectrophotometry.

In Fig. 1 we present the results from our spectrophotometry using the DDO bands. We got plenty of subgiants with some contamination from dwarfs and giants. Preliminary results from the low-resolution spectra provide accurate values with errors in temperature around 100 K and 0.2 dex for metallicity and gravity.

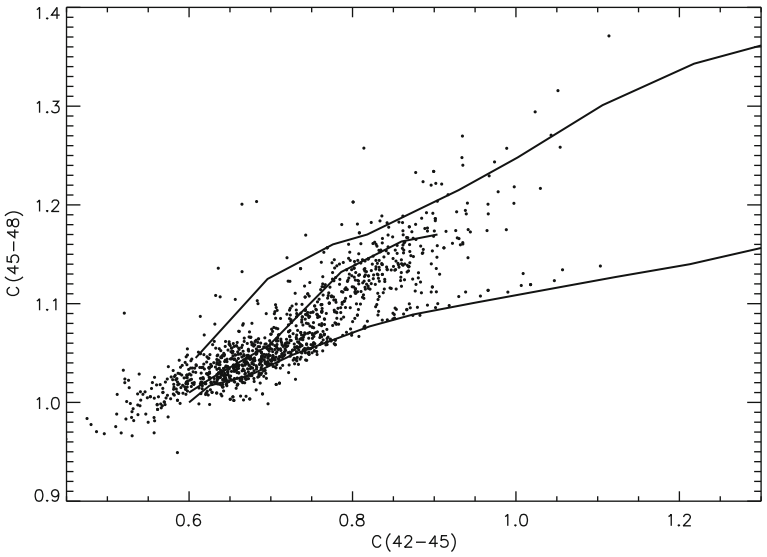


Fig. 1 The DDO indices C(42–45) versus C(45–48) derived from the spectrophotometry for the selected sample. The *lines* represent the luminosity class V, IV, and III (from *bottom* to *top*) from the mean values of the DDO indices for MK spectral classes listed in the catalogue of homogeneous photometry of bright stars from McClure and Forrester [10]

3 The Age of Stars

Stellar ages are crucial to place the observed chemical and kinematical properties in an evolutionary context. The most common method for deriving ages of stars is via isochrones, but not all stars can be dated in this way. Moreover we find remarkable differences between different authors and techniques (see Fig. 2, also [1]). Another way is via the chromospheric activity of the stars. Rocha-Pinto and

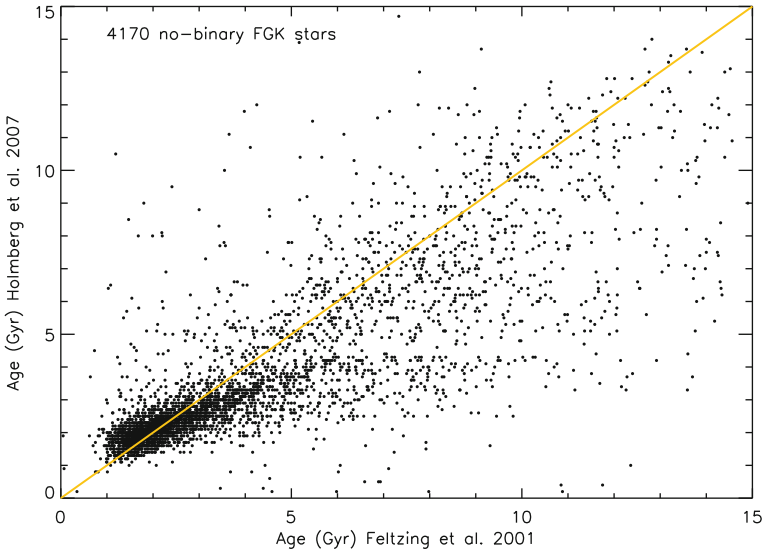


Fig. 2 Comparison between the ages of common stars derived in Feltzing et al. [6] and Holmberg et al. [9]. Both works use isochrone fitting but different grid of isochrones to derive the age of stars. There are clearly systematic errors in the age determination

Maciel [15] find that chromospheric ages tend to be lower than the isochrones age for metal-poor stars; however, a systematic study for intermediate-old stars is missing.

To derive the age of our subgiants we use the $M_V - T_{eff}$ plane for the Hipparcos stars and the log g–T_{eff} plane for the RAVE stars. Ten percent error in parallax or 0.1 dex in log g correspond to 2 Gyr in age uncertainty. For the older stars the error in age could be up to 3 Gyr [19].

4 The Velocity–Metallicity Relation in the Solar Neighborhood

In Fig. 3 we represent how the velocity dispersion changes with metallicity for the solar volume using the sample from Edvardsson et al. [4]. The velocity dispersion increases when we are getting more metal-poor stars. As the abundance decreases, we are probably getting more thick disk stars in the sample and the dispersion gradually increases.

The velocity dispersion slightly increases for the high-metallicity stars. These stars could be associated to the solar neighborhood moving groups, some of which have significant mean velocities relative to the LSR. That would contribute to the higher velocity dispersion.

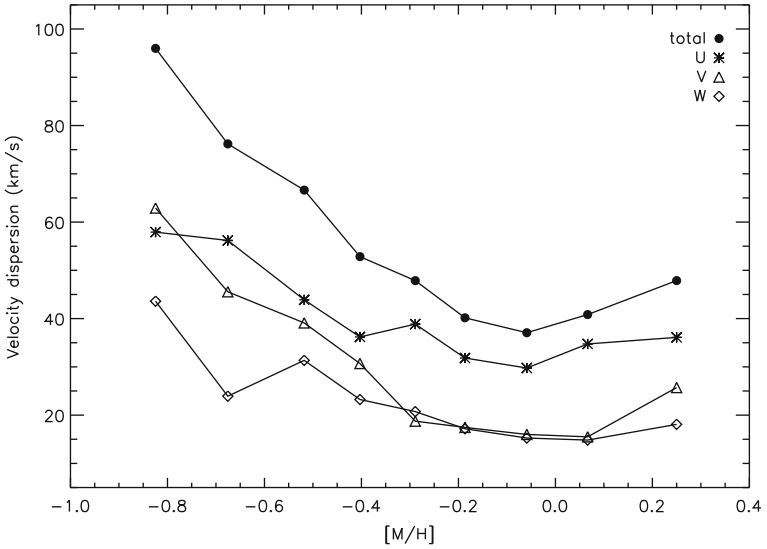


Fig. 3 The velocity dispersion versus metallicity derived from the Edvardssoon et al. [4] sample. Each point contains 22 stars. The typical error in the abundances is around 0.1 dex while for the space velocity components (U,V,W) is around 1.5 km/s in each component

References

1. Anguiano, B., Freeman, K., Steinmetz, M., Wylle de Boer, E., Siebert, A., The Rave Collaboration. 2009, IAU Symp., Cambridge University Press, Cambridge, 254, 3
2. Bessell, M.S. 2007, PASP, 119, 605
3. Dehnen, W., Binney, J.J. 1998, MNRAS, 298, 387
4. Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D.L., Nissen, P.E., Tomkin, J. 1993, A&A, 275, 101
5. Eggen, O.J., Sandage, A.R. 1969, ApJ, 158, 669
6. Feltzing, S., Holmberg, J., Hurley, J.R. 2001, A&A, 377, 911
7. Freeman, K.C. 1991, Dynamics of Disc Galaxies, Göteborg University, Göteborg, p. 15
8. Friel, E.D. 1995, ARAA, 33, 381
9. Holmberg, J., Nordström, B., Andersen, J. 2007, A&A, 475, 519
10. McClure, R.D., Forrester, W.T. 1981, Publ Dominion Astrophys Obs, Vic, 15, 439
11. Meusinger, H., Stecklum, B., Reimann, H.-G. 1991, A&A, 245, 57
12. Munari, U., Sordo, R., Castelli, F., Zwitter, T. 2005, A&A, 442, 1127
13. Nordström, B., et al. 2004, A&A, 418, 989
14. Quillen, A.C., Garnett, D.R. 2001, Galaxy Disks and Disk Galaxies, ASP, San Francisco, CA, vol. 230, p. 87
15. Rocha-Pinto, H.J., Maciel, W.J. 1998, MNRAS, 298, 332
16. Rocha-Pinto, H.J., Maciel, W.J., Scalo, J., Flynn, C. 2000, A&A, 358, 850
17. Soubiran, C., Bienaymé, O., Mishenina, T.V., Kovtyukh, V.V. 2008, A&A, 480, 91
18. Steinmetz, M., et al. 2006, AJ, 132, 1645
19. Thorén, P., Edvardsson, B., Gustafsson, B. 2004, A&A, 425, 187
20. Twarog, B.A. 1980, ApJ, 242, 242
21. Wielen, R. 1977, A&A, 60, 263
22. Zwitter, T., et al. 2008, AJ, 136, 421

The HERMES Project: Reconstructing Galaxy Formation

Kenneth C. Freeman

Abstract HERMES is a new high-resolution multi-object spectrometer on the Anglo Australian Telescope. The primary science driver for HERMES is Galactic archaeology. We are planning a spectroscopic survey of about a million stars, aimed at using chemical tagging techniques to reconstruct the star-forming aggregates that built up the disk and halo of the Galaxy. This project will benefit greatly from the stellar distances and transverse motions from the Gaia mission. We expect that HERMES will also be used for a wide range of other stellar and interstellar projects

1 Introduction

HERMES is a new high-resolution multi-object spectrometer on the Anglo Australian Telescope. It uses the existing 2dF fiber positioner with 400 fibers over an area of π deg^2 and has a spectral resolution of 27,000. (A higher resolution mode with a resolution of about 42,000 will also be available). Four VPH gratings and four cameras give a total wavelength coverage of about 1,000 Å in four non-contiguous wavelength bands chosen to include measurable lines of the elements to be used for chemical tagging. First light is expected in 2012.

The primary science driver for this multi-object high-resolution spectrometer is Galactic archaeology. We seek signatures or fossils from the epoch of Galaxy formation, to give us insight about the processes that took place as the Galaxy formed. The goal is to reconstruct the star-forming aggregates that built up the disk and halo of the Galaxy. Other HERMES science includes

- stellar physics
- the structure dynamics and chemical properties of the galactic bulge and disk

K.C. Freeman (✉)
Research School of Astronomy and Astrophysics, Mount Stromlo Observatory,
ANU, Canberra, Australia
e-mail: kcf@mso.anu.edu.au

D.L. Block et al. (eds.), *Galaxies and their Masks*,

DOI 10.1007/978-1-4419-7317-7_27,

- stellar populations in the Magellanic Clouds
- star clusters
- interstellar medium

The Galactic stellar science will benefit greatly from proper motion and parallax data from the Gaia mission which is expected to start generating data in the middle of the next decade.

2 Galactic Archaeology

The aim of the HERMES Galactic archaeology survey is to reconstruct the star-forming aggregates that built up the disk, bulge, and halo of the Galaxy. These aggregates have dispersed and mixed their debris around the Galaxy. Some of these dispersed aggregates can still be recognized kinematically as stellar moving groups. For others, the dynamical information was lost through disk heating processes, but their stars are still recognizable by their common chemical signatures via chemical tagging.

A major goal of Galactic archaeology is to identify how important mergers and accretion events were in building up the Galactic disk and the bulge. Cold dark matter (CDM) simulations predict a high level of merger activity which appears to conflict with many observed properties of disk galaxies and it would be very valuable to have a direct observational assessment of the contribution of mergers and accretion of small objects as the Galaxy came together. We can try to find the debris of groups of stars, now dispersed, that were associated at birth, either because they were born together in a single Galactic star-forming event or because they came in to the Galaxy from a common accreted galaxy.

2.1 The Galactic Halo

As an example, we can consider the Galactic stellar halo. The halo comprises only about 2% of the stellar mass of the Galaxy, but may be relatively easy to reconstruct. The halo includes stars with abundances [Fe/H] < -1 and forms a slowly rotating component of the Galaxy, in contrast to the rapidly rotating disk. The halo is believed to have been built up at least partly by the accretion of small metal-poor satellite galaxies (see [6] for a recent review). Most of these halo-building events occurred long ago, but some are still ongoing, like the accretion of the Sgr dwarf galaxy. Halo events are likely to be the easiest to reconstruct dynamically, because they are minimally affected by dissipation and the long orbital periods allow dynamical structures to survive. Accreted objects leave long-lived kinematic substructure in the galactic halo. This substructure is usually too faint to see in configuration space but is visible in phase space or integral space (e.g., the space of energy and angular momentum: see [11]).

2.2 The Galactic Disk

We can extend this approach of reconstruction to other components of the Galaxy, like the disk and bulge, which contain most of the Galactic stellar baryons. The Galactic halo shows kinematical substructure, believed to be the remains of accreted objects which contributed to building up the halo. The Galactic disk also shows kinematical substructure, usually seen in the solar neighborhood as stellar moving groups. The stars of these moving groups are all around us and are recognized by their common space motions. Some of these moving groups are associated with dynamical resonances involving the Galactic bar or spiral structure, like the Hercules moving group [7]. Some are the debris of star-forming aggregates in the disk, which have dispersed into extended regions of the Galaxy; the HR1614 moving groups appear to be such a system (e.g., [8]). Others may be the debris of infalling objects; Navarro et al. [13] suggested that the Arcturus moving group is an example but see Williams et al. [19] for a different view.

The HR1614 group contains stars with an age of about 2 Gyr and an abundance [Fe/H] = +0.2. De Silva et al. [8] measured precise abundances for several chemical elements in HR1614 group stars and found no detectable spread in abundance. This group appears to be chemically homogeneous, as do several open clusters studied by de Silva and others, and appears to be an example of debris from a dispersed star-forming event. See also de Silva's chapter "Elemental Abundance Patterns of Disk Substructure; this volume.

On the other hand, many of the moving groups are likely to be associated with resonances, as for the Hercules groups. Bensby et al. [2] showed that the stars of the Hercules group cover a range of abundances and cannot be chemically distinguished from the field stars of the thin and thick disk. Such groups are dynamically interesting but are probably not relevant to the goals of Galactic archaeology.

Although the disk does show some kinematical substructure, and some of this substructure appears to be associated with the debris of star-forming events, a lot of dynamical information was lost in the dissipation that led to disk formation and the subsequent heating by spiral arms and giant molecular clouds. A further loss of dynamical information comes from the radial mixing associated with resonant interaction of stars with transient spiral structure (e.g., [15–17]). This radial mixing comes from the interaction of a transient spiral wave with stars that are near the corotation radius for the wave. This interaction changes the angular momentum of the stars. The effect is to move stars from one near-circular orbit to near-circular orbits at larger or smaller radii. As an aggregate disperses around the Galaxy, it seems likely that the effects of radial mixing will be different for debris at different azimuthal locations in the Galaxy, so some stars from the aggregate will be flipped to larger radii and some to smaller radii. We can therefore expect that many dispersed stellar aggregates will not be recognizable dynamically. In the context of this meeting, the loss of dynamical information acts like a mask which obscures our ability to view the debris of these dispersed aggregates.

However, we are not restricted to dynamical techniques for reconstructing dispersed stellar aggregates. Much fossil information is locked up in the detailed

distribution of chemical elements in stars, and we can use this chemical information to see behind the mask and identify stars that were born together.

2.3 The Galactic Thick Disk

Probing the origin of the Galactic thick disk will be a prime application of this chemical tagging technique. The Galactic thick disk is an old component, with a mass of about 10% of the thin disk, and is significantly more metal-poor than the thin disk. Most of its stars have [Fe/H] between about -0.5 and -1, with a long tail of more metal-poor stars extending down to about -2.2. The stars of the thick disk are enhanced in α-elements, indicating that the thick disk formed on a timescale less than about a Gyr.

Thick disks are very common in other disk galaxies and are probably ubiquitous. They appear to be old ($> 6 - 10$ Gyr) and moderately metal poor, as in our Galaxy. They are clearly a significant feature in the formation and evolution of disk galaxies, but their formation is not yet understood. Some possible formation routes include

- the heating of the early stellar disk by accretion events or minor mergers [14]
- the disruption of ancient merging or accreted galaxies to form an old disk-like structure [1]
- star formation associated with early major gaseous accretion events [5]
- the disruption of large star-forming aggregates, releasing their individual binding energies to form a more extended disk-like structure [12]

The debris from star-forming events in the thick disk will have phase mixed and probably radially mixed and is likely to be kinematically unrecognizable. On the other hand, the debris of ancient chemically homogeneous star-forming aggregates will maintain its chemical abundance distribution over many elements. Even if we cannot recognize its stars from their kinematics and spatial distributions, we can identify them from their common chemical signature via chemical tagging.

3 Chemical Tagging

This technique uses the detailed chemical abundances of stars (light elements, α-elements, iron-peak elements, *r*- and *s*-process elements) to tag or associate stars with common ancient star-forming aggregates with similar abundance patterns (e.g., [10]). The detailed abundance patterns reflect the chemical evolution of the gas from which the aggregate formed. Differences in the details of the local chemical evolution lead to scatter in the detailed abundance distributions of different star-forming sites, particularly at lower metallicities. Different supernovae (SN) provide different chemical yields, depending on their mass, metallicity, detonation details, and ejected mass, and at lower metallicities only a few SN may be involved in the enrichment process.

For chemical tagging to work, a few conditions are needed such as

- stars form in large aggregates: this is believed to be true from evidence in nearby galaxies and galaxies at high redshift [3]
- star-forming aggregates are chemically homogeneous
- aggregates have unique chemical signatures defined by several elements whose abundances do not vary in lockstep from one aggregate to another. We need sufficient spread in abundances from aggregate to aggregate so that chemical signatures can be distinguished by measurements with the achievable accuracy (about 0.05 dex differentially)

Testing the last two conditions were the goals of Gayandhi de Silva's work on abundances of stars in open clusters: they appear to be true. See de Silva et al. [9] for more discussion of chemical tagging.

Chemical studies of the old disk stars in the Galaxy can help to identify disk stars that are the debris of common dispersed star-forming aggregates and also those which came in from outside in disrupting satellites.

We can think of a chemical space of abundances of elements (O, Na, Mg, Al, Ca, Mn, Fe, Cu, Zr, Ba, Eu, for example), defined by about 35 measurable elements. Not all of these elements vary independently; the dimensionality of this space is probably between about 7 and 9. The stars of a chemically homogeneous star cluster occupy a very small region of this space. Most disk stars inhabit a restricted subregion of the chemical space. In the surviving dwarf satellite galaxies of the Milky Way, the chemical evolution was different from satellite to satellite and different from the Galactic disk, as seen by the distribution of stars in say the [Mg/Fe] vs [Fe/H] plane (e.g., [18]). Stars which came in from satellites will stand out from the disk stars in chemical space.

With this chemical tagging approach, we expect to identify the debris of ancient star-forming aggregates in the galactic disk. We also expect to be able to detect or put observational limits on the satellite accretion history of the galactic disk.

We should stress that chemical tagging is not just assigning stars chemically to a particular population like the thin disk, thick disk, or halo. It is intended to assign stars chemically to a common origin in substructure which has lost its phase space identity. Chemical tagging needs a high-resolution spectroscopic survey of about a million stars, homogeneously observed and analyzed. This is the prime science driver for HERMES.

4 Galactic Archaeology with HERMES

Imagine a large complete stellar survey down to a V-magnitude of 14; at this magnitude, the stellar density matches the fiber density at intermediate galactic latitude. A survey covering about half of the southern sky (10,000 square degrees, with 3,000 pointings) would give about 1.2×10^6 stars. At V = 14 and with a spectral resolution of 30,000, we expect a SNR of 100 per resolution element in a 60-min exposure.

With about 8 pointings per night, this survey would take about 400 clear bright nights.

Thin disk stars provide about 78% of the sample (58% dwarfs, 20% giants). About 17% of the stars come from the thick disk (10% dwarfs, 7% giants) and the remaining 5% from the Galactic halo (2% dwarfs, 3% giants). The old disk dwarfs can be observed out to a distance of about 1 kpc, the disk giants to about 5 kpc, and the halo giants to about 15 kpc.

We assume that the debris of all of the stellar formation sites are now azimuthally mixed, right around the Galaxy. About 9% of the Galactic thick disk stars and about 14% of the thin disk stars pass through the 1 kpc dwarf horizon, so that all of their formation sites are represented within our horizon. For the halo, the HERMES halo giants are visible out to about 15 kpc, so we sample a large fraction of the Galactic halo.

Simulations [4] show that a random sample of 1.2×10^6 stars with $V < 14$ would allow detection of about 20 thick disk dwarfs from each of about 4,500 star-formation sites and about 10 thin disk dwarfs from each of about 35,000 star formation sites. We note that a smaller survey means less stars from a similar number of sites.

Is it possible to detect about 35,000 different disk sites using chemical tagging techniques? We would need about seven independently varying chemical elements or element groups, each with five measurable abundance levels, to get enough independent cells (5^7) in chemical abundance space. Are there seven independently varying elements or element groups? The answer appears to be yes. Evidence from star clusters indicates that the following groups of elements would serve the purpose:

- light elements (O, Na, Al)
- Mg
- other α-elements (Ca, Si, Ti)
- Fe and Fe-peak elements
- light s-process elements (Zr, Y)
- heavy s-process elements (Ba, La)
- r-process elements (Eu)

Additional dimensionality comes from abundance variations of elements within some of these groups.

5 HERMES and Gaia

HERMES offers an exciting synergy with Gaia, which will provide precision astrometry for about a billion stars in the next decade. At the relatively bright magnitudes of the HERMES chemical tagging survey, Gaia is performing at its best. We can expect a parallax precision of about 10 μas and a proper motion precision of about 10 μas year^{-1}. This corresponds to a 1% distance error at a distance of 1 kpc and a 0.7 km s^{-1} transverse velocity error at a distance of 15 kpc. We can

expect accurate distances and transverse velocities distances for all stars in the HERMES survey sample. We can then precisely locate all of the survey stars in a color–(absolute magnitude) diagram, to provide an independent check that the stars in chemically tagged groups have common ages.

It is likely that Gaia will detect many apparent moving groups in velocity space. From the HERMES survey, we will be able to determine which of these phase space substructures are the debris of dispersed star-forming aggregates and which are just due to dynamical resonances (like the Hercules group discussed above).

6 Other HERMES Science

The ultimate goal of the Galactic archaeology program is unravelling the star-formation history of the thin and thick disk and halo via chemical tagging. In the shorter term, the data products will include

- distribution of about a million stars in (position, velocity, chemical) space and isochrone ages for about 2,00,000 stars
- distribution of [Fe/H], [α/Fe], and [X/Fe] for many elements in large samples of stars from each of the thin and thick disks and the halo
- detailed abundance gradients in each component of the Galaxy
- chemical and kinematical correlations in the inner and outer thin and thick disks.

We can also expect many other programs to be proposed for HERMES. Some of these will involve longer exposures, reaching down to V $\sim$ 16. Examples include

- abundances of large samples of stars in the outer disk of the Galaxy: measuring the galactic abundance gradient in young and old stars
- chemical evolution and kinematics in the galactic bulge, hopefully complemented by accurate astrometry from JASMINE
- globular clusters, open star clusters, and superclusters: membership, dynamics, internal age–metallicity correlations and detailed abundance distributions
- AGB stars in the Magellanic Clouds and the evolution of the *s*-process element abundances.

Acknowledgments The ideas described in this chapter have been developed in collaboration with Joss Bland-Hawthorn. We acknowledge support from ARC grant DP0988751. The HERMES spectrometer concept was developed by the AAO technical group, led by Sam Barden.

References

1. Abadi, M., et al. 2003, ApJ, 597, 21
2. Bensby, T., et al. 2007, ApJ, 655, L89
3. Bland-Hawthorn, J., et al. 2010, ApJ, 713, 166
4. Bland-Hawthorn, J., Freeman, K. 2004, PASA, 21, 110
5. Brook, C., et al. 2004, ApJ, 612, 894

6. Carollo, D., et al. 2010, ApJ, 712, 692
7. Dehnen, W. 1998, AJ, 115, 2384
8. De Silva, G., et al. 2007, AJ, 133, 694
9. De Silva, G., et al. 2009, PASA, 26, 11
10. Freeman, K., Bland-Hawthorn, J. 2002, ARAA, 40, 487
11. Helmi, A., White, S. 1999, MNRAS, 307, 495
12. Kroupa, P. 2002, MNRAS, 330, 707
13. Navarro, J., et al. 2004, ApJ, 601, L43
14. Quinn, P., Goodman, J. 1986, ApJ, 309, 472
15. Roskar, R., et al. 2008, ApJ, 675, L65
16. Schönrich, R., Binney, J. 2009, MNRAS, 396, 203
17. Sellwood, J., Binney, J. 2002, MNRAS, 336, 785
18. Venn, K., et al. 2004, AJ, 128, 1177
19. Williams, M., et al. 2010, in preparation

Stellar Halos: Unmasking a Galaxy's History

Amina Helmi

Abstract We present the first results from an ongoing study to characterize the substructure in stellar halos as predicted by semi-analytic galaxy formation models in the context of the Aquarius project. The model stellar halos depict a wealth of substructure: broad overdensities akin those observed around the Milky Way as well as very narrow faint streams. Substructures are distributed anisotropically on the sky, a property which should become apparent in the next generation of photometric surveys. The solar neighborhood presents a large amount of kinematic groups which are related to the various building blocks of the stellar halo. This implies that the merger history of the galaxy may be unmasked in the very near future with surveys such as Gaia and LSST.

1 Stellar Halos

Stellar halos may contain *the* clues to the formation histories of galaxies. They result (at least in part) from the mergers and accretion of smaller systems, are repositories of their debris, and hence are key to unravel assembly histories. In the case of the Milky Way, in the halo we find the oldest and most metal-poor stars, fossils whose atmospheres have retained imprints of the earliest generations of stars formed in the Universe.

Observational studies of stellar halos are very challenging. These components are exceedingly faint (with typical surface brightness lower than 26 mag/arcsec^2), making their detection in external systems very difficult (e.g., [19, 26]). In the case of the MW, halo stars are largely outnumbered by the more common disc and bulge stars, and may be isolated by their characteristic MSTO color, by their lower metallicity, or high proper motion (see for references, [9]).

A. Helmi (✉)
Kapteyn Astronomical Institute, University of Groningen, 9700 AV Groningen, The Netherlands
e-mail: ahelmi@astro.rug.nl

D.L. Block et al. (eds.), *Galaxies and their Masks*,

DOI 10.1007/978-1-4419-7317-7_28,

Recent studies have revealed a plethora of substructures in the Milky Way halo, many of which may be directly associated to accretion events. The textbook example is the Sagittarius dwarf, which was discovered in 1994 by Ibata et al. [12] and whose streams were found later in surveys such as the APM [13], 2MASS [15], and SDSS [14, 25]. However, the nature of some of the other known overdensities [3, 16, 23] is much less clear as these could perhaps also be due to global asymmetries in the Galaxy [20]. Studies of the resolved stellar populations and of the integrated light in the halos of external systems have begun to reveal extended streams, such as the Giant Stream in M31 [17]. Nonetheless, we are still far from having an unbiased census of the commonality of such features or events in Milky Way-type systems.

In this chapter I will present the first results from an ongoing project on the formation of stellar halos in a cosmological context using semi-analytic models of galaxy formation in the Aquarius project.

2 Methods

Cooper et al. ([5], hereafter C10) have coupled the cosmological simulations from the Aquarius project to the semi-analytic galaxy formation code GALFORM. The Aquarius simulations [22] model the formation of six dark matter halos of mass comparable to that of the Milky Way, i.e., 10^{12} $M_{\odot}$ [1, 2, 21, 24]. Having an ensemble of objects allows one to establish how different formation histories can affect the properties of the resulting object.

In parallel to the dynamical evolution available from the simulations, C10 model the physical processes affecting the baryons associated to the various dark matter halos that are present in the simulations. Very complex physical processes, such as gas cooling, star formation and feedback, are included and allow one to establish the characteristics of the stellar populations formed in the course of time, as well as their location and dynamics thanks to the coupling to the dark matter simulations.

In Fig. 1 we show the distribution of dark matter particles in six different snapshots for Aq-C. At each timestep the algorithm by C10 identifies halos which may host galaxies and follows their most bound particles as a function of time. In practise, C10 select at each timestep the 1% most-bound particles present in star forming halos and assign them the properties of the stellar populations formed in that timestep. When an object is accreted, tidal forces can lead to its full disruption, and its "stars" thus become field stars. In this way, a stellar halo may be built.

In this contribution we have resampled the particle distribution such that each particle corresponds to a red giant branch star (for more details we refer the reader to Helmi et al., in prep.).

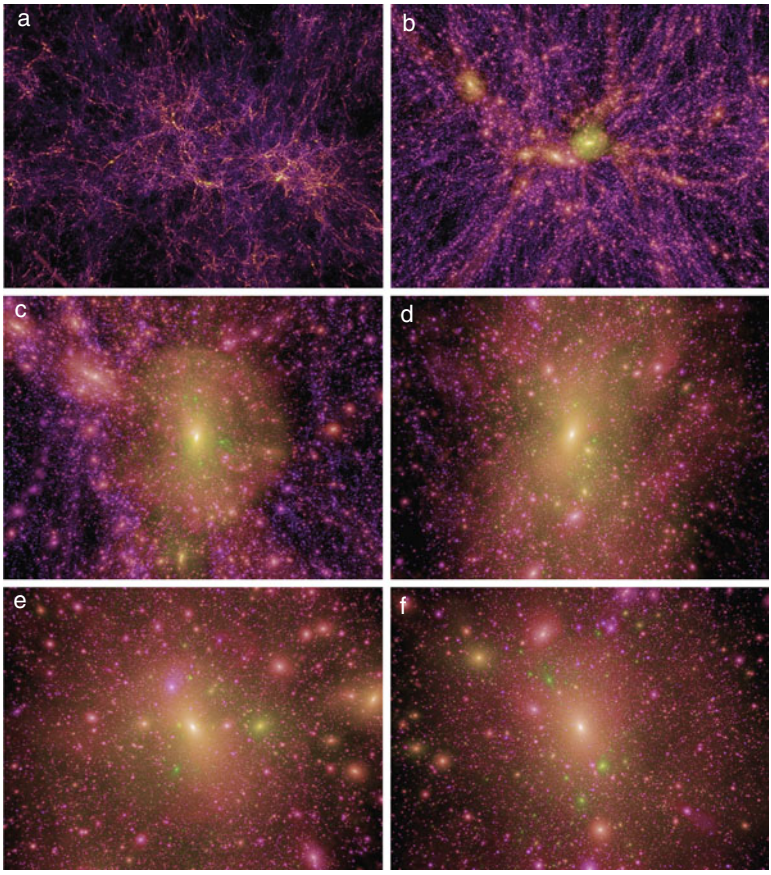


Fig. 1 The evolution of the dark matter distribution for Aquarius halo C (from [22]). The *top panels* correspond to simulations at redshifts greater than 6 in our early Universe, while the *central panels* correspond to intermediate redshifts ranging between 1 and 2. The two *bottom panels* show the results of simulations at redshifts closer to zero

3 Results I

3.1 Sky Distribution

Figure 2 shows the sky distribution of red giants (modeled as described above) for the Aquarius halo C. This stellar halo has a mass of $\sim 3.6 \times 10^8\,M_\odot$. We notice the presence of several very broad features, similar to some extent to the Hercules-Aquila clouds or to the Virgo overdensity, particularly in the distance range 10–30 kpc. Narrow streams are also present, they typically originate from small objects on elongated orbits with large apocenters.

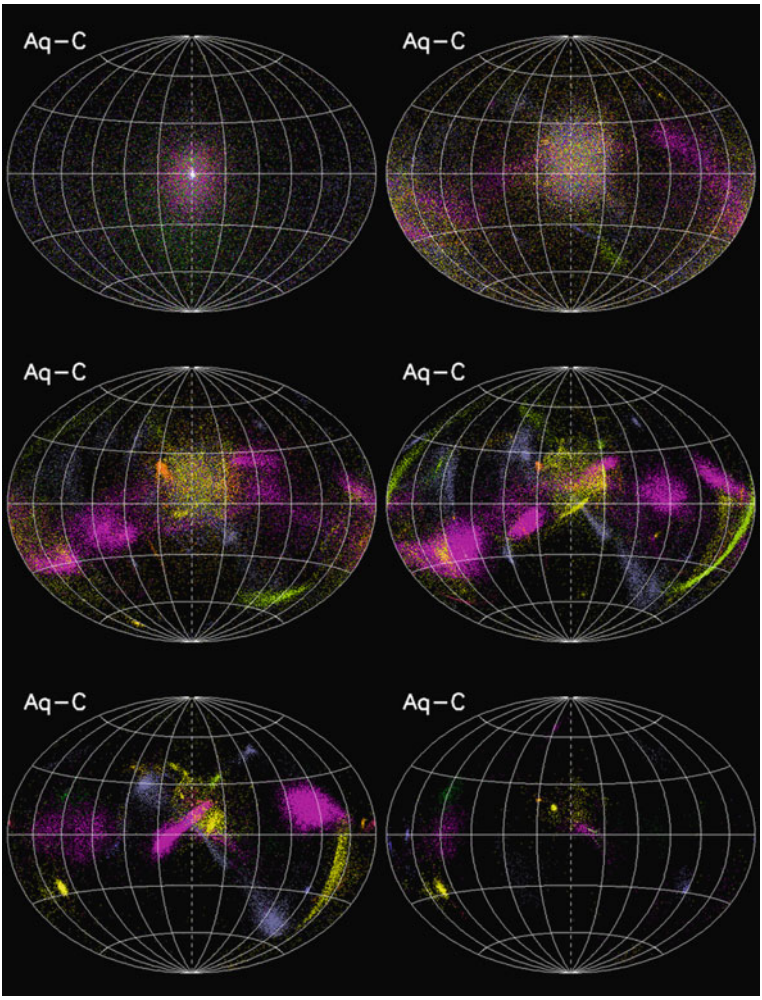


Fig. 2 Distribution of RGB stars on the sky for various distances from the "Sun" for the stellar halo of Aq-C: 0–10, 10–30, 30–50, 50–80, 80–120, 120–150 kpc (from *top left* to *bottom right*). The *different colors* correspond to stars originating in different progenitors

Also apparent in this figure is the somewhat anisotropic distribution of stars on the sky. This anisotropy is present at varying distances from the "Sun" and is due to streams from several objects. This suggests that the progenitors of the streams fell in from preferred directions.

3.2 Radial Velocities

Figure 3 shows the radial velocity from the Galactic center as a function of distance from the center for Aq-halo C. As in previous figures, different colors indicate dif-

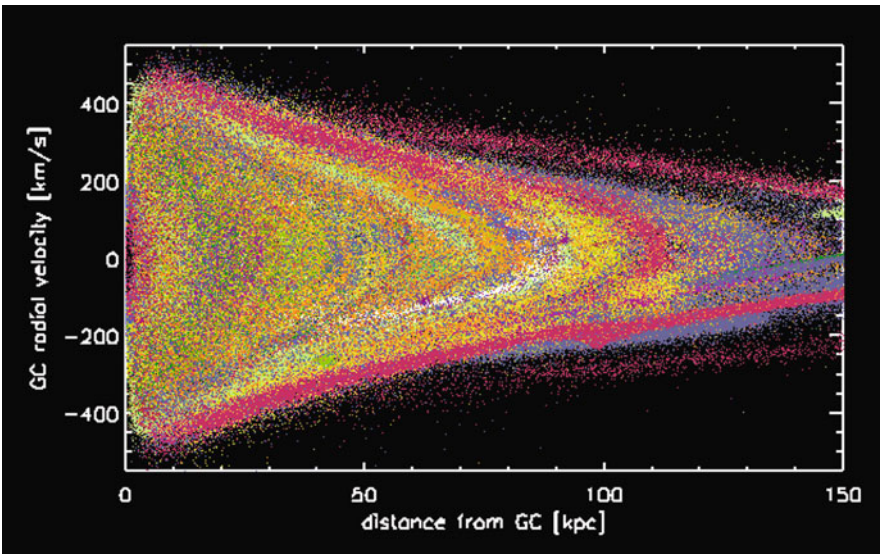


Fig. 3 Radial velocity as function of distance from the halo center for Aq-C. The *different colors* correspond to stars originating in different progenitors

ferent progenitors. We note that there is significant substructure in the kinematics, especially at distances greater than 10 kpc from the center. At small radii, streams are also apparent, particularly at very high velocities. These probably originate in the most recently accreted objects.

4 Results II. Solar Neighborhood

4.1 Kinematics

The inner regions of the halo are typically the first to form. Thus they contain unique imprints of the physical conditions present in the very early Universe. In the Aq-halos, the inner 10 kpc is composed by stars which have been accreted in objects at redshifts $z > 2.5 - 3$ (see C10, also [4, 6, 11]). Furthermore, the majority of these stars typically originate in a few progenitors.

Previous figures have shown a large amount of substructure at large radii. Mixing timescales are very short in the inner galaxy, which results in streams that are no longer coherent in configuration space. However, since the system is collisionless, phase space density should be conserved, which implies that substructure should still be present in the form of moving groups crisscrossing the solar neighborhood [10].

Figure 4 shows the velocity distribution of particles in a "solar neighborhood sphere." This volume is located at 8 kpc from the halo center and has a radius of 3 kpc. The particles shown here are those that have been "tagged" by our scheme,

but have not been resampled to represent RGB stars (although we expect a similar distribution for such stars). We note a great amount of substructure that remains in this volume, despite the very chaotic build up of the parent halo.

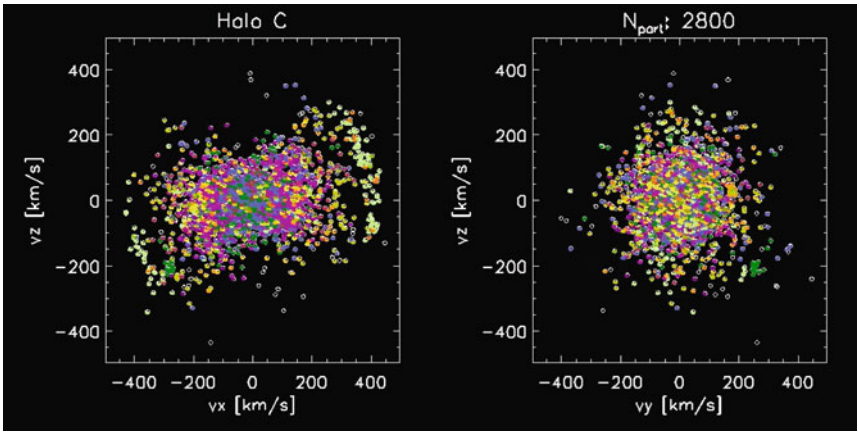


Fig. 4 Velocities of tagged dark matter particles in Aq-C located in a "solar neighborhood sphere." The *different colors* indicate different progenitors

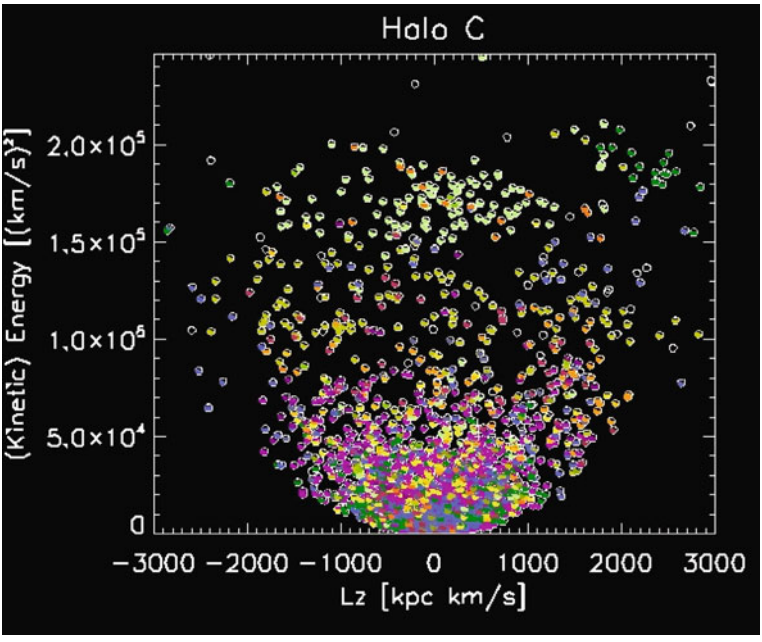


Fig. 5 A proxy for "Integrals of motion" space: Kinetic energy vs. z-angular momentum for the particles shown in Fig. 4

A possibly better way to visualize and recover the streams from the various building blocks of the stellar halo is to study the space of "integrals of motion." In our case, these are not known, and we simply plot the kinetic energy of our stars as a function of the z-component of their angular momentum. Note that kinetic energy is a proxy for the total energy, since the stars are located in a sufficiently small volume that one may, to zero order, assume the gravitational potential is nearly constant. None of the components of the angular momentum vector are conserved, since our dark halos are triaxial. Nonetheless, Fig. 5 shows very strong clumping, which is much more clearly visible than in velocity space. We see now that stars from the same building block are strongly clustered in this space. All these substructures should be visible with the Gaia mission.

5 How to Retrieve the History of the Stellar Halo

We have seen that stars from accreted objects are strongly clumped in "integrals of motion" space. A suitable clustering algorithm should be able to retrieve the various structures and hence this would allow us to establish the number as well as the characteristics of the progenitors of the stellar halo.

To establish when such objects were accreted (i.e., the history), we turn to the space of orbital frequencies. In this space, streams from the same object are distributed in a regular pattern whose characteristic scale is directly related to the time of accretion.

Gómez and Helmi ([7], see also [18]) have shown that when merger debris phase-mixes away, streams that cross a small volume in configuration space can be easily identified in frequency space. This is highlighted in Fig. 6, where the top panels show the evolution of the distribution of stars in configuration space from an accreted satellite in a static spherical potential. The bottom panels show the distribution in various spaces of a subset of these stars (those located inside the circle shown in the top right panel, i.e., a "solar neighborhood sphere"). The left middle panel corresponds to velocity space: here a broad swath of stars is clearly apparent (a "banana"-shape) which is composed by more than ten streams, which are barely distinguishable from one another. The right middle panel shows the often used "energy-L_z" space. Now the various streams become more apparent as small elongated structures at nearly constant energy. The bottom panel shows the space of orbital frequencies and demonstrates a clear and very regular distribution of streams. The orbital frequencies have been computed using the (known) underlying gravitational potential, but even if this is known only approximately, the regularity of the pattern remains. This is likely due to the fact that the streams are localized in a small volume in space (i.e., a solar neighborhood sphere), in which the gravitational potential is roughly constant.

The characteristic scale $\delta\Omega$ that separates neighboring streams in frequency space is directly related to the time since the accretion of a satellite. In action-angle space, the separation in angles in the i-th direction is

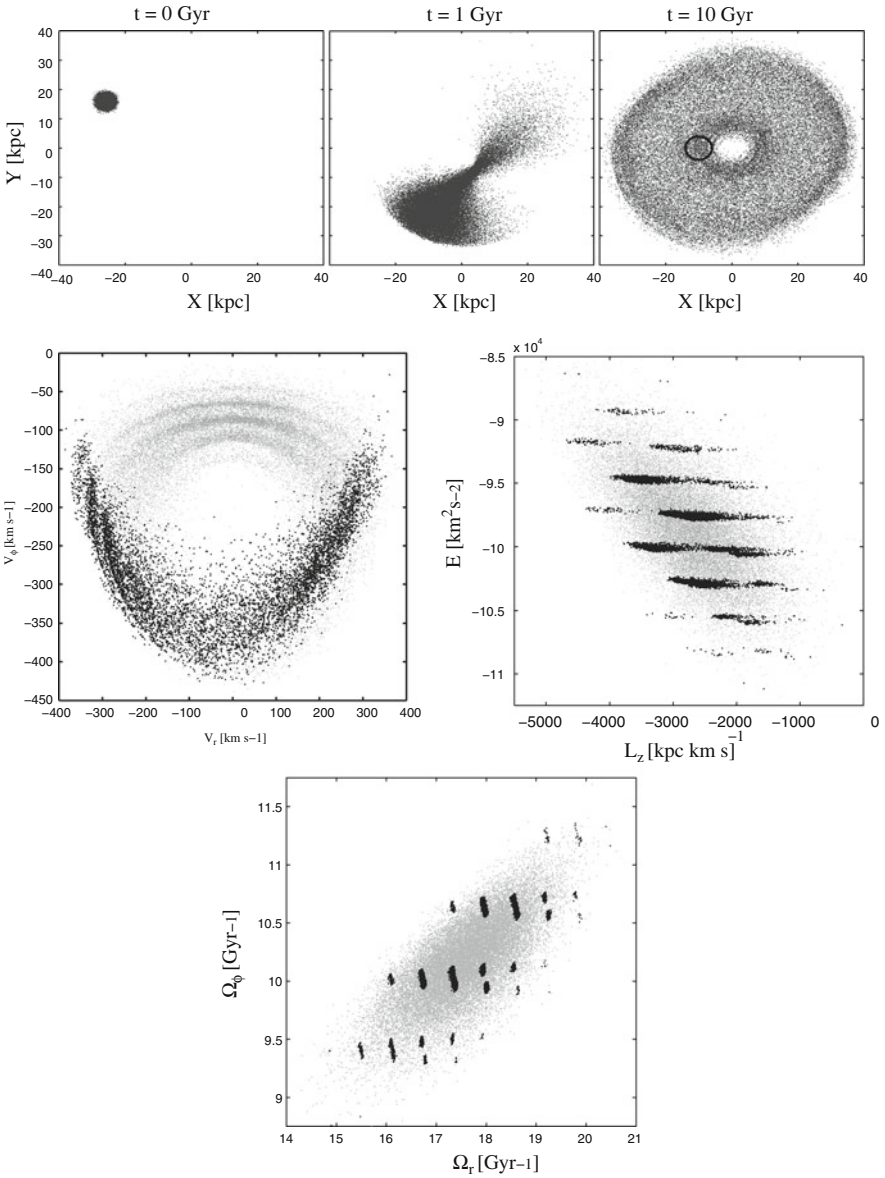


Fig. 6 The *top panels* show the spatial distribution of particles originating in a satellite galaxy accreted onto a Galactic potential. The *middle panels* correspond to different projections of phase-space for those particles located in a "solar neighborhood" sphere (as indicated in the *top right panel*). The *bottom panel* corresponds to the distribution in frequency space for the same set of particles [7]

$$\Delta\Theta_i(t) = [\Delta\Theta_i(0) + \Delta\Omega_i t]_{mod2\pi} = \Delta\Omega_i t \quad (1)$$

For sufficiently long times the number of streams present in a local volume is simply

$$N_{streams} = [\Delta\Omega_i t/(2\pi)] \quad (2)$$

and the characteristic separation between streams is

$$\delta\Omega_i = \Delta\Omega_i / N_{streams} \sim 2\pi/t \quad (3)$$

Therefore, $t_{\rm acc} = 2\pi/\delta\Omega$. We may thus estimate this timescale by measuring the characteristic separation between streams. This is straightforwardly done by computing the Fourier transform of the distribution of streams in frequency space, its power spectrum, and by finding the characteristic wavenumber which has the largest power.

We have carried out many experiments to test how accurately the time since accretion t_{acc} can be retrieved. We have tested cases in which the gravitational potential evolves strongly in time (such that the frequencies are not adiabatically varying), or in live N-body simulations of mergers, and in all cases we are able to measure t_{acc} to within 15–25% [7, 8].

6 Conclusions

We have presented the characteristics of merger debris in the high-resolution Aquarius cosmological simulations of the formation of a dark matter halo. By coupling these simulations with a semi-analytic model of galaxy formation, we have been able to study the assembly of stellar halos as well as their present day properties. In this contribution we have focused on the characteristics of substructure in the sky, in radial velocities and in "solar neighborhood" volumes.

We find that our stellar halos (which are purely accreted by definition) contain a wealth of substructure. At large distances from the galactic center, this is in the form of spatially coherent streams. These streams and various broad overdensities are distributed anisotropically on the sky, in a manner likely to be related to the filamentary pattern present in CDM cosmologies.

At smaller radii, the streams are well-mixed spatially but remain coherent in their kinematics, and are very apparent, for example, in energy–angular momentum space. We thus should expect to be able to disentangle the various building blocks of the stellar halo in the Gaia dataset. This will allow us to unveil the merging history of the Galaxy.

Acknowledgments It is a great pleasure to thank Ken Freeman. Much of the work presented here has been, in one way or another, influenced by his ideas and vision. But more importantly, Ken's unique approach to science and especially his warm and humble attitude have left a mark and set a standard for many of us, collaborators, and friends. I would also like to thank the organizers

for the wonderful celebration of Ken's 70th birthday, and I very much look forward to the next birthday party Ken! My collaborators, especially Andrew Cooper and Facundo Gómez, are gratefully acknowledged. This work has been supported by the Netherlands Organization for Scientific Research through a VIDI research grant and by the European Research Council under ERC-StG grant GALACTICA-240271.

References

1. Battaglia, G., et al. 2005, MNRAS, 364, 433
2. Battaglia, G., et al. 2006, MNRAS, 370, 1055
3. Belokurov, V., et al. 2007, ApJ, 657, L89
4. Bullock, J.S., Johnston, K.V. 2005, ApJ, 635, 931
5. Cooper, A.P., et al. 2010, MNRAS, 406, 744
6. De Lucia, G. Helmi, A. 2008, MNRAS, 391, 14
7. Gómez, F.A., Helmi, A. 2010, MNRAS, 401, 2285
8. Gómez, F.A., Helmi, A., Brown, A.G.A., Li, Y.-S. 2010, MNRAS, arXiv:1004.4974
9. Helmi, A. 2008, A&ARv, 15, 145
10. Helmi, A., White, S.D.M. 1999, MNRAS, 307, 495
11. Helmi, A., White, S.D.M., Springel, V. 2003, MNRAS, 339, 834
12. Ibata, R., Gilmore, G., Irwin, M.J. 1994, Nature, 370, 194
13. Ibata, R., Irwin, M., Lewis, G.F., Stolte, A. 2001, ApJ, 547, L133
14. Ivezić, Ž., et al. 2000, AJ, 120, 963
15. Majewski, S.R., Skrutskie, M.F., Weinberg, M.D., Ostheimer, J.C. 2003, ApJ, 599, 1082
16. Martínez-Delgado, D., Peñarrubia, J., Jurić, M., Alfaro, E.J., Ivezić, Z. 2007, ApJ, 660, 1264
17. McConnachie, A.W., et al. 2009, Nature, 461, 66
18. McMillan, P.J., Binney, J.J. 2008, MNRAS, 390, 429
19. Morrison, H.L., Boroson, T.A., Harding, P. 1994, AJ, 108, 1191
20. Newberg, H.J., Yanny, B. 2006, J Phys Conf Ser, 47, 195
21. Smith, M.C., et al. 2007, MNRAS, 379, 755
22. Springel, V., et al. 2008, MNRAS, 391, 1685
23. Watkins, L.L., et al. 2009, MNRAS, 398, 1757
24. Xue, X.X., et al. 2008, ApJ, 684, 1143
25. Yanny, B., et al. 2000, ApJ, 540, 825
26. Zibetti, S., White, S.D.M., Brinkmann, J. 2004, MNRAS, 347, 556

The Outer Halos of Elliptical Galaxies

Ortwin Gerhard

Abstract Recent progress is summarized on the determination of the density distributions of stars and dark matter, stellar kinematics, and stellar population properties in the extended, low-surface brightness halo regions of elliptical galaxies. With integral field absorption spectroscopy and with planetary nebulae as tracers, velocity dispersion and rotation profiles have been followed to $\sim$ 4 and $\sim$ 5−8 effective radii, respectively, and in M87 to the outer edge at $\sim$ 150 kpc. The results are generally consistent with the known dichotomy of elliptical galaxy types, but some galaxies show more complex rotation profiles in their halos and there is a higher incidence of misalignments, indicating triaxiality. Dynamical models have shown a range of slopes for the total mass profiles and that the inner dark matter densities in ellipticals are higher than in spiral galaxies, indicating earlier assembly redshifts. Analysis of the hot X-ray emitting gas in X-ray bright ellipticals and comparison with dynamical mass determinations indicate that non-thermal components to the pressure may be important in the inner $\sim$ 10 kpc and that the properties of these systems are closely related to their group environments. First results on the outer halo stellar population properties do not yet give a clear picture. In the halo of one bright galaxy, lower [α/Fe] abundances indicate longer star formation histories pointing toward late accretion of the halo. This is consistent with independent evidence for ongoing accretion and suggests a connection to the observed size evolution of elliptical galaxies with redshift.

1 Introduction

There are indications that the outer halos of bright elliptical galaxies formed later and through different processes than their inner parts. It is clearly important to establish whether this is true for some, most, or all ellipticals. Signatures of the formation

O. Gerhard (✉)
MPE, Giessenbachstrasse, D-85748 Garching, Germany
e-mail: gerhard@mpe.mpg.de

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_29,

processes may be preserved longer in the outer halos than in the well-mixed inner parts, because of the longer dynamical timescales at large radii. Owing to the low surface brightness in these halos, the outer kinematics and stellar population properties can be best studied in local galaxies. Combining the results with predictions from simulations and with high-redshift studies of ellipticals one may expect to obtain a global picture of halo formation in elliptical galaxies. The following sections will discuss some recent results on the stellar distribution and kinematics in the outer halos (§2), the dark matter distribution and dynamics (§3), and finally the stellar populations and assembly history of nearby elliptical galaxy halos together with related results from high-redshift studies and simulations (§4).

The subject of the outer halos of ellipticals is one that has benefitted enormously from the vision of Ken Freeman, whose scientific work we celebrate with this conference. Ken was one of the first to realize the potential of using planetary nebulae (PNe) as kinematic tracers in regions where the surface brightness is too low for traditional absorption line spectroscopy. In 1987, in the course of a project with X. Hui, H. Ford, and M. Dopita, he measured the first PN radial velocities in the halo of Centaurus A, the nearest elliptical galaxy, using a multi-fiber instrument at the Anglo-Australian Telescope. Together with later AAT and CTIO measurements, they were able to obtain 433 PN velocities in this galaxy [32]. Together with M. Arnaboldi and others, he extended this work to the fainter PNe in the halos of more distant ellipticals such as NGC 1399, M86, and NGC 1316 [2–4], in order to constrain the kinematics and dark matter mass distribution in these galaxies.

The potential shown by these papers led to the idea of a special-purpose instrument, the Planetary Nebula Spectrograph (PNS), which is based on the principle of using counterdispersed imaging to simultaneously find PNe and measure their radial velocities [21, 23]. This instrument is currently mounted at the William Herschel Telescope in La Palma. As two highlights, some 2600 PN velocities were measured in nearby M31 [39] and within a still ongoing project ∼100–200 PN velocities have so far been obtained for each of a sample of some 15 ellipticals [13]. The 1996 paper on M86 also reported three discrepant PN velocities, ascribed to the intracluster stellar population in Virgo. Since then, Ken has enthusiastically participated in using PNe to study the dynamics of intracluster stars in nearby galaxy clusters [26]; see also Chapter "Diffuse Light and Galaxy Intaractions in the Core at nearby clusters" by M. Arnaboldi's, this volume. For the first multi-slit imaging observations of PNe in the Coma cluster [28] he went observing at the Subaru telescope. Ken is a coauthor of some 30 refereed papers related to the use of PNe for kinematic measurements, including also the Galactic bulge [5] and, again, Centaurus A [45].

2 Density Distribution and Kinematics of Outer Halo Stars

Accurate brightness profiles over large radius ranges were constructed by Kormendy et al. [35] for Virgo cluster elliptical and spheroidal galaxies. The brightest elliptical galaxies in this sample have profiles characterized by large Sérsic indices $n \simeq 4-12$,

large effective radii $R_e > 10$ kpc at faint surface brightnesses, $\mu_{e,V} \simeq 23$–25, and central cores (depressions relative to the Sérsic profile). Fainter ellipticals have Sérsic indices $n \sim 2$–4, R_e's of a less than a few kiloparsec $\mu_{e,V} \simeq 19$–22, and central cusps (extra light relative to the Sérsic profile). While the dichotomy in the central profiles of both classes is clear and is their primary distinction, there are some transition cases in the large-radius profiles. The important point for this talk is that the (bright) core ellipticals have the most extended luminous halos and that within each of the two classes n (extent) does not correlate with luminosity [6, 35].

Outer halo kinematics based on PN velocities were analyzed by Coccato et al. [13] for a sample of 15 early-type galaxies. These data typically reach to $\sim 5R_e$ and in some galaxies to $8R_e$. The derived profiles of PN number distribution and kinematics agree with surface brightness and absorption line kinematics where the data sets overlap, within their respective errors. The radial profiles of mean rms velocity $v_{\rm rms}$ for this sample fall within two groups, with part of the galaxies having slowly decreasing $v_{\rm rms}(R)$ and the remainder having steeply falling $v_{\rm rms}(R)$. There is a large overlap of these groups with the core and cusp ellipticals, respectively [30, 31, 35]. The outer halos of ellipticals tend to be more rotationally dominated than the central parts. The slow-rotator versus fast-rotator classification of Emsellem et al. [24] based on inner Sauron data is largely preserved in the halos; however, some more complex profiles of specific angular momentum parameter λ_R are seen in the halos. Slow rotators typically reach $\lambda_R \simeq 0.1$ in their outer parts. Lastly, twists and misalignments in the velocity fields, commonly interpreted as signs of triaxiality, seem to be more frequent at large radii [13].

Proctor et al. [47] have pioneered a technique to measure absorption line kinematics for the integrated halo stellar population using data from multi-object slitlet spectroscopy and to construct 2D velocity fields to $3R_e$ from these data. They find a variety of stellar rotation profiles beyond $1R_e$, constant, decreasing, and increasing, confirming that the fast–slow rotator classification may be more complicated in the halos than within R_e. Halo kinematics obtained with integral field units (IFUs) such as Sauron and VIRUS-P reach somewhat further to $\sim 4R_e$ [40, 57]. In these studies, the signal from many lenslets or fibers is co-added to increase the signal-to-noise. Because of the limited field-of-view, several IFU fields may need to be exposed. For the galaxies where kinematics is available both from these absorption line techniques and from PNe, the velocity dispersion profiles agree within the errors. The integrated light techniques yield measurements also of the line profile shape parameters, giving valuable constraints on the anisotropy of the halo orbit distribution. These are harder to obtain from PNe unless large samples can be obtained. Since PNe can be found out to larger radii, the best dynamical constraints will come from combining both techniques.

The extreme outer halo of the Virgo-centric giant elliptical galaxy M87 was studied by Doherty et al. [20]. They used small samples of PNe in fields at 60 and 150 kpc to measure the velocity dispersion, after removing Virgo intracluster stars. PNe centered around the systemic velocity of M87 were found only out to 150 kpc, but not further, and the velocity dispersion profile was found to decline steeply toward the outer edge. Given that the X-ray-determined mass profile is rising steeply

at these radii and is dominated by the cluster, this is dynamically possible only if the halo of M87 is truncated. McNeil et al. [38] studied the central bright galaxy in the Fornax cluster, NGC 1399. Again using PNe, they found that a component of low-velocity PNe (by $\sim 800\,\mathrm{km\,s^{-1}}$) is superposed on the main galaxy population. When this component is removed, the velocity dispersion profile of NGC 1399 is flat at $\sim 200\,\mathrm{km\,s^{-1}}$ from PNe and red globular clusters out to 80 kpc. Both studies show that the outermost halos of elliptical galaxies are in interaction with the environment and that discrete tracers are important to disentangle the different (mixed and unmixed) components in the velocity distribution.

3 Dark Matter and Dynamics

The dark matter distribution in elliptical galaxies can be studied with a variety of techniques, including strong and weak lensing, X-ray emitting hot gas, and dynamics. The lensing results are discussed in K. Kuijken's talk in this volume; in brief, while with weak lensing data the dark halos of elliptical galaxies can be followed to hundreds of kiloparsec, the strong lensing analyses give accurate mass measurements within the Einstein radius, and together with central velocity dispersions constrain the slopes of the mass density profiles to an average value of very nearly -2 (isothermal), with small scatter.

Early dynamical analysis of stellar velocity dispersions and line profile shape data with non-parametric spherical models first indicated that the inner circular velocity curves (CVC) of elliptical galaxies are flat to $\sim 10\%$, out to approximately 1–2 R_e [29, 36]. Recent work by Thomas et al. [51] on a sample of ellipticals in the Coma cluster with slightly more extended data (1–3 R_e) has shown more varied CVC slopes, with some falling, some flat, and some rising. Using axisymmetric Schwarzschild models to constrain the dark matter halos in the Coma ellipticals, Thomas et al. [52] found that the dark matter densities in these galaxies are on average $7\times$ higher than in spiral galaxies of the same luminosity, and $13\times$ higher than in spirals of the same baryonic mass, consistent with the earlier analysis of round nearby galaxies by Gerhard et al. [29]. Baryonic contraction is not sufficient to explain the difference, implying that the inner halos of ellipticals presumably formed earlier than those of spiral galaxies. If one assumes that the halo densities measure the density of the Universe at the time of collapse [29], the assembly redshift $(1 + z_{\mathrm{ass}})$ for the Coma ellipticals is twice that for typical spiral galaxies, and puts their $z_{\mathrm{ass}} = 1, \ldots, 3$ [52], with brighter ellipticals assembling later.

Several intermediate-luminosity ellipticals have steeply falling outer PN velocity dispersion profiles [13, 48]. The prototype galaxy for this class is NGC 3379, for which some 200 PN velocities are available. It was thus argued [22, 48] that this galaxy might be "naked", i.e., lack a significant dark matter halo. Extensive recent dynamical modeling by Lorenzi et al. [17], using the made-to-measure particle code NMAGIC [16], has shown that a variety of dark matter halos are allowed by the data, implying radially anisotropic orbit distributions, and that the no-dark matter halo is

ruled out (see also [57]). However, the allowed halo density distributions imply that the CVC in this galaxy falls by at least 25% between the center and $2R_e$, i.e., they are of significantly lower density than would be required for a flat CVC. Similar results were found for two other galaxies of this class, NGC 4697 and NGC 4494 [18, 43].

There is a long history of mass determination in X-ray bright elliptical galaxies, using the hot gas as a tracer in assumed hydrostatic equilibrium in the gravitational potential (e.g., [27, 33, 42, 44]). Based on a comparison of the gravitational potentials and circular velocities inferred from X-rays and from dynamics, Churazov et al. [8, 9] argued that the non-thermal sources of pressure contribute on average 20–30% of the thermal gas pressure. Das et al. [15] recently devised a non-parametric Bayesian technique to determine CVCs and confidence ranges from deprojected hot gas pressure and temperature profiles. Comparing with dynamical mass determinations, they also find differences indicating that non-thermal pressure sources are important in the central 10 kpc of their analyzed six galaxies. Other possible interpretations for these differences are violations of hydrostatic equilibrium or biases introduced by the assumptions in the dynamical models; this is an important issue that needs to be clarified. The dark matter mass fractions inferred by Das et al. [15] are 35–80% at $2R_e$, rising to 80–90% at the outermost radii. Das et al. [15] also find that for all of their galaxies the outer CVCs are rising, implying steeper than isothermal mass profiles. Furthermore, the inferred circular velocities and the luminosities of their galaxies are found to correlate with the velocity dispersion of the environment. Both results suggest that the properties of X-ray bright ellipticals are closely related to their group environments.

4 Stellar Population and Assembly History

Studies of the stellar populations in the outer halos of early-type galaxies may give important constraints on their assembly history. Due to the faint surface brightnesses in the outer regions this work is only in its beginning, requiring large telescopes and/or special techniques.

In a recent study, Coccato et al. [11] used deep-medium resolution spectroscopy with FOCAS at the Subaru telescope to measure Lick line-strength indices far into the halo of one of the two Coma brightest cluster galaxies (BCGs), NGC 4889. Combining these with literature data, Coccato et al. [12] constructed radial profiles of metallicity, [α/Fe] abundance ratio and age, from the center out to $\sim$ 60 kpc ($\sim 4R_e$). These profiles show evidence for different chemical and star formation histories for stars inside and outside $\sim 1.2R_{\rm e} = 18$ kpc radius. The inner regions have a steep metallicity gradient and high [α/Fe] at $\sim$ 2.5 solar value, pointing to a rapid formation process. This is consistent with a quasi-monolithic, dissipative merger collapse at early redshifts [34], followed by one or at most a few dry mergers between several such units so that the original steep metallicity gradient [7, 46] was only partially erased [19]. In the halo of NGC 4889, between 18 and 60 kpc, the metallicity is near-solar with a shallow gradient, while [α/Fe] shows

a strong negative gradient, reaching solar values at 60 kpc, and the inferred ages are 9–13 Gyr. This argues for later accretion of stars from old systems with more extended star formation histories [37], presumably through minor mergers from the galaxy's group environment [1, 41].

Rudick et al. [49] study the color distribution in the outer halo of M87 and the adjacent intracluster light. They find that the M87 halo becomes bluer with radius to the outermost radii and argue that the common colors of the outer halo and of some surrounding intracluster tidal features suggest that the galaxy's envelope may have formed from similar streams. Foster et al. [25] used the technique from Proctor et al. [47] to study three galaxies. Their most extended data set is for NGC 2768 which they suggest, based on comparison with simulations by Hopkins et al. [30], is consistent with a dissipative merger origin. The two other datasets have too large errors beyond R_e to allow firm conclusions. Weijmans et al. [57] find from co-added Sauron field data that the metallicity gradients in the two intermediate luminosity ellipticals NGC 3379 and NGC 821 remain approximately constant to $4R_e$, with old and subsolar metallicity populations in the halos. Thus outer halo stellar population studies to date do not yet combine to a clear picture.

The bimodal stellar populations found by Coccato et al. [12] in the core and outer halo of NGC 4889 are consistent with recent results on the size evolution of early-type galaxies (ETGs) with redshift (e.g., [10, 14, 53]), such that ETGs at $z \sim 1$ ($z \sim 2$) have sizes a factor of 2 (3–5) smaller than ETGs with similar mass today [50, 54, 55]. Recently, van Dokkum et al. [56] find from stacked rest-frame R-band images that massive ETGs have nearly constant mass inside 5 kpc with redshift, but increase their envelope mass by a factor ~ 4 since $z = 2$, with effective radius evolving as $R_e \propto (1 + z)^{-1.3}$. Minor mergers may play an important role in this process [41], and the accreted galaxies are likely to have gone through longer star formation histories, leading to less α-enriched stellar populations as was found in the halo of NGC 4889. Further work on the stellar populations in the outer halos of local elliptical galaxies may become a valuable tool to understand what part of the elliptical galaxy population participates in the observed size evolution.

Acknowledgments First, it is a pleasure on this occasion to thank Ken Freeman for the long-standing collaboration and the many stimulating discussions over the years. I would also like to thank David Block and his team for organizing this memorable conference in Ken's honour, and for all the extra work they did in order to be able to hold it in this fantastic location in the Namibian desert.

References

1. Abadi, M.G., Navarro, J.F., Steinmetz, M. 2006, MNRAS, 365, 747
2. Arnaboldi, M., Freeman, K.C., Gerhard, O., et al. 1998, ApJ, 507, 759
3. Arnaboldi, M., Freeman, K.C., Hui, X., Capaccioli, M., Ford, H. 1994, Messenger, 76, 40
4. Arnaboldi, M., Freeman, K.C., Mendez, R.H., et al. 1996, ApJ, 472, 145
5. Beaulieu, S.F., Freeman, K.C., Kalnajs, A.J., Saha, P., Zhao, H. 2000, AJ, 120, 855
6. Caon, N., Capaccioli, M., D'Onofrio, M. 1993, MNRAS, 265, 1013
7. Chiosi, C., Carraro, G. 2002, MNRAS, 335, 335

8. Churazov, E., Forman, W., Vikhlinin, A., et al. 2008, MNRAS, 388, 1062
9. Churazov, E., Tremaine, S., Forman, W., et al. 2010, MNRAS, 404, 1165
10. Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, A&A, 482, 21
11. Coccato, L., Arnaboldi, M., Gerhard, O., et al. 2010a, A&A in press, arXiv:1006.1894
12. Coccato, L., Gerhard, O., Arnaboldi, M. 2010b, MNRAS, 407, L26
13. Coccato, L., Gerhard, O., Arnaboldi, M., et al. 2009, MNRAS, 394, 1249
14. Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, ApJ, 626, 680
15. Das, P., Gerhard, O., Churazov, E., Zhuravleva, I. 2010, MNRAS in press, arXiv:1007.5322
16. de Lorenzi, F., Debattista, V.P., Gerhard, O., Sambhus, N. 2007, MNRAS, 376, 71
17. de Lorenzi, F., Gerhard, O., Coccato, L., et al. 2009, MNRAS, 395, 76
18. de Lorenzi, F., Gerhard, O., Saglia, R.P., et al. 2008, MNRAS, 385, 1729
19. di Matteo, P., Pipino, A., Lehnert, M.D., Combes, F., Semelin, B. 2009, A&A, 499, 427
20. Doherty, M., Arnaboldi, M., Das, P., et al. 2009, A&A, 502, 771
21. Douglas, N.G., Arnaboldi, M., Freeman, K.C., et al. 2002, PASP, 114, 1234
22. Douglas, N.G., Napolitano, N.R., Romanowsky, A.J., et al. 2007, ApJ, 664, 257
23. Douglas, N.G., Taylor, K., Freeman, K.C., Axelrod, T.S. 1997, In: IAU Symp., Planetary Nebulae, ed. H.J. Habing, H.J.G.L.M. Lamers, Kluwer, Dordrecht, vol. 180, p. 493
24. Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, MNRAS, 379, 401
25. Foster, C., Proctor, R.N., Forbes, D.A., et al. 2009, MNRAS, 400, 2135
26. Freeman, K.C., Arnaboldi, M., Capaccioli, M., et al. 2000, In: Astronomical Society of the Pacific Conference Series, Dynamics of Galaxies: From the Early Universe to the Present, eds. F. Combes, G.A. Mamon, V. Charmandaris, ASP, San Francisco, CA, vol. 197, p. 389
27. Fukazawa, Y., Botoya-Nonesa, J.G., Pu, J., Ohto, A., Kawano, N. 2006, ApJ, 636, 698
28. Gerhard, O., Arnaboldi, M., Freeman, K.C., et al. 2005, ApJ, 621, L93
29. 29 Gerhard, O., Kronawitter, A., Saglia, R.P., Bender, R. 2001, AJ, 121, 1936
30. Hopkins, P.F., Cox, T.J., Dutta, S.N., et al. 2009, ApJS, 181, 135
31. Hopkins, P.F., Lauer, T.R., Cox, T.J., Hernquist, L., Kormendy, J. 2009, ApJS, 181, 486
32. Hui, X., Ford, H.C., Freeman, K.C., Dopita, M.A. 1995, ApJ, 449, 592
33. Humphrey, P.J., Buote, D.A., Gastaldello, F., et al. 2006, ApJ, 646, 899
34. Kobayashi, C. 2004, MNRAS, 347, 740
35. Kormendy, J., Fisher, D.B., Cornell, M.E., Bender, R. 2009, ApJS, 182, 216
36. Kronawitter, A., Saglia, R.P., Gerhard, O., Bender, R. 2000, A&AS, 144, 53
37. Matteucci, F. 1994, A&A, 288, 57
38. McNeil, E., Arnaboldi, M., Freeman, K.C., et al. 2010, A&A, 518, A44
39. Merrett, H.R., Merrifield, M.R., Douglas, N.G., et al. 2006, MNRAS, 369, 120
40. Murphy, J., Gebhardt, K. 2007, BAAS, 38, 747
41. Naab, T., Johansson, P.H., Ostriker, J.P. 2009, ApJ, 699, L178
42. Nagino, R., Matsushita, K. 2009, A&A, 501, 157
43. Napolitano, N.R., Romanowsky, A.J., Coccato, L., et al. 2009, MNRAS, 393, 329
44. Nulsen, P.E.J., Böhringer, H. 1995, MNRAS, 274, 1093
45. Peng, E.W., Ford, H.C., Freeman, K.C. 2004, ApJ, 602, 685
46. Pipino, A., D'Ercole, A., Matteucci, F. 2008, A&A, 484, 679
47. Proctor, R.N., Forbes, D.A., Romanowsky, A.J., et al. 2009, MNRAS, 398, 91
48. Romanowsky, A.J., Douglas, N.G., Arnaboldi, M., et al. 2003, Science, 301, 1696
49. Rudick, C.S., Mihos, J.C., Harding, P., et al. 2010, ApJ, 720, 569
50. Saracco, P., Longhetti, M., Andreon, S. 2009, MNRAS, 392, 718
51. Thomas, J., Saglia, R.P., Bender, R., et al. 2007, MNRAS, 382, 657
52. Thomas, J., Saglia, R.P., Bender, R., et al. 2009, ApJ, 691, 770
53. Trujillo, I., Conselice, C.J., Bundy, K., et al. 2007, MNRAS, 382, 109
54. van derWel, A., Holden, B.P., Zirm, A.W., et al. 2008, ApJ, 688, 48
55. van Dokkum, P.G., Franx, M., Kriek, M., et al. 2008, ApJ, 677, L5
56. van Dokkum, P.G., Whitaker, K.E., Brammer, G., et al. 2010, ApJ, 709, 1018
57. Weijmans, A., Cappellari, M., Bacon, R., et al. 2009, MNRAS, 398, 561

Galaxies: Lighthouses in the Shoals of Dark Halos

R. Brent Tully

Abstract It is anticipated from hierarchical clustering theory that there are scaling relationships between halos over a wide range of mass. Observationally it can be difficult to identify the markers that characterize these relationships because of the small numbers of visible probes and confusion from contaminants in projection. Nonetheless, in favorable circumstances it is possible to identify a very useful marker: the radius of the caustic at second turnaround. In a few favorable circumstances it is possible to identify the radius of first turnaround or zero velocity surface about a collapsed region. It will be shown that specifically the radius of second turnaround scales as anticipated over three orders of magnitude in mass from 10^{12} to 10^{15} $M_{\odot}$. Halos are characterized by zones of dispersed velocities within the second turnaround caustic and zones of infall between the first and the second turnaround radii. The inner zone is populated in the majority by gas-poor morphologies and the outer zone is populated in the majority by gas-rich morphologies. The numbers of dwarfs within the inner zone is roughly constant per unit halo mass.

1 Introduction

In the imaginary world of simulations, researchers have a well-developed picture of the collapse of matter into halos. Over time, small halos are absorbed into larger units. Collapsed regions filled with substructure can be defined by hundreds or thousands of particles. Halos can be identified with precision within those simulations [3].

In the real world, most galaxies are observed to lie in groups or clusters. However, membership may be so limited that the structure is ill defined. There is no concensus among observers about what is meant by the terms "group" and "cluster." For example, it is commonly accepted that we live in something called the Local Group. A quantitative boundary that would include the traditional members is the zero-velocity surface [12]. Galaxies inside this surface are infalling or on more

R.B. Tully (✉)
Institute for Astronomy, University of Hawaii, Honolulu, Hawaii
e-mail: tully@ifa.hawaii.edu

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_30,

complex-bound orbits. Galaxies outside this surface are participating in the Hubble expansion. But now consider the Virgo Cluster. This entity would traditionally be taken to include the region within 2 Mpc where several thousand galaxies are following randomized orbits [2]. The zero-velocity surface around the Virgo Cluster lies at a radius of ~ 7 Mpc, almost half of our distance of 16.7 Mpc from the cluster. The implicit traditional definitions of the Local Group and the Virgo Cluster are inconsistent.

It does not seem useful to draw a distinction between the terms "cluster" and "group." Clusters contain a lot of galaxies and groups contain only a few, but there is no clear demarcation between the two. In this chapter, the two terms will be used interchangeably.

So how should an observer define a group? Structure forms as sufficient matter accumulates through gravitational attraction to decouple from the expansion of the Universe. The matter falls together on nearly radial orbits and, given enough time, evolves toward dynamic equilibrium. Two natural dimensions to describe a collapsed structure are the gravitational or virial radius, r_g, and the radius at 200 times the "critical" density for a matter-dominated closed universe, r_{200}. Here, lower case r implies a 3-dimensional radius and an upper case R will be used to denote a projected radius. Accordingly, if all galaxies in a group sample are given equal weight (i.e., galaxies of whatever luminosity are considered to be test particles within an environment dominated by distributed dark matter) then the virial radius is defined as

$$R_g = \frac{\mathrm{N}^2}{\sum_{i<j} 1/R_{ij}} \tag{1}$$

where R_{ij} is the projected distance between pairs and a virial mass estimate is given for the group by

$$M_\mathrm{v} = \alpha \sigma_p^2 r_g / G = (\pi/2)\alpha \sigma_p^2 R_g / G \tag{2}$$

where the line-of-sight velocity dispersion for the sample of N galaxies is

$$\sigma_p = \sqrt{\sum_i (v_i - < v >)^2 / \mathrm{N}} \tag{3}$$

The group mean velocity is $< v >$ and $\alpha = 3$ if orbits are isotropically distributed.

The alternative dimension r_{200} is defined to coincide with a density $200\rho_{crit}$ where $\rho_{crit} = 3\mathrm{H}_0^2/8\pi G$ at redshift zero. This parameter can be determined in simulations with large numbers of test particles but it is not such a useful construct in the context of observations of small groups. However, Mahdavi et al. [16] calculate a ratio between r_{200} and r_g assuming $M(r) \propto r$ that can be reduced to the relationship (at $z = 0$)

$$r_{200} = \frac{\sqrt{\alpha}\sigma_p}{10\mathrm{H}_0} \tag{4}$$

There may be other dimensions that are observationally useful. Consider the collapse of a spherically symmetric overdense region in the expanding universe [1, 6]. The time of collapse, t_c, depends on density, ρ: $t_c \propto \rho^{-1/2}$. In the approximation of spherical collapse, all across the universe at $t_c = today$ overdense regions with a common density will be separating from the cosmic expansion, creating zero velocity surfaces. The radius that encloses one of these regions will be called the first turnaround radius, r_{1t}.

At earlier times, regions of successively higher density collapsed. There was a time and a corresponding density that lead to collapse and re-expansion to a pause before recollapse. This pause and turnaround create a caustic [21] at what will be called the second turnaround radius, r_{2t}.

At yet earlier times, regions of yet higher density collapsed and resulted in caustics of higher order turnarounds. In the real world, departures from spherical symmetry would result in violent relaxation [15], departures from radial orbits, and blurring of the caustics. The collapsed region will begin to evolve toward energy equipartition. Such regions will be referred to as "quasi-virialized."

These theoretical musings provide a basis for a definition of groups. It will be argued that radii of second turnaround, and more problematically of first turnaround, are observable.

The caustic of second turnaround for a group can potentially be defined by two features: a density discontinuity and a velocity dispersion discontinuity. These features were displayed by Bertschinger [1] in the simple case of spherical infall of an initial top hat density excess. In this simple case, density falls off as a power law with radius within r_{2t} and velocity dispersions are large. Beyond r_{2t}, the density is expected to take a downward step and velocities locally are expected to manifest coherent infall with modest dispersion. The same salient features are recovered with N-body simulations based on the same spherical infall model [27]. These features might be hard to identify in observational information. The density step might be obscured by noise from line-of-sight contamination. The step from large to small velocity dispersions can easily be confused by the limitation of access to only radial velocities and lack of precise distances that would distinguish backside from frontside infall. The situation with many groups will be messy but there are relatively clean cases where the desired information can be recovered. For example, one can identify the caustics in redshift–radius plots for rich clusters [5, 19]. The challenge given attention here is to find analogous features in groups with much smaller populations. The challenge can be met if the entities are sufficiently isolated or so nearby that their three-dimensional structure can be discerned.

The musings suggest that there should be interesting scaling relations between groups. In the spherical approximation, the density at r_{2t} is the same for all groups

$$\frac{M(r_{2t})}{r_{2t}^3} = constant \tag{5}$$

With the approximation $M(r) \propto r$

$$M(r_{2t}) = M_{\rm v}(r_{2t}/r_g) \tag{6}$$

so

$$\frac{M_{\rm v}}{r_{2t}^3} \cdot \frac{r_{2t}}{r_g} = \frac{\sigma_p^2}{r_{2t}^2} = constant \tag{7}$$

Hence

$$r_{2t} \propto \sigma_p \propto M_{\rm v}^{1/3} \tag{8}$$

In subsequent sections, data will be presented that confirm these relationships and provide correlation zero points.

There is the prospect that equivalent scaling relationships can be found based on the first turnaround radius, r_{1t}. The interest here is that the ratio r_{1t}/r_{2t} is dependent on the dark energy content of the Universe. Collapse conditions at a given time depend almost exclusively on the matter density but dark energy affects the clock, hence the timing between collapse events. In a flat universe with 70% of the cosmological density in vacuum energy then $r_{1t}/r_{2t} = 3.3$ whereas in a flat universe with the critical density in matter then $r_{1t}/r_{2t} = 3.7$ (Gary Mamon, private communication).

Finally it is to be entertained if there are observational manifestations of the "zero gravity" surface at a radius r_{zg} around a halo. Only material within this surface will ever participate in collapse onto the halo. It has been suggested [4] that the region of the zero gravity surface is expected to be evacuated and can be identified by this property. The importance of this feature depends on the properties of dark energy.

2 From Big to Small

The questions to be addressed are whether there are observational features that distinguish the quasi-virialized and infall regions of halos and, if so, how these features scale with halo mass. Most of the discussion in this chapter will focus on modest to very small groups of galaxies, the visible manifestations for the overwhelming majority of identifiable halos. The sites of intermediate mass halos, those in the range $5 \times 10^{12} - 10^{14}$ $M_\odot$, have been investigated with a wide-field imaging survey at the Canada–France–Hawaii Telescope (CFHT), initially using the 0.3 deg^2. 12K CCD camera and then the 1 deg^2 Megacam detector. [7]. Follow-up spectroscopy was undertaken with Subaru and Keck telescopes. The study is extended to very small halos in the range $10^{11} - 5 \times 10^{12}$ $M_\odot$ by giving consideration to the region within 4 Mpc where, outside the zone of obscuration, almost every galaxy brighter than $M_B = -11$ has probably been identified [11]. Almost all of these nearby galaxies have now been sufficiently observed with Hubble Space Telescope (HST)

that an accurate distance is available from a measurement of the luminosity of the tip of the red giant branch, the TRGB method. There is now detailed information on the grouping properties of galaxies down to the scales of associations of dwarfs [24]. Page limitations prevent a detailed discussion of the many individual group environments that have been given attention. The following is an expurgated version of a study that will be flushed out in a journal article.

The Coma Cluster rests in the nearest massive and evolved halo at 2×10^{15} $M_{\odot}$ and the Virgo Cluster is the dominant structure in the Local Supercluster at 8×10^{14} $M_{\odot}$. Each is continuing to grow by infall along filaments. The dimensions orthogonal to these filaments are well defined. It is shown in Fig. 1 that density discontinuities give good definition of the caustics of second turnaround. In the case of the Virgo Cluster where there is improving distance resolution the location of the caustic is supported by velocity information.

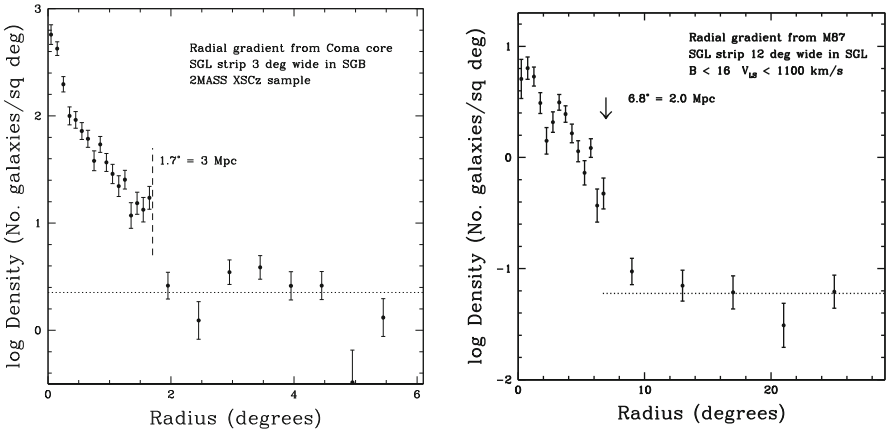


Fig. 1 Radial distributions of galaxies in the Coma and Virgo clusters. The density discontinuities at 3 and 2 Mpc, respectively, are associated with the caustics of second turnaround

The CFHT wide-field imaging survey gave attention to three groups with masses in the range $2 - 8 \times 10^{13}$ $M_{\odot}$ at distances $25 - 29$ Mpc [17, 23, 26]. The evidence for the detection of second turnaround caustics is shown for two of these cases in Fig. 2. Two of these intermediate mass halos are dynamically evolved, with central dominant ellipticals and gas-poor companions, while the third presents a situation with mixed morphologies.

In an effort to understand what is going on in low-density environments, a substantial area around the galaxy NGC 1023 was imaged with the CFHT cameras [22]. The central halo has a mass of 6×10^{12} $M_{\odot}$. The detection of the second turnaround caustic and the morphological segregation of galaxies are seen in Fig. 3.

The very complete information on the three-dimensional distribution of galaxies within 4 Mpc lets us look at very small halos. The largest of the nearby halos is centered on Centaurus A and has a mass of 1×10^{13} $M_{\odot}$. It is seen in Fig. 4 that the

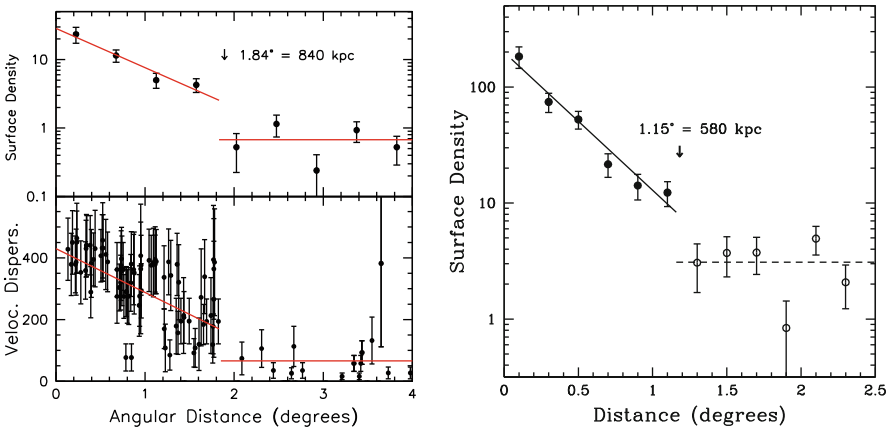


Fig. 2 Radial distributions of galaxies in the NGC 5846 and N5353/4 groups. The density discontinuities at 840 and 580 kpc, respectively, are associated with the caustics of second turnaround. The velocity discontinuity at the caustic is shown in the case of NGC 5846

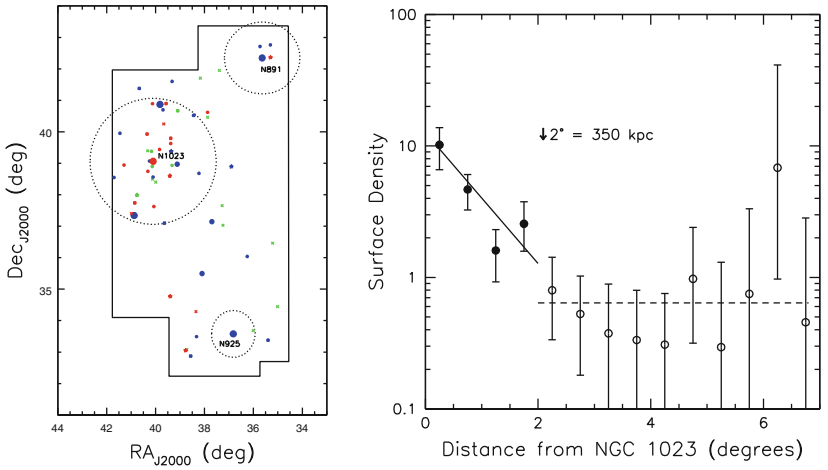


Fig. 3 *Left:* Galaxies in the vicinity of NGC 1023. *Red*: early; *blue*: late; *green*: transition. *Irregular box*: region of CFHT wide-field imaging survey. *Dotted circles*: inferred caustics of second turnaround for the three dominant halos. *Right:* Radial distribution of galaxies with respect to NGC 1023. The caustic of second turnaround is associated with the discontinuity at 350 kpc

Cen A and M83 halos are discrete and the Cen A halo is host to a substantial family of gas-poor dwarfs [13, 14].

Figure 5 gives attention to the historical Local Group, an environment with two discrete halos with masses $1 - 2 \times 10^{12}$ $M_\odot$ within a common infall zone. There is clear morphological segregation between the regimes inside and outside the second turnaround caustics. There is reasonable definition of the surface of first turnaround [12].

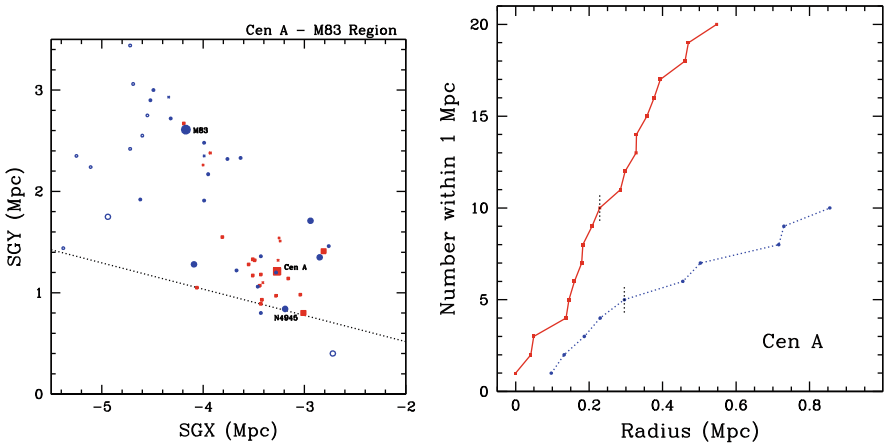


Fig. 4 *Left:* Distribution of galaxies in the region of Cen A and M83. The *slanted dotted line* is a projection of galactic latitude +15°. Accurate distances reveal that Cen A and M83 are the dominant galaxies in two distinct groups. *Right:* The cumulative number of early (*solid red*) and late (*dotted blue*) galaxies as a function of projected radius from Cen A

The M81 group provides a relatively clean situation [9]. The central halo of 4×10^{12} $M_\odot$ is characterized by a large dispersion of velocities and it is surrounded by a pattern of infall extending to a first turnaround or zero velocity surface at 1.3 kpc (see Fig. 6). The second panel of this figure illustrates the more uncertain situation with the $\sim 1 \times 10^{12}$ $M_\odot$ halo centered on NGC 253 [8, 10].

Finally, Fig. 7 illustrates the velocity patterns around two halos in the range $4 - 7 \times 10^{11}$ $M_\odot$ that only contain minor galaxies with luminosities of the LMC or less [23]. These are interesting because luminosities are so small that mass to light ratios are high. It is suspected that smaller halos become unidentifiable because of a lack of galaxy signposts.

3 Scaling Relations

It was anticipated in the introduction that the two independent variables r_{2t} and σ_p would be directly correlated. The relationship is seen in the left panel of Fig. 8. The constant of proportionality is found empirically to be 350 km s^{-1} Mpc^{-1}.

Two interesting relations are shown in the right panel of Fig. 8 and the left panel of Fig. 9. Each involves the virial mass calculated according to

$$M_v = \frac{\sigma_p^2 r_g}{G} \tag{9}$$

where the virial radius r_g was defined in Eq. (1) so is calculated independently from r_{2t}. What is being plotted at the right in Fig. 8 is r_{2t} versus $\sigma_p^{2/3} r_g^{1/3}$.

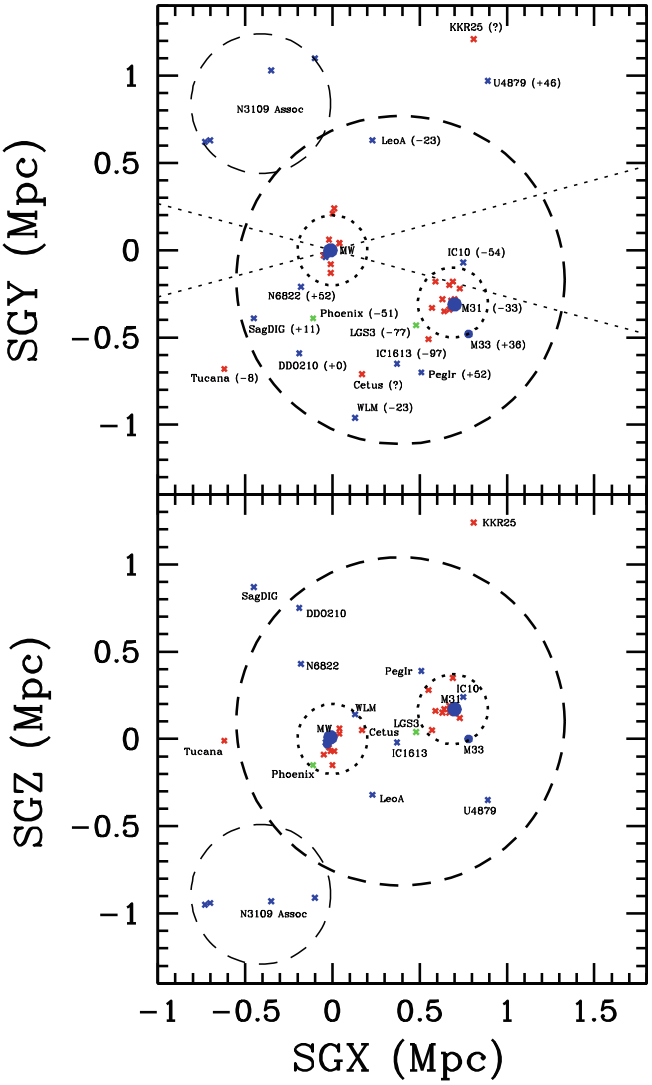


Fig. 5 *Left:* Two projections of the distribution of galaxies around the Local Group. *Blue*: late types; *red*: early types; *green*: transition. The *slanted dotted lines* are projections of galactic latitude ±15°. The *inner dotted circles* approximate the radii of second turnaround for the M31 and Milky Way halos. The *outer heavy dashed circle* approximates the first turnaround radius for the combined halos. A *light dashed circle* encloses the separate NGC 3109 association. *Right:* Velocities as a function of distance from the nearer of M31 or the Milky Way. Inside the caustic of second turnaround at roughly 270 kpc around each of the major galaxies, the velocity dispersion is large and galaxies generally have early types. Outside second turnaround most galaxies are late types and there is a pattern of infall with the first turnaround located at 940 kpc

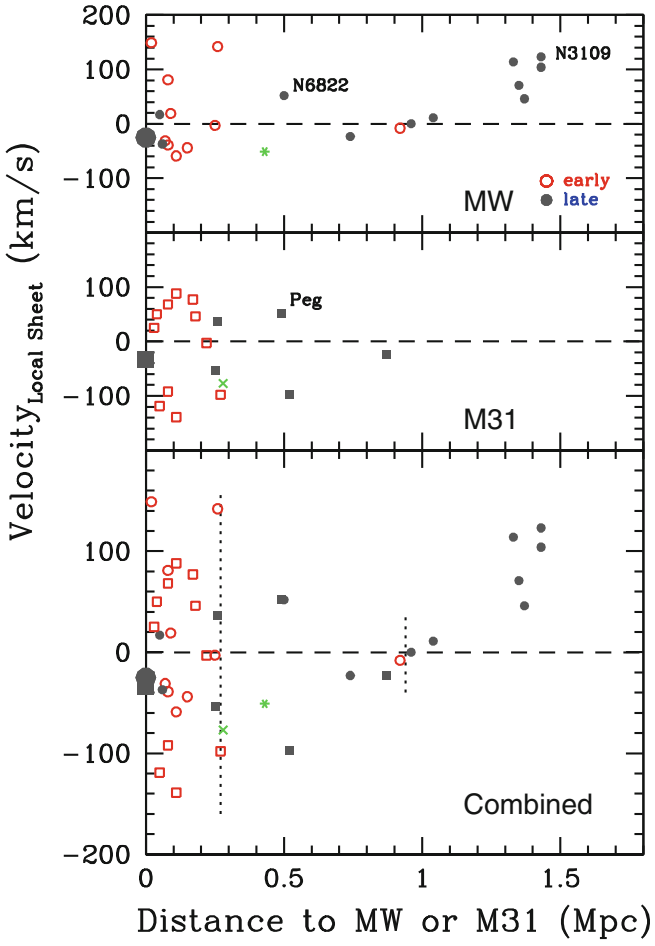


Fig. 5 (continued)

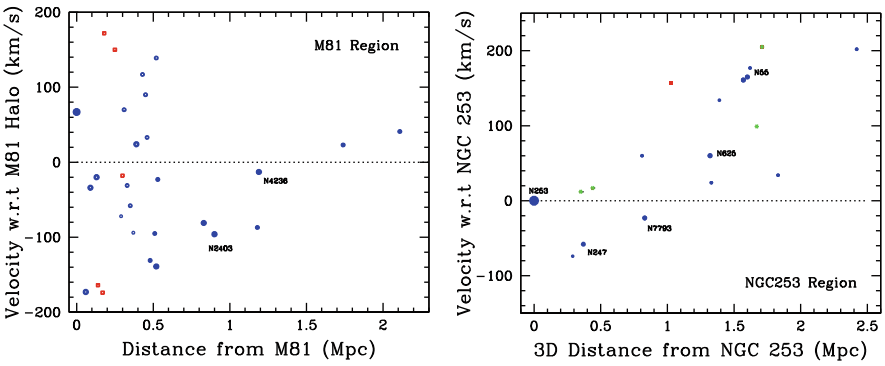


Fig. 6 Infall patterns onto M81 *(Left)* and NGC 253 *(Right)*. With M81 the differentiation between the region of high-velocity dispersion within the second turnaround and the region of infall is clear. The situation is less clear in the case of the smaller halo containing NGC 253

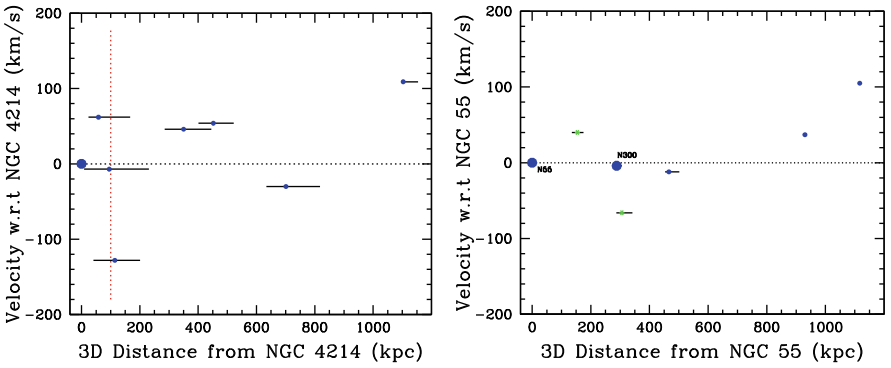


Fig. 7 Velocities for galaxies in two associations of dwarfs with halo masses below 10^{12} $M_\odot$

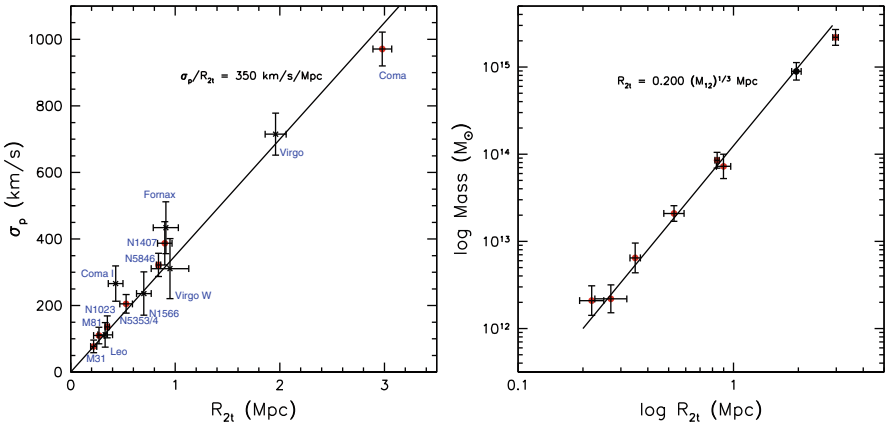


Fig. 8 *Left:* Correlation between projected second turnaround radius and line-of-sight velocity dispersion. *Right:* Correlation between projected second turnaround radius and halo virial mass

The derived correlation is

$$r_{2t} = 0.200\, M_{12}^{1/3}\ \mathrm{Mpc} \tag{10}$$

if the virial mass is in units of 10^{12} $M_\odot$. The tight correlation implies a close correlation between r_{2t} and r_g. It follows that the correlation seen at the left in Fig. 9 would exist. In this plot the slope is fixed and the zero-point is determined by the relations seen in the two panels of Fig. 8. The mass–dispersion relation for groups obeys the law

$$M_\mathrm{v} = 2.5 \times 10^6\, \sigma_p^3 \tag{11}$$

where mass M_v is in solar units and velocity dispersion σ_p is in km s^{-1}.

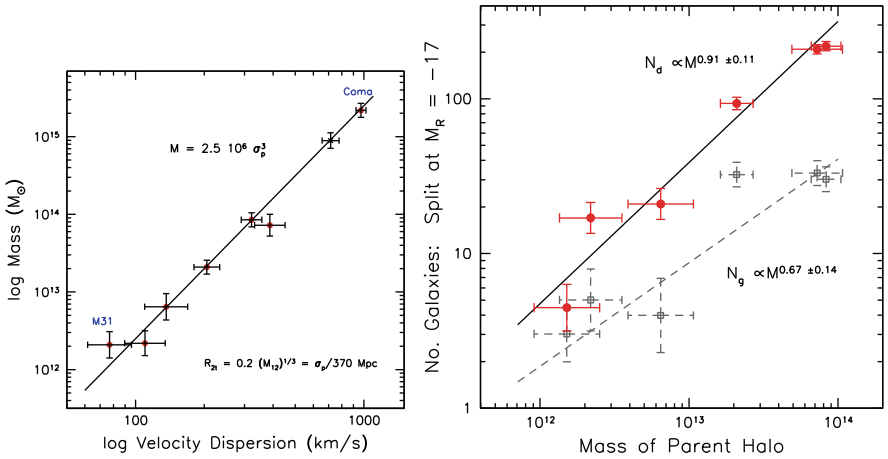


Fig. 9 *Left:* Correlation between the halo virial mass and the line-of-sight velocity dispersion. *Right:* Number of galaxies as a function of the mass of the parent halo, separated between dwarfs (*red points and solid line fit*) and giants (*grey points and dashed fit*). The separation between giants and dwarfs is made at $M_R = -17$

4 The Radius of First Turnaround

The best opportunities to measure r_{1t}, the radius of first turnaround, are presented by the groups within 4 Mpc including the Local Group. An initial discussion by Sandage [20] was carried forward most recently by Karachentsev et al. [12]. The situation in the Local Group is shown in Fig. 5. The value of $r_{1t} = 940$ kpc enclosing a mass of 2×10^{12} $M_\odot$ seems reasonably constrained to an uncertainty of 10%. However, spherical symmetry is a poor approximation for the zero velocity surface of the Local Group because of the dumbbell distribution of mass between M31 and the Milky Way. Moreover, only five galaxies are well placed to define the zero-velocity surface.

A second relatively clean opportunity to measure the surface of first turnaround is presented by the pattern of infall around M81 seen in Fig. 6. The zero-velocity surface lies at 1.3 Mpc from the center of a 4×10^{12} $M_\odot$ halo.

This limited information nonetheless gives us a preliminary calibration of the relationship between r_{1t} and the mass within this radius:

$$r_{1t} = 0.78\, M_{12}^{1/3} \text{ Mpc} \tag{12}$$

An interesting potential confirmation of the validity of this relation could come from the definition of the zero-velocity surface around the Virgo Cluster. It is anticipated to lie at $r_{1t} \sim 7$ Mpc, a value in accordance with long-standing models of Virgo infall [25]. A dense grid of galaxy distances across the Local Supercluster is needed to resolve the issue.

The relationship between a specific radius like r_{2t} or r_{1t} and the mass internal to the specific radius is virtually independent of the dark energy content of the Universe. However, because of the relative differences with respect to the age of the universe between the onset of collapse that is associated with r_{2t} and the current age associated with r_{1t}, the ratio r_{1t}/r_{2t} has a dependence on the dark energy content.

Currently the best estimate of the ratio of these radii is based on observations of the M81 halo. From the data used to generate the left panel of Fig. 6 a tentative ratio $r_{1t}/r_{2t} = 3.5 \pm 0.4$ is found where the error estimate includes random but not systematic effects. At present the errors encompass the full range of reasonable cosmological models.

Another quantity of potential interest is the surface of zero gravity, the surface around an overdense region that is at the limit of what is destined to ever collapse onto the overdensity. This surface is well specified as a function of cosmological model in the case of spherical collapse. With $\Omega_\Lambda = 0.7$ in a flat universe then $r_{ZG} \sim 1.4 r_{1t}$ [18]. However, given the close proximity of separate collapsed halos, often with more than one halo within an infall region and separate infall regions close to overlapping, the assumption of spherical collapse is generally a bad approximation. It must be a very bad approximation with r_{ZG}.

5 Parent Halos and Their Children

Groups of both the early and the late types were studied during the CFHT Megacam imaging campaign. The completeness limit for membership studies was typically $M_R = -11$ or fainter. In an effort to study the properties and possible variations in the luminosity function, galaxies were separated into two bins: "giants" with $M_R < -17$ and "dwarfs" with $-17 < M_R < -11$. Small but significant variations in luminosity functions were found. The variations could be characterized by variations in the ratio of dwarfs to giants. The groups of predominantly early types have larger dwarf/giant ratios. An interesting result is found if the number of giants or dwarfs is plotted against the group virial mass. The correlation is poor with giants but very pronounced with dwarfs. The result is shown in Fig. 9 (right panel). The straight line describes the relation

$$N_d = 5.2\, M_{12}^{0.91 \pm 0.11} \tag{13}$$

where N_d is the number of dwarfs with $-17 < -M_R < -11$ in the group. The slope is consistent within the errors with unity. The number of dwarfs per unit halo mass is roughly constant. A count of the number of dwarfs is a good measure of the mass of a group halo.

It is evident from Figs. 3, 4, and 5 that the morphological separation with environment familiar from studies of rich clusters is also seen in small groups. The caustic of second turnaround is a clear marker of the transition between preponderant morphological types. Secondary galaxies must loose their supplies of cold gas within of order the halo crossing time of

$$t_X = \frac{(\pi/2)R_{2t}}{\sqrt{3}\sigma_p} = 2.5 \times 10^9 \text{ years} \tag{14}$$

6 Summary

Halos can be observationally delineated by a density drop and a transition between dispersed orbits and infall at the radii of second turnaround, r_{2t}. The scaling relations that are theoretically anticipated to exist over a wide range of halo masses are confirmed from observations of nearby collapsed regions. Specifically

- Over the mass range from the halo of M31 to the Coma Cluster there is a correlation between projected second turnaround radius R_{2t} and line-of-sight velocity dispersion σ_p: $\sigma_p/R_{2t} = 350 \text{ km s}^{-1} \text{ Mpc}^{-1}$.
- Over the same mass range, there is the correlation between R_{2t} and the virial mass M_{12} measured in units of $10^{12}\,\mathrm{M}_\odot$: $R_{2t} = 0.200(M_{12})^{1/3}$ Mpc.
- These two correlations can be combined to give the alternative, but not independent, relation: $M_\mathrm{v}/\mathrm{M}_\odot = 2.5 \times 10^6 \sigma_p^3$.

Dwarf galaxies serve to identify collapsed halos because they are numerous. Dwarfs and giants alike that are located within the radii of second turnaround tend to be gas-poor "early" types while galaxies of all sizes outside the second turnaround tend to be gas-rich "late" types. The transition between predominantly early and late populations is another indicator of the radius of second turnaround.

Dwarf galaxies delineate collapsed halos in a way that is surprisingly quantitative. If the number of dwarf galaxies with $-17 < M_R < -11$ are counted, then the number of these dwarfs in a halo depends linearly on the mass of the halo. Count the number of dwarfs and one has a measure of the mass of the halo.

There are differences in the luminosity function of galaxies with environment [22]. More dynamically evolved environments have somewhat steeper faint end slopes. However, it appears that the reason is not more dwarfs per unit mass in evolved environments, but fewer intermediate luminosity systems in the vicinity of $L^\star$, the luminosity that characterizes the exponential cutoff from the faint end power law distribution. It is suspected that intermediate luminosity systems are being lost through mergers with the central dominant galaxy. At faint luminosities, the production and depletion mechanisms are such that the number of dwarfs per unit halo mass remains roughly constant.

Acknowledgments My collaborators in this enterprise are Kristin Chiboucas, Hélène Courtois, Brad Jacobs, Igor Karachentsev, Andisheh Mahdavi, Luca Rizzi, Ed Shaya, and Neil Trentham. A major wide-field imaging program has been carried out with the Canada–France–Hawaii Telescope. Spectroscopic follow-up has been carried out with the Subaru and Keck telescopes. Another major part of the program has involved observations with Hubble Space Telescope. Support has been provided by the Space Telescope Science Institute in connection with the satellite observations and by the National Science Foundation through the grants AST0307706 and NSF0908846.

References

1. Bertschinger, E. 1985, ApJS, 58, 39
2. Binggeli, B., Tammann, G.A., Sandage, A. 1987, AJ, 94, 251
3. Boylan-Kolchin, M., Springel, V., White, S.D.M., Jenkins, A., Lemson, G. 2009, MNRAS, 398, 1150
4. Chernin, A.D., Teerikorpi, P., Valtonen, M.J., Dolgachev, V.P., Domozhilova, L.M., Byrd, G.G. 2009, A&A, 507, 1271
5. Diaferio, A., Geller, M.J. 1997, ApJ, 481, 633
6. Gunn, J.E., Gott, J.R.I. 1972, ApJ, 176, 1
7. Jacobs, B.A., Rizzi, L., Tully, R.B., Shaya, E.J., Makarov, D.I., Makarova, L. 2009, AJ, 138, 332
8. Jerjen, H., Freeman, K.C., Binggeli, B. 1998, AJ, 116, 2873
9. Karachentsev, I.D., Dolphin, A.E., Geisler, D., Grebel, E.K., Guhathakurta, P., Hodge, P.W., Karachentseva, V.E., Sarajedini, A., Seitzer, P., Sharina, M.E. 2002, A&A, 383, 125
10. Karachentsev, I.D., Grebel, E.K., Sharina, M.E., Dolphin, A.E., Geisler, D., Guhathakurta, P., Hodge, P.W., Karachentseva, V.E., Sarajedini, A., Seitzer, P. 2003, A&A, 404, 93
11. Karachentsev, I.D., Karachentseva, V.E., Huchtmeier, W.K., Makarov, D.I. 2004, AJ, 127, 2031
12. Karachentsev, I.D., Kashibadze, O.G., Makarov, D.I., Tully, R.B. 2009, MNRAS, 393, 1265
13. Karachentsev, I.D., Sharina, M.E., Dolphin, A.E., Grebel, E.K., Geisler, D., Guhathakurta, P., Hodge, P.W., Karachentseva, V.E., Sarajedini, A., Seitzer, P. 2002, A&A, 385, 21
14. Karachentsev, I.D., Tully, R.B., Dolphin, A., Sharina, M., Makarova, L., Makarov, D., Sakai, S., Shaya, E.J., Kashibadze, O.G., Karachentseva, V., Rizzi, L. 2007, AJ, 133, 504
15. Lynden-Bell, D. 1967, MNRAS, 136, 101
16. Mahdavi, A., Geller, M.J., Böhringer, H., Kurtz, M.J., Ramella, M. 1999, ApJ, 518, 69
17. Mahdavi, A., Trentham, N., Tully, R.B. 2005, AJ, 130, 1502
18. Peirani, S., de Freitas Pacheco, J.A. 2008, A&A, 488, 845
19. Rines, K., Geller, M.J., Kurtz, M.J., Diaferio, A. 2003, AJ, 126, 2152
20. Sandage, A. 1986, ApJ, 307, 1
21. Shandarin, S.F., Zeldovich, Y.B. 1989, RvMP, 61, 185
22. Trentham, N., Tully, R.B. 2009, MNRAS, 398, 722
23. Trentham, N., Tully, R.B., Mahdavi, A. 2006, MNRAS, 369, 1375
24. Tully, R.B., Rizzi, L., Dolphin, A.E., Karachentsev, I.D., Karachentseva, V.E., Makarov, D.I., Makarova, L., Sakai, S., Shaya, E.J. 2006, AJ, 132, 729
25. Tully, R.B., Shaya, E.J. 1984, ApJ, 281, 31
26. Tully, R.B., Trentham, N. 2008, AJ, 135, 1488
27. Vogelsberger, M., White, S.D.M., Mohayaee, R., Springel, V. 2009, MNRAS, 400, 2174

Dark Haloes as Seen with Gravitational Lensing

Konrad Kuijken

Abstract Dark matter is an important ingredient of galaxies, as was recognized early on by Ken Freeman himself! Evidence for dark matter halos is still indirect, based on analysing motions of tracers such as gas and stars. In a sense the visible galaxy is the mask through which we can study the dark matter. Light rays are also sensitive to gravitational fields and dark haloes cause observable gravitational lensing effects. There are three regimes: microlensing (which probes the clumpiness of dark matter haloes), strong lensing (sensitive to the inner mass distribution) and weak lensing (which can probe haloes out to 100 s of kpc from the centre). This review will concentrate on weak lensing and describe a new survey, the Kilo-Degree Survey (KiDS), that is designed to study galaxy halo masses, extents, and shapes as a function of environment, galaxy type, and redshift.

1 Ken Freeman's Dark Side

Dark matter haloes are an important ingredient of galaxies and indeed of the universe as a whole. They dominate the mass of material in galaxies and drive the gravitational instability that leads to galaxies condensing and clustering in the first place.

The most convincing indications for dark matter haloes in galaxies are still from rotation curves of extended HI disks. In fact, the first mention of extra mass in the literature appears to be in [7], where exponential galaxy disks are described, their gravitational potentials are calculated analytically, and where (in an appendix!) the calculations are compared with then available HI rotation curves of four bulgeless galaxies. For NGC 300, Ken prophetically wrote, without further comment, *"If the HI rotation curve is correct, then there must be undetected matter beyond the optical*

K. Kuijken (✉)
Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands
e-mail: kuijken@strw.leidenuniv.nl

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_31,

extent of NGC 300; its mass must be at least of the same order as the mass of the detected galaxy".

Further influential work on dark matter haloes includes the work with van der Kruit (described elsewhere in this volume) on understanding the internal dynamics of stellar disks and hence their contribution to the total mass budget of the galaxy (are disks' contributions to the rotation curve "maximal" or not?), which included some of the first deep spectroscopy and integrated light velocity dispersion measurements of galaxy disks ever attempted. See van der Kruit's chapter for more details.

A final contribution to be highlighted here is Ken's part in the MACHO project [1], which used the (now sadly burnt down) Mt. Stromlo 50-in. telescope to monitor millions of stars in the Large Magellanic Cloud for many years and discovered tens of gravitational microlensing events. This experiment probes the clumpiness of the matter distribution along the line of sight to the LMC and the low number of events seen by MACHO (and by the similar EROS experiment [2]) proved that the galactic dark halo does not consist predominantly of compact dark objects. Any substantial population of dark objects more massive than Earth is ruled out.

2 Gravitational Lensing and Dark Matter

Even though dynamical tracers were the first to provide evidence for dark matter, photons are a very useful alternative. Because of their speed they are only deflected slightly as they pass a massive galaxy 'lens', but the effect is large enough to be measurable and simple enough to be free from some of the uncertainties inherent in dynamical tracer studies. A brief comparison of dynamics and lensing as gravitational potential tracers is given below.

Dynamics	**Lensing**
Sensitive to the 3D potential of the lens	Sensitive to the 2D, projected potential on the sky
Requires assumption of equilibrium	Sensitive to the instantaneous mass distribution
Orbit structure unknown	Mass-sheet degeneracy
Strong lensing: few data points	Good S/N in baryon-dominated parts
Weak lensing at large radii requires stacking	Few good tracers at large radii

Both techniques have their fundamental limitations. For dynamics this is the need to understand the orbit structure of the tracer: the most famous example is the radial orbit anisotropy degeneracy: the same dispersion profile in a spherical galaxy can be explained with radial orbit anisotropy or with a rising mass-to-light ratio [4]. For stellar tracers disentangling this degeneracy can be complicated requiring higher order measurements of the shape of the velocity distribution. In the case of lensing,

the main uncertainty is the mass-sheet degeneracy, which is a transformation of the lens that leaves the observed image positions intact but changes the (unobservable) source plane scale. The effect on the lens model is to trade some of the lens mass for a constant mass sheet, in such a way that the projected mass enclosed in the Einstein radius remains unchanged. The mass-sheet degeneracy makes it hard to measure the slope of the projected density of a lens accurately.

Lensing manifests itself in three different settings: microlensing, weak lensing, and strong lensing. Microlensing is the twinkling that occurs when a source is viewed through a screen of effectively point-mass lenses. Occasionally a lens will pass very close to the line of sight to a source, and temporarily magnify it with a characteristic lightcurve. Observing the statistics and duration of these lightcurves allows the clumpiness of the foreground mass distribution to be studied. Strong lensing occurs when the lightrays are deflected sufficiently for several images of the same source to be observed. These special geometries allow some of the best mass determinations in cosmology to be made. Finally, weak lensing covers the other cases: a weak deflection that is not sufficient to cause multiple images, but can be detected statistically from the correlated distorting effect it has on the shapes of the sources.

3 Microlensing

As noted above, the MACHO and EROS experiments established that the galactic dark halo is not predominantly composed of massive objects, at least along the sightline to the Magellanic clouds.

A later study of M31, modestly called the 'MEGA' project, found similar results [5]. In ground-based images of M31 the stars are terribly blended, so that variables can only be found with a different technique, difference image photometry. This involves subtracting exposures of the same field after carefully correcting for seeing and extinction differences and it has the effect of turning a crowded field of stars into an uncrowded field of difference flux. The main difference in the analysis of difference image time series is that the baseline flux of the star in its unlensed state is lost, but it turns out that it is still possible to constrain the lens mass spectrum from such data. In fact, M31 offers several advantages over LMC studies because the entire M31 halo can be mapped with a single experiment. In particular, since M31 is inclined to the plane of the sky, sight lines to the near and far sides of the disk cross different path lengths of M31's halo. This implies that the two sides of the disk should see different microlensing rates, making it possible to distinguish disk self-lensing (which should be the same on both sides) from halo microlensing. The asymmetric lensing signal also affords (in principle) a constraint on the shape of the halo.

Results from a 4-year campaign using the Isaac Newton Telescope on La Palma, using about 200 nights of data, revealed 14 microlensing events over the face of M31's disk. Analysis is complicated because much of M31's disk is extincted to

varying degree and because the two fields monitored did not cover the full galaxy. A full Monte-Carlo simulation of the experiment concluded that the number of events, and the asymmetry in their distribution between both sides of the disk, are fully consistent with the expectation for self-lensing of/by disk and bulge stars (Fig. 1).

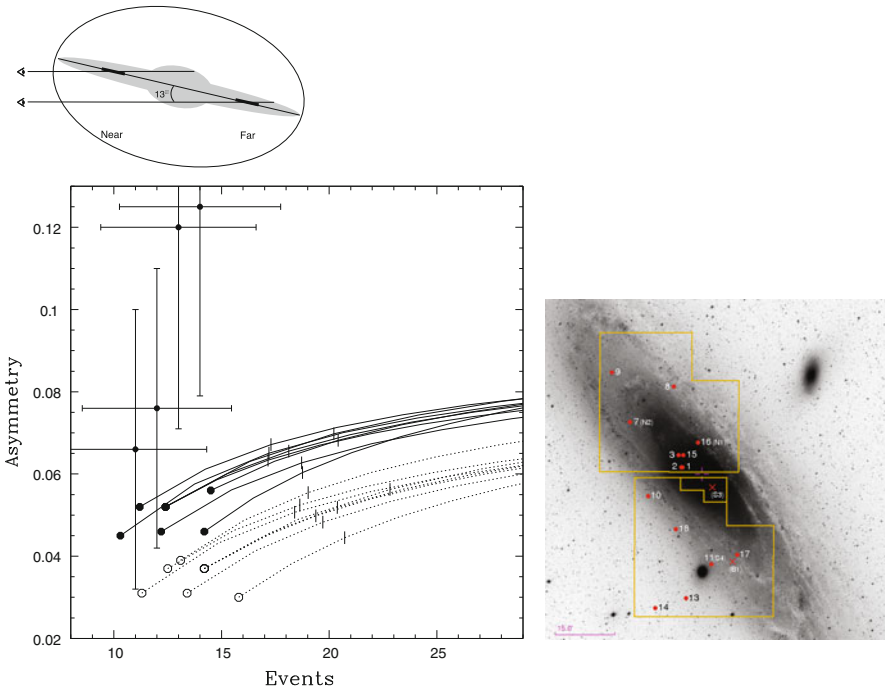


Fig. 1 *Top left*: illustration of the asymmetry to be expected in the M31 halo microlensing rates between the near and the far sides of the stellar disk. *Right*: field layout with detected events. *Bottom left*: Model predictions for number of events, and asymmetry in their distribution (*curves*) and measurements employing various cuts (*data points*). Different curves make different assumptions on extinction correction. From [5]

4 Strong Lensing

Many of the most spectacular astronomical images are caused by strong gravitational lensing, which splits the light from a distant source into multiple images. It turns out that these images can provide some of the best-measured masses in astronomy.

The angle through which a gravitational lens at distance[1] D_l moves the image of a source at distance D_s on the sky is determined by its surface mass density, scaled by the 'critical surface mass density' Σ_{cr} given by

[1] All distances are defined as angular diameter distances.

$$\Sigma_{cr} = \frac{c^2}{4\pi G} \frac{D_s}{D_l D_{ls}} \tag{1}$$

where D_{ls} is the distance of the source from the lens. In the ideal case in which the source lies directly behind the lens, and the lens has a circularly symmetric mass distribution on the sky, the source will be distorted into an 'Einstein Ring', inside which the average surface density is exactly equal to Σ_{cr}. Even for not-quite axisymmetric lenses, or not-quite perfect alignment, it is possible to estimate the Einstein radius of the lens and hence anchor the mass distribution of the lens. Typical values for the Einstein radius are an arcsecond, which corresponds roughly to an effective radius of the lens. In other words, strong lenses probe the mass distribution in the visible central parts of massive galaxies.

The most comprehensive study of the mass distribution of galaxies from strong lensing has been undertaken by the SLACS team [11]. They identified candidate strong lenses from the SDSS spectra, by looking for early-type galaxies with incompatible emission lines in the spectra that could be due to amplified background emission-line galaxies. Subsequent HST imaging revealed a very high success rate of the technique and resulted in a sample of dozens of new lenses. Unlike the existing lensed-quasar samples, many of these lenses have resolved sources so that the mass distribution of the lens can be probed at many positions. A complete modelling effort, including stellar dynamics measurements of many of the lenses, resulted in the surprising conclusion that a simple singular isothermal ellipsoid mass model is able to account for all the data.

5 Weak Lensing

Strong lensing probes the central, high surface mass density regions of galaxies. At larger radii gravitational lensing still occurs, but not at the strength required to split images. Instead the effect is to distort the background sources (since different parts of a source are deflected by different amounts). The leading-order effect of this distortion is expressed as a two-dimensional symmetric matrix that describes the mapping from image to (unlensed) source coordinates as a *shear* γ and a *convergence* κ:

$$\begin{pmatrix} x_s \\ y_s \end{pmatrix} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \equiv (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} \tag{2}$$

Here $g \equiv \gamma/(1-\kappa)$ is called the *reduced shear* and it is related to the axis ratio of b/a of the ellipse into which a circular source is sheared, via $g = (a-b)/(a+b)$. It is the only quantity that can be measured from galaxy shapes, a fact that is related to the mass-sheet degeneracy.

The effect of a massive lens can be understood qualitatively as follows. Gravitational lensing acts to push images on the sky away from the lens, a consequence of

the ‘detour’ that the light has to make around the lens. This distorts the background sky, effectively squeezing the sky around a lens radially. The result is a pattern of tangentially distorted background sources around every foreground mass concentration.

Weak lensing consists of extracting the shear field on the sky from measurements of galaxy shapes, essentially by trying to derive the shape of the ‘average’ galaxy, which is intrinsically round. This subject is technically complex, due to the many distorting effects in astronomical images that have nothing to do with gravitational lensing, but can be much stronger: smearing by the point spread function, distortion by the camera optics, pixelation of the image on the detector, charge transfer inefficiency, etc. The central issue is the need to average over large numbers of background galaxies in order to extract the shear against the intrinsic variety of galaxy shapes, the so-called shape noise.

Weak lensing is not only an effective way to study the mass distribution around galaxies and in clusters, it has also been recognized as a powerful cosmological probe. Lensing strength is sensitive to the angular–diameter distances of lens and source and in combination with redshift measurements can be used to reconstruct the angular diameter distance–redshift relation, i.e. the expansion history of the universe. The statistical power of a large, three-D weak lensing map of the sky is sufficient to yield percent-level measurements of dark energy equation of state parameters and is currently the subject of several design studies for space missions.

Recently we have demonstrated that it is possible to control the systematic errors to below the statistical errors, for what is currently the largest space-based weak lensing map, the COSMOS field [14]. Combining photometric redshifts with weak lensing measurements, we were able to show that the strength of the gravitational lensing signal increases with source redshift as expected in the standard cosmological model, and could even add independent, lensing-based confidence intervals on the acceleration parameter. Weak lensing is fast becoming established as a reliable probe for galactic structure and cosmology (Fig. 2).

Individual galaxy lenses have too few nearby background galaxies around them to yield useful mass measurements at large radii: the shape noise simply overwhelms any shear signal. A clear shear measurement can only by achieved by stacking large numbers of similar galaxies. Fortunately, modern wide-field imagers are capable of the necessary observations.

An impressive demonstration of the power of numbers is provided by the weak lensing analyses of the SDSS image survey. Already the early data release showed a convincing signal [6], but the more recent results have refined the result to the point where the average mass distribution can be studied as a function of galaxy stellar mass (derived from SDSS spectroscopy and SED fitting), luminosity, Hubble type and environmental density [13]. Tentative results on the shapes of galaxy haloes have also been reported [10, 12]. A separate study [8], using HST images, of the SLACS lenses discussed above has also been done and shows that the singular isothermal lens model continues to fit this population well out to radii of several hundreds kiloparsec.

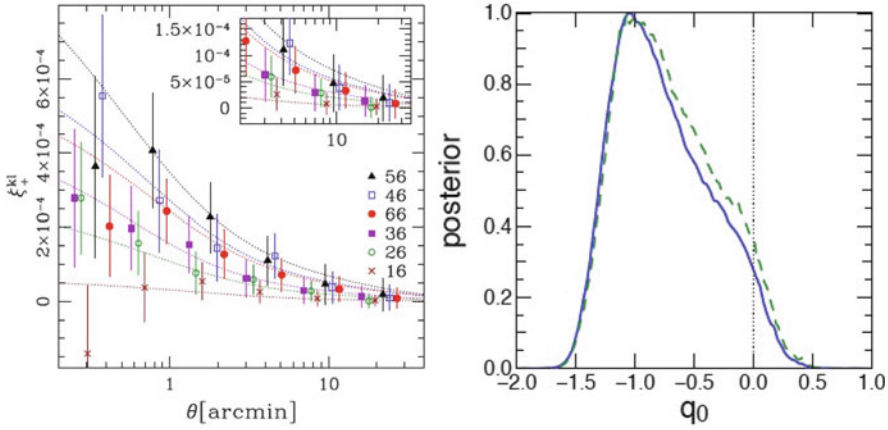


Fig. 2 *Left*: cross-correlation functions (*symbols*) of the shear field in the COSMOS field, for different redshift slices. The amplitudes of the cross-correlation functions depend only on the angular-diameter distances to these redshift slices and they scale as expected (*solid lines*). *Right*: Confidence intervals on the acceleration parameter, showing that these lensing data strongly favour an accelerating universe. From [14]

5.1 Future Directions

Weak lensing focuses on the most visible effect of lensing, the shear, but these are not the only measures of lensing strength. Also the magnification, and the flexion, are showing promise as independent measures.

Magnification is in principle not possible to measure for an individual source whose intrinsic brightness is unknown. (In strong lensing the flux ratio of two images of the same source provides some of this information, but in the weak lensing regime there is only one image). However, the brightness distribution of a population of sources can be measured, as can its change due to lensing. Magnifying a piece of sky containing a population of sources with a steep luminosity function will boost the individual source fluxes, but will also dilute the sky density by the same factor. For a sufficiently steep LF the boost will win and the net result will be an increase in the number of sources above a given flux limit. For a flatter luminosity function, on the other hand, the dilution effect dominates and the number of sources will actually decrease as a result of the magnification (Fig. 3).

This effect is difficult to apply to individual cluster lenses because it can be affected by clustering of the background source populations. But it has recently been demonstrated to work very nicely on foreground galaxy lenses in the deep fields of the Canada–France–Hawaii Legacy Survey [9], using u, g, and r dropout galaxies as the sources: brighter dropouts tend to show a positive correlation with the foreground sources, whereas the fainter ones (which have a flatter luminosity function) show an anticorrelation. Also qualitatively the strength of the signal is as expected (Fig. 4).

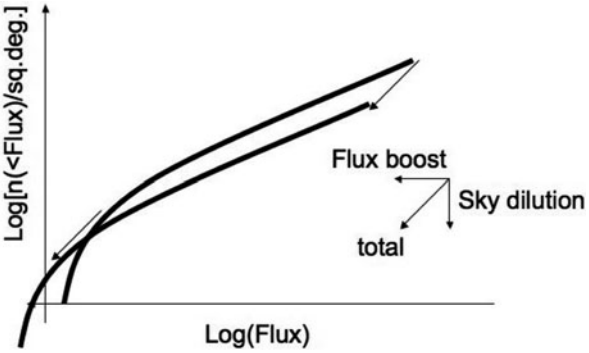


Fig. 3 Illustration of the magnification effect on the luminosity function of a population of sources. Depending on the slope of the luminosity function the numbers of observed sources per unit area may increase or decrease

A second effect is *flexion* [3], the next-to-leading order distortion after shear, due to the fact that the shear will be different in different parts of a source. Flexion gives rise to the arc-like shapes of lensed galaxies, for example. While this is a weaker effect than shear and it drops off faster with distance to the lens, it is a valuable addition of information because it probes the inner regions around lenses where there are not many background sources. Flexion also has the advantage that, unlike shear, it distorts galaxies into 'unnatural' shapes, so that the shape noise is considerably lower than it is for shear.

6 The KiDS project

After many years of waiting, we are finally looking forward to the VLT Survey Telescope on Paranal being completed. It has been designed with image quality over a wide field in mind and even though it is late to arrive it will still provide a unique capability in the Southern hemisphere. The hope (and design) is that the telescope will deliver natural seeing images over a full square degree field of view, well sampled with the small pixels of the OmegaCAM camera that was built by a consortium of institutes in the Netherlands (NOVA), Germany (Göttingen, Bonn, München, and ESO) and Italy (Naples and Padua).

The system allows an unprecedented weak lensing survey to be carried out and this is the aim of the largest survey, KiDS (the Kilo-Degree Survey), that is planned for the VST as soon as it enters operations (Fig. 5). Together with a companion survey on the near-IR VISTA survey telescope, the project will make images of 1,500 $\deg^2$ of sky, in nine photometric bands from u to K (Table 1). The survey area includes many galaxies with known redshift, ensuring that the foreground mass distribution is mapped in detail and can be 'weighed' with gravitational lensing. Because all observations will be done in queue-scheduling, it will be possible to

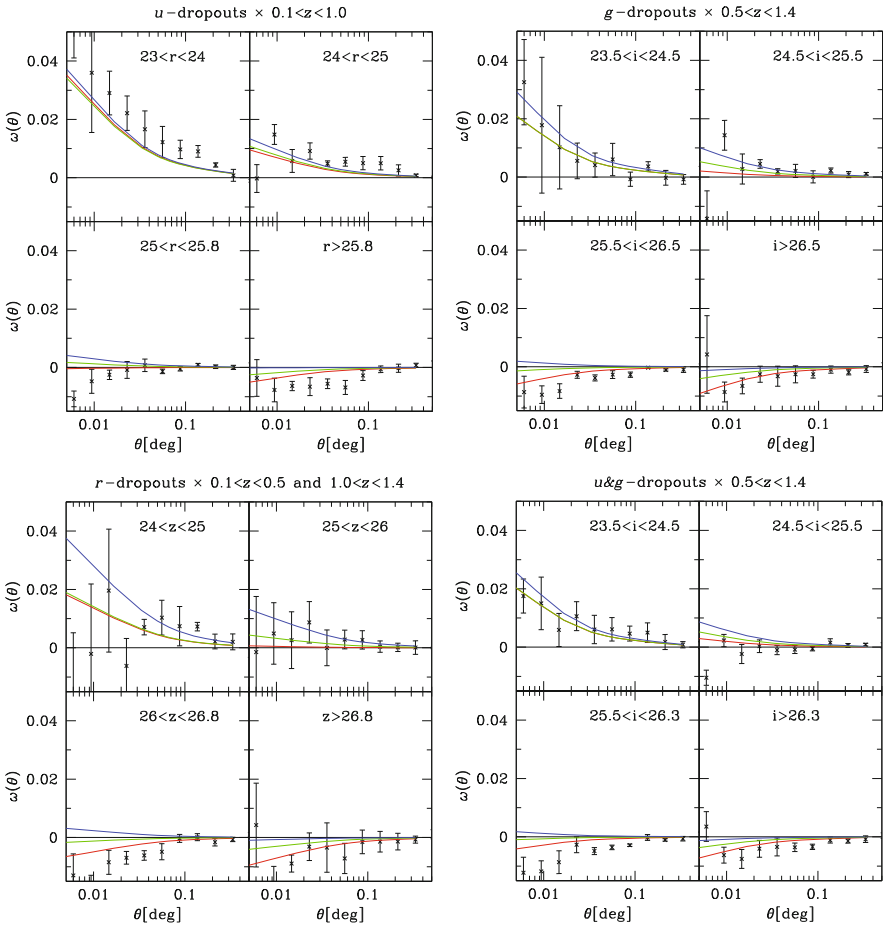


Fig. 4 Demonstration of magnification as a weak lensing measure, from [9]. Each *panel* shows the cross-correlation between foreground lenses and background galaxies identified with a dropout technique. Notice how in each case the brighter sources show a positive correlation, but the fainter ones anticorrelate, as expected from the shape of their luminosity function

optimize the use of the seeing distribution, so that the best seeing dark time will be used for deep r band exposures suitable for weak lensing measurements, with brighter or worse seeing time being devoted to the other photometric bands. As a result the survey should be able to deliver high-fidelity weak lensing *and* photometric redshift measurements over a large part of the sky, with homogeneous data quality. In terms of the usual tradeoff between area and depth, the survey sits between the wide and shallower SDSS images (9,000 deg^2) and the 170-deg^2 CFHTLS. But the unique combination of image quality and 9-band photometry should make it a very valuable resource.

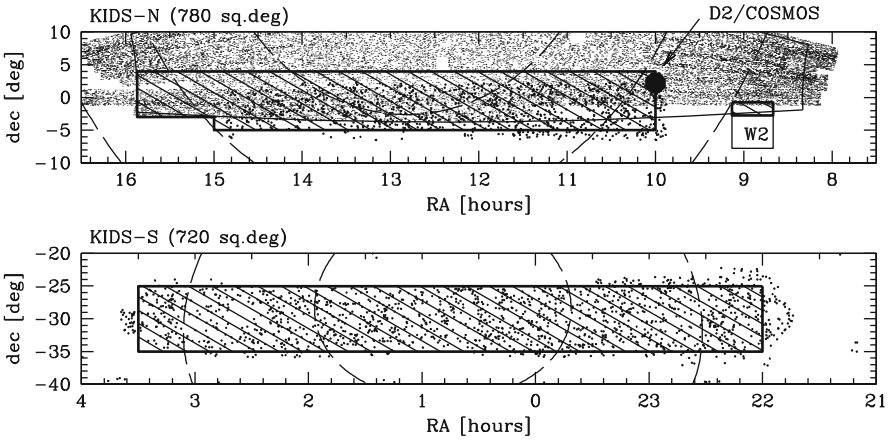


Fig. 5 Field layout of the KiDS survey. The *hashed areas* indicate the survey area to be imaged, 1,500 deg^2 in total. *Small and large dots* delineate the areas where massive spectroscopic surveys of the brighter galaxies have taken place with SDSS and 2dF, respectively

Table 1 Main parameters of the KiDS+VIKING survey

Band	Exp.time	seeing ($''$)	5-σ 2$''$ limit	Telescope
u	900	0.85–1.1	24.8	VST
g	900	0.7–0.85	25.4	VST
r	1, 800	< 0.7	25.2	VST
i	1, 080	0.7–1.1	24.2	VST
Z	500	< 1	23.1	VISTA
Y	400	< 1	22.4	VISTA
J	400	< 1	22.2	VISTA
H	300	< 1	21.6	VISTA
K	500	< 1	21.3	VISTA

Science goals for the survey are broad, but the design centred on using weak lensing to measure the mass distribution in and around galaxies as a function of environment, luminosity, and type, out to a radius of several 100 kpc. This is an interesting regime because it is far beyond the baryon-dominated regions of the galaxy and so can be compared to robust predictions from structure formation theory and simulations. A particularly interesting measurement will be the average flattening of galaxy haloes from edge-on disk galaxies: most alternatives to dark matter should make robust predictions for the variation of the quadrupolar field with radius and these can be tested directly with such data.

The survey will also serve as a pathfinder for yet more ambitious lensing surveys for cosmology, including dark energy experiments from the ground and from space.

7 Conclusions

Galaxies provide one of the best environments for studying the dark matter. In the parlance of this conference, they are the mask that covers the mystery, but that also shows us where to look.

Gravitational lensing is one of the most promising ways of studying the dark matter distribution, particularly at large radii where no classical dynamical tracers are available. New surveys such as KiDS will be able to map the dark matter in and between galaxy haloes with great accuracy, which will enable new tests of structure formation as well as the cosmological model itself.

Acknowledgments It is a pleasure to acknowledge the important role that Ken Freeman has played throughout my career as an astronomer, first as the author of a number of seminal papers, later as a source of occasional excellent advice, and now as a collaborator in the PN.S project, but foremost as a great person to be around. He and David Block organized a fantastic, memorable conference at Sossusvlei.

This chapter benefited from the collaboration and hard work of my students and postdocs, in particular Jelte de Jong, Tim Schrabback, and Hendrik Hildebrandt.

References

1. Alcock, C., et al. 1993, Nature, 365, 621
2. Alcock, C., et al., 1998, ApJ, 499, L9
3. Bacon, D.J., Goldberg, D.M., Rowe, B.T.P., Taylor, A.N. 2006, MNRAS, 365, 414
4. Binney, J.J., Mamon, G.A.1982, MNRAS, 200, 361
5. de Jong, J.T.A., et al. 2006, A&A, 446, 855
6. Fischer, P., et al. 2000, AJ, 120, 1198
7. Freeman, K.C. 1970, ApJ, 160, 811
8. Gavazzi, R., et al. 2007, ApJ, 667, 176
9. Hildebrandt, H., van Waerbeke, L., Erben, T. 2009, A&A, 507, 683
10. Hoekstra, H., Yee, H.K.C., Gladders, M.D. 2004, ApJ, 606, 67
11. Koopmans, L.V.E., Treu, T., Bolton, A.S., Burles, S., Moustakas, L.A. 2006, ApJ, 649, 599
12. Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C.A., Brinkmann, J. 2006, MNRAS, 368, 715
13. Mandelbaum, R., Hirata, C.A., Broderick, T., Seljak, U., Brinkmann, J. 2006, MNRAS, 370, 1008
14. Schrabback, T., et al. 2010, A&A, 516, A63

Behind the Mask: Resolving the Core–Cusp Problem in Spiral Galaxies

George Rhee

Abstract We discuss two aspects of the core–cusp problem in spiral galaxies. The first is observational. In deriving rotation curves from HI data cubes it is essential to take into account projection effects, non-circular motions due to bars and other deviations from axisymmetry, and the effect of pressure gradients on the gas rotation velocity. While these effects may be individually small in themselves, the combined effect is significant as we demonstrate from our analysis of the galaxy NGC 6822. The second issue concerns the effect of gas flows on the survival of density cusps. Winds created by multiple supernova explosions selectively remove low angular momentum baryons from the center of galaxies. Recent simulations show that supernova winds acting throughout the history of a dwarf galaxy can substantially reduce the baryonic and dark matter density in the centers of dwarf galaxies.

1 The Core–Cusp Problem

Cosmological models based on the inflationary paradigm, cold dark matter and dark energy (ΛCDM), have been very successful at explaining the large-scale structure of the universe [17]. Nevertheless, ΛCDM models face a number of challenges on galactic scales [2, 4, 12, 14]. One of these in particular, the core–cusp problem, has persisted and been the subject of much discussion. At issue is the fact that galaxy kinematics seem to be at odds with the structure of cosmological dark matter halos. Cosmological simulations predict that halos have a central density cusp [13] while many galaxies appear to have rotation curves that suggest the presence of a constant density core at their centers. Because dwarf galaxies and low surface brightness galaxies are expected to be dark matter dominated at all radii, they have been at the center of the debate. Much of the controversy concerns the interpretation of the

G. Rhee (✉)
Department of Physics and Astronomy, University of Nevada Las Vegas, 89154 NV, USA
e-mail: grhee@physics.unlv.edu

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_32,

HI and Hα observations of nearby galaxies. We turn our attention in the following section to one of the best studied of these, the nearby dwarf galaxy NGC 6822.

2 NGC 6822

NGC 6822 is a good candidate for dark matter studies. The fact that it is nearby ensures that the central kiloparsec is well resolved in HI observations. The fact that it is small and has low surface brightness suggests that it is dark matter dominated.

NGC 6822 is a local group galaxy, the third nearest dwarf irregular galaxy after the LMC and SMC. It is located at a distance of 0.5 Mpc from the Milky Way. The stellar distribution shows a rather strong bar that is prominent in visible light. The bar has been considered the main stellar component of the galaxy; however, deeper photometry reveals a faint blue component tracing the HI distribution [6]. There is also a faint red stellar component extending to several radial scale lengths. This red component is elongated almost perpendicular to the HI disk and is supported by rotation, leading to the suggestion that NGC 6822 is a polar ring galaxy [8]. The central surface brightness of NGC 6822 is 27 mag/arcsec^2 which is comparable to the surface brightness of large low surface brightness galaxies [6].

We make use of the data published by Weldrake et al. [20]. The data were very kindly made available to us by Professor Erwin de Blok. The data set was obtained with the Australia Telescope Compact Array for 15 $\times$ 12 h in its 375 (1 $\times$ 12 h), 750D (2 $\times$ 12 h), 1.5A (4 $\times$ 12 h), 6A, and 6D (8 $\times$ 12 h) configurations over the period from June 1999 to March 2000. A total of eight pointings were observed covering the entire HI extent of the galaxy. A bandwidth of 4 MHz was used with a channel separation of 0.8 km/s. The data were reduced using the MIRIAD data reduction package. The data cube we examine in this chapter has channel separations of 1.6 km/s and an effective resolution of 1.9 km/s. The synthesized beam size is 42.4 $\times$ 12 arcsec and the pixel size is 4 $\times$ 4 arcsec.

We have computed the rotation curve of this galaxy using the intensity-weighted method and the maximum envelope-tracing method. The intensity-weighted method yields results that are consistent with the results of Weldrake et al. [20]. The envelope-tracing method, as expected, yields velocities that are slightly higher by $\sim$5 km/s. We have made a correction for the gas pressure when calculating the circular velocity from the rotation velocity measurements;

$$V_c^2 = V_{\rm rot}^2 + \frac{KT}{\mu m_H} \frac{\partial \ln \Sigma(r)}{\partial \ln R}$$

We list here only the relevant term (for a full treatment, see [19]). The data are shown in Fig. 1. An NFW model with $M_{\rm vir} = 7.5 \times 10^{10}\,M_\odot$ and a concentration of 11 gives a reasonable fit to the data. The key difference is that we have included a pressure term up to a radius of 2.4 kpc. At this distance there is a sudden drop in the gas density and a break in the spheroid light distribution. This effect may be either due to the effects of tidal interaction or, for the HI, to the ionization of the HI gas. In

either case, the break does not reflect a pressure gradient in the gas and we are not justified in applying this pressure correction. Furthermore, we have found evidence for a bar in the HI data by examining the residuals of the tilted ring fit to the data. The residuals are largest at the location of the optical bar [15].

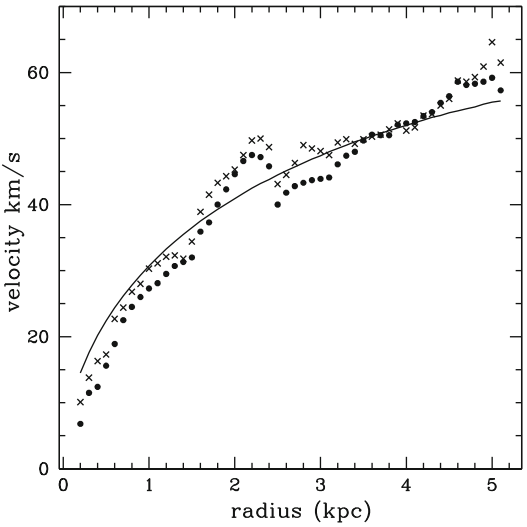


Fig. 1 HI rotation curve for the galaxy NGC 6822. The *dots* show the rotation computed using a tilted ring analysis using intensity-weighted moments. The *crosses* show the rotation computed using the envelope-tracing method. A correction has been made for gas pressure at radii less than 2.4 kpc. The *solid line* is an NFW model that gives a reasonable fit to the data

We conclude that it is indeed possible to reconcile the rotation curve of the galaxy NGC 6822 with an NFW profile when the appropriate corrections are applied to the rotation velocities.

3 Dwarf Galaxy Models with Supernova Winds

We have analyzed recent simulations [5] that suggest that cusps are not an inevitable consequence of galaxy formation. These simulations take into account baryonic processes such as gas cooling, heating from the cosmic ultraviolet field, star formation, and supernova-driven gas heating. The resolution is such that dense gas clumps as small as $10^5\,M_\odot$ are resolved. We show results for a single dwarf galaxy (Fig. 2), but we obtain equivalent results with galaxies of similar mass and different assembly histories and halo spin. We also address the related problem of the absence of bulges in dwarf galaxies.

In the ΛCDM models galaxy disks form as gas cools and collapses inside spinning halos of collisionless dark matter. Models that assume that the stellar component of galaxies inherits the angular momentum distribution of their host dark matter

halos also predict the formation of a centrally concentrated stellar bulge and a cuspy dark matter profile. However, the vast majority of dwarf galaxies have no stellar bulges and the observed rotation curves rise linearly in the central kiloparsec which may be a sign of a shallow dark matter distribution [16, 18]. We propose to explain the existence of these bulgeless galaxies using gas winds. These winds, created by multiple supernova (SN) explosions, selectively remove low angular momentum baryons from the center of galaxies [3, 22]. Supernova winds are observed in both local and high redshift galaxies and are efficient at removing gas from the disks of nearby galaxies at a rate of few times the current star formation rate [9, 21].

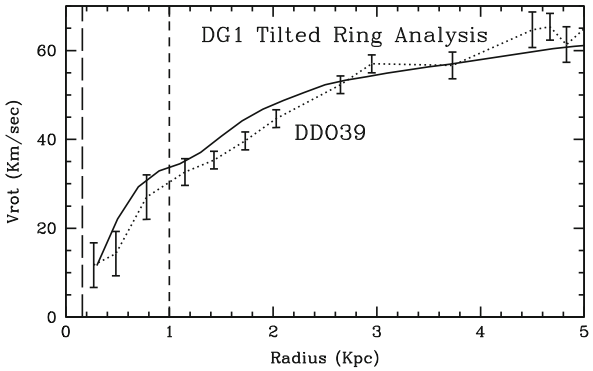


Fig. 2 Rotation curve of the simulated dwarf compared to that measured for a real galaxy. The *continuous black line* shows the rotation velocity of the galaxy using actual projected velocity field and the tilted ring analysis. The *dotted line* shows the same for the galaxy DD039. The *long dashed line* shows the resolution of the simulation and the *short dashed line* shows the scale of the core [5]

These supernova winds have not until now been included in cosmological hydrodynamical simulations of present-day dwarf galaxy systems (Vrot $\sim$ 60 km/s). Dynamical arguments [10, 11] suggest that bulk gas motions (possibly SN induced) and gas clouds orbital energy loss due to dynamical friction can transfer energy to the center of the DM component. Sudden gas removal through outflows then causes the DM distribution to expand.

The model galaxy has a rich merger history: three proto-galaxies of similar mass merge at redshift z $\sim$ 3, and a large satellite is accreted at $z = 1.2$ that has a mass one-third that of the central galaxy (a major merger). Several other satellites are accreted, including one at low redshift. Star formation occurs in bursts as observed in nearby dwarfs [7] peaking at 0.25 $M_\odot$/year during interactions at $z \sim 2$. The disk component assembles shortly after that. At $z = 0$ an exponential stellar disk is surrounded by a neutral hydrogen (HI) disk that extends out to six disk scale lengths. The galaxy shows no sign of a stellar spheroid. The star formation rate declines after $z = 1$, and at present it is down to 0.01 $M_\odot$/year.

A rotation curve was obtained by determining the rotational motion of cold ($T < 10^4$ K) gas as a function of radius using the standard tilted ring method, which is

used on real data and hence includes the effects of non-circular motions induced by disk distortions warping and bars [19]. The rotation curve rises almost linearly out to one stellar disk scale length and is still rising at four scale lengths (4 kpc), replicating the rotation curves of real dwarf galaxies. Previous simulations have persistently produced rotation curves which rise rapidly in the inner regions and peak inside one scale length, reflecting their steep central mass distributions [1]. The dark matter central density has a shallow profile over a core of roughly 1 kpc in size. The dark matter density averaged over the same radius is $10^{7.5}$ $M_{\odot}/kpc^3$, about 50% lower than in a control run where the gas is not allowed to cool or form stars, and where the slope of the inner dark matter profile is as steep as in typical dark matter simulations.

Outflows are mainly responsible for altering the central density profile of the baryonic component. The strongest outflows correlate with star formation bursts, caused by mergers and strong interactions. When dense gas regions form from disk instabilities and sink to the center due to dynamical friction, supernova feedback destroys these gas clumps as soon as they start forming stars. Outflows thus selectively remove most low angular momentum gas before it is transformed into stars, effectively halting the processes that would lead to a concentrated baryon distribution and to the formation of stellar bulges. By the present time the angular momentum distribution of stars formed from the remaining gas has a median value higher than the dark matter and lacks its low angular momentum tail.

The removal of centrally concentrated, low angular momentum gas is also closely connected to the origin of shallow DM profiles. As the galaxy assembles, gas starts collecting at its center via clumps and filaments while the dark matter remains smoothly distributed. This spatial decoupling between the gas and dark matter can lead to efficient orbital energy transfer from the gas to the dark matter through gas bulk motions and gas orbital energy loss. Furthermore, the gas outflows rapidly remove a large fraction of the gas, leading to a significant loss of dark matter binding energy, causing a net expansion and the formation of a shallow profile [10, 11]. These simulations provide direct confirmation of the above two mechanisms, as the expansion of the collisionless dark matter component occurs over several gigayears, closely following the strongest outflows. Outflows occur in both smaller mass progenitors and then at the center of the main galaxy, where the process of core formation is essentially complete by $z \sim 0.5$ by which time the dark matter profile in the inner kiloparsec settles to a shallow slope with $\rho \propto r^{-0.6}$ quite different to that produced in the dark matter only control run that has the canonical, much steeper profile.

Gas outflows generated by powerful SN explosions within an inhomogeneous interstellar medium provide a mechanism to explain the properties of small disk galaxies within the cold dark matter model. This work demonstrates how this mechanism could plausibly result in the production of low-mass, bulgeless disk galaxies with shallow dark matter central profiles. The simulations account for the correlation of these two properties that are observed in samples of nearby galaxies.

Acknowledgments G.R. acknowledges support from the NSF AST-04-07072 grant to the University of Nevada, Las Vegas. The research presented here is the result of collaborations with O. Valenzuela, A. Klypin, F. Governato, C. Brook, L. Mayer, A. Brooks, J. Wadsley, P. Jonsson, B. Willman, G. Stinson, T. Quinn, and P. Madau. We are grateful to E. Brinks for help with the data and to E. de Blok for making the data available to us.

References

1. Abadi, M.G., Navarro, J.F., Steinmetz, M., Eke, V.R. 2003, ApJ, 591, 499
2. de Blok, W.J.G., McGaugh, S.S. 1997, MNRAS, 290, 533
3. Ceverino, D., Klypin, A. 2009, ApJ, 695, 292
4. Flores R., Primack J.R. 1994, ApJ, 427, L1
5. Governato, F., et al. 2010, Nature, 463, 203
6. Komiyama, Y., et al. 2003, ApJ, 590, L17
7. Lee, J.C., Kennicutt, R.C., José, G., Funes, S.J., Sakai, S., Akiyama, S. 2009, ApJ, 692, 1305
8. Letarte, B., Demers, S., Battinelli, P. Kunkel, W.E. 2002, AJ, 123, 832
9. Martin, C.L. 1999, ApJ, 513, 156
10. Mashchenko, S., Couchman, H.M.P., Wadsley, J. 2006, Nature, 442, 539
11. Mo, H.J., Mao, S. 2004, MNRAS, 353, 829
12. Moore, B. 1994, Nature, 370, 629
13. Navarro, J.F., Frenk, C.S., White, S.D.M. 1997, ApJ, 490, 493
14. Ostriker, J.P., Steinhardt, P. 2003, Science, 300, 1909
15. Rhee, G., Valenzuela, O., Klypin, A. 2010, in preparation
16. Simon, J.D., Bolatto, A.D., Leroy, A., Blitz, L., Gates, E.L. 2005, ApJ, 621, 757
17. Spergel, D.N., et al. 2003, ApJS, 148, 175
18. Swaters, R.A., Madore, B.F., van den Bosch, F.C., Balcells, M. 2003, ApJ, 583, 732
19. Valenzuela, O., Rhee, G., Klypin, A., Governato, F., Stinson, G., Quinn, T., Wadsley, J. 2007, ApJ, 657, 773
20. Weldrake, D.T.F., de Blok, W.J.G., Walter, F. 2003, MNRAS, 340, 12
21. Wilman, R.J., Gerssen, J., Bower, R.G., Morris S.L., Bacon, R., de Zeeuw, P.T., Davies, R.L. 2005, Nature, 436, 227
22. Robertson, B., Kravtsov, A. 2008, ApJ, 680, 1083

A GALAXY BASELINE: Multiwavelength Study of a Sample of the Most Isolated Galaxies in the Local Universe

Lourdes Verdes-Montenegro, J. Sulentic, G. Bergond, Daniel Espada, S. Leon, U. Lisenfeld, V. Martinez-Badenes, J.E. Ruiz, J. Sabater, and Simon Verley

Abstract We introduce and describe main results of the AMIGA project. AMIGA involves a panchromatic characterization for many of the most isolated galaxies in the local universe. Many of these galaxies have avoided major interaction events for all, or most, of their lives. Our studies show these galaxies to be the most IR/radio/optically "quiet" sample that is known. AMIGA data are publicly released under a VO interface at http://amiga.iaa.es/ and are also accessible by standard VO tools, as, e.g. TOPCAT.

1 Introduction

The AMIGA project (Analysis of the interstellar Medium in Isolated GAlaxies, http://www. iaa.csic.es/AMIGA.html; [30]) is focused on identifying a statistically significant sample of the most isolated galaxies in the local universe. A major goal is to quantify properties of the interstellar medium in these galaxies and the environmental relationship to star formation (SF) and nuclear activity. The AMIGA results fit naturally within the context of this conference representing the galaxy sample least masked by galaxy evolution. Like the San People of Namibia (the host country of this conference), AMIGA galaxies constitute the best footprints of the ancestors. But why is a sample of isolated galaxies needed? Observational evidence for the interaction enhancement connection is difficult to establish quantitatively. The involved processes and their amplitudes are not well measured or understood: (1) pairs show a star formation excess but no HI deficit [36, 37]; (2) Hickson Compact groups show morphology changes and HI depletion, but no excess SF [3, 14, 32]; (3) dependence of nuclear activity on environment is confused [2, 7, 18, 24]; as well as (4) the enhancement of molecular gas in interacting samples [5, 19, 25, 31]. On the other hand, most definitions of "normal" or "isolated" in the literature are

L. Verdes-Montenegro (✉)
Instituto de Astrofísica de Andalucía – CSIC, Glorieta de la Astronomía s/n, 18008 Granada, Spain
e-mail: lourdes@iaa.es

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_33,

ambiguous, e.g. field samples [17] might include any galaxy not belonging to a cluster, so galaxies in pairs, triplets, and loose/compact groups would not necessarily be excluded. In other cases galaxies without velocity data are not considered companions [16]. When a well-defined isolation criterion is used, the largest samples of isolated galaxies in the literature in most cases involve monochromatic observations of subsamples from the Catalog of Isolated Galaxies (CIG: [15, 30] and references therein). The CIG catalogue constitutes the starting point for the AMIGA project.

2 AMIGA Sample Description and Revisions

We began with the CIG because of its numerous strengths: (1) size: the sample is composed of n = 1,050 galaxies; (2) isolation: no similar-sized galaxies (factor 4) lie within 20 times the diameter of the companion, likely containing many of the most isolated galaxies in the local universe; (3) morphological diversity; (4) depth: the CIG samples a large enough volume of space to allow sampling the majority of the optical luminosity function; (5) completeness: between 80 and 95% brighter than 15.0 mag. Complementarity also extends to existing multiwavelength data for the CIG (e.g. [1, 12, 36]). We are building upon those foundations by refining the basic catalog, utilizing new multiwavelength databases and obtaining new radio and optical measures.

In previous works we revised all of the CIG positions [20] and compiled redshift/distances for more than 90% of the sample. We also performed an optical characterization of the sample [30], but our two largest refinements involved uniform reevaluation of isolation degree and morphology.

2.1 Isolation Revision

In [33, 34] we processed large digitized POSS-I fields for 950 galaxies from the CIG and evaluated their degree of isolation down to $m_B \sim 17.5$ within a projected radius R≥0.5 Mpc. Pairs that remained in Karachentseva's revision were rejected while minor interactions were quantified, calculating either the local surface density or the external tidal force affecting each galaxy. Hence we defined various subsamples of galaxies according to their degree of isolation. The same parameters were calculated for samples of triplets, compact groups, and clusters for comparison.

2.2 Morphology Refinement

We used the digitized POSS2 for a uniform revision of optical morphologies [28]. This allowed us to exploit the high resolution and dynamic range of the photographic

emulsions used for that survey. About 20% of galaxies in the catalog (n = 193) were flagged for the presence of nearby companions or signs of distortion likely due to interaction, and this subsample is treated separately in our subsequent studies. The dominant (2/3) CIG subsample involves late-type Sb-Sc spirals while 14% of the sample is classified as early type and distributed approximately equally between elliptical and lenticular galaxies. We have identified the most nurture-free population of E's which appear to be less luminous even than those found in less restrictive field samples. These are perhaps the best candidate population of primordial or at least minimally nurtured, early-type galaxies.

3 Detailed Morphological Study of Selected Subsamples

More recently [8, 9] presented surface photometric and Fourier analysis of a representative "prototypical" isolated sample of 100 Sb-c galaxies from the vetted CIG ($M_i = -19$ to -23) exploiting SDSS images. This structural analysis showed that most isolated galaxies in our sample (1) likely host pseudobulges rather than classical bulges (bulge/total flux ratios were found to lie B/T$\leq$ 0.2 with Sérsic indices generally $n_{BULGE} \leq 2.5$) and (2) host longer bars, are more symmetric, less concentrated, and less clumpy than less isolated samples such as OSUBGS (Ohio State University Bright Galaxy Survey; [10]) or CSRG (Catalog of Southern Ringed Galaxies [6]).

We obtained Hα and Gunn r photometric data for more than 200 spiral galaxies from AMIGA in order to separate the effects of local environment from internal dynamical processes. The 45 largest and least inclined galaxies were selected for morphological study including the modes of spiral structure as well as bar strength. We also estimated the torques between gas and the bulk of the optical matter using Fourier analysis [35]. We concluded that isolated galaxies do not appear to be preferentially barred or unbarred. Comparison of the Hα distributions with the results of numerical simulations helped to set constraints on the SF law, which is likely to differ from a genuine Schmidt law. The frequently observed phenomenon of star formation avoiding the bar, in spite of large gas density there, suggests that it is probable that the relative velocity of the gas in the bar also needs to be taken into account.

4 FIR Emission, Radio Continuum Emission, and Nuclear Activity

We studied IR emission properties of the AMIGA sample [22] since MIR/FIR emission are regarded as sensitive interaction diagnostics. The four IRAS bands were coadded for 1030 CIG galaxies increasing the archival signal-to-noise ratio by a factor 3–5. The complete sample was used for characterization of the LFIR−LB

correlation as a reference for evaluating IR properties of interacting samples. Galaxies suspected interacting in [28] were excluded. Only 2% of the galaxies showed log(LFIR) > 10.5 $L_{\odot}$. Comparison with 2,445 galaxies from the magnitude limited CfA sample, selected without environmental discrimination, showed a significantly higher mean log(LFIR), <log(LFIR)CfA>=<log(LFIR)AMIGA> +0.26, whereas the mean log(LB) was the same as the AMIGA sample. The galaxies flagged as interacting lie above this correlation showing a systematic enhancement in LFIR.

Radio continuum data at 325, 1,420, and 4,850 MHz were extracted from the WENSS, NVSS/FIRST, and GB6 surveys [21]. Comparison between NVSS and FIRST results indicate that radio continuum emission in AMIGA spirals is disk-dominated and likely driven by low to moderate star formation. In contrast, compact group members show an excess of nuclear radio emission. Furthermore, the derived radio luminosities for detections reveal a very radio-quiet sample, with most detections showing radio/optical luminosity ratios R= 1 − 10.

We also considered the radio–FIR correlation for our sample in order to study nuclear activity in non-interacting galaxies [27]. We searched for radio-excess galaxies which are candidate AGNs and FIR colours to find obscured AGN candidates. There are no radio-excess galaxies in our sample above a factor 5 of radio excess, which is the lowest rate found in comparison with other samples in denser environments. We conclude that the environment plays a crucial and direct role in triggering radio nuclear activity and not only via the morphology–density or the density–luminosity relations. We also used SDSS nuclear spectra for AMIGA galaxies to search for spectroscopic signatures of AGN and to study the properties of the underlying stellar populations. We produced a final catalogue of AGN candidate galaxies indicating 22% as an upper limit for the AGN fraction. We are currently studying this rate for samples with higher environmental densities in the bibliography, with a special care to perform the analysis in exactly the same way, in order to avoid biased results.

5 Atomic Gas

We also explored properties of the cold gas in AMIGA galaxies using HI data for ∼800 galaxies which is about three times more than the classical study in [12]. We presented the preliminary results in [11]. We have also quantified the degree of asymmetry of the HI single dish profiles and revised the results from the bibliography suggesting that more than half of the galaxies present a perturbed HI independently of the environment ([4, 13, 23, 26, 29]). Our study shows that AMIGA exhibits the lowest HI asymmetry level in the local universe. We found that field samples present an excess of ∼20% more asymmetric HI profiles than those in CIG. Still a small percentage of galaxies in our sample present large asymmetries. We have started a follow-up of isolated and asymmetric systems using high-resolution VLA and GMRT maps to give insight into the origin of such asymmetries.

6 Concluding Remarks

We suggest that the size of AMIGA sample and the strength of the isolation criterion make it the optimal statistical reference sample for studying the low-density tail of the two-point correlation function and/or morphology–density relation. The AMIGA-vetted CIG is the largest compilation of the most isolated galaxies in the local universe. Other samples in the bibliography appear to contain more luminous early types and spirals with large bulges than found in the CIG. The fact that we find differences with other samples at optical, FIR, radio continuum, and HI wavelengths argue that these galaxies have spent most of their lives in relative isolation. The low-luminosity early-type fraction and the large fraction of small bulge spirals cannot have experienced major nurture events (i.e., mergers) in their lives. This is the closest we are likely to come to identifying a population of galaxies whose properties are dominated by nature rather than nurture.

Acknowledgments The authors are partially supported by DGI Grant AYA2008-06181-C02 and Junta de Andalucía (Spain) TIC-114 and P08-FQM-4205-PEX.

References

1. Adams, M., Jensen, E., Stocke, J. 1980, AJ, 85, 1010
2. Best, P.N., Kauffmann, G., Heckman, T.M., Brinchmann, J., Charlot, S., Ivezić, Ž., White, S.D.M. 2005, MNRAS, 362, 25
3. Bitsakis, T., Charmandaris, V., Le Floc'h, E., Diaz-Santos, T., Slater, S.K., Xilouris, E., Haynes, M.P. 2010, A&A, 517, 75
4. Bournaud, F., Combes, F., Jog, C.J., Puerari, I. 2005, A&A, 438, 507
5. Braine, J., Combes, F. 1993, A&AS, 97, 887
6. Buta, R. 1995, ApJS, 96, 39
7. De Robertis, M.M., Yee, H.K.C., Hayhoe, K. 1998, ApJ, 496, 93
8. Durbala, A., Sulentic, J., Buta, R., Verdes-Montenegro, L. 2008, MNRAS, 390, 881
9. Durbala, A., Buta, R., Sulentic, J., Verdes-Montenegro, L. 2009, MNRAS, 397, 1756
10. Eskridge, P.B., et al. 2002, ApJS, 143, 73
11. Espada, D., et al. 2009, arXiv:0909.2736
12. Haynes, M.P., Giovanelli, R. 1984, AJ, 89, 758
13. Haynes, M.P., van Zee, L., Hogg, D.E., Roberts, M.S., Maddalena, R.J. 1998, AJ, 115, 62
14. Iglesias-Páramo, J., Vílchez, J.M. 1999, ApJ, 518, 94
15. Karachentseva, V. 1973, Commun Spec Astrophys Obs, 8, 1
16. Kelm, B., Focardi, P. 2004, A&A, 418, 937
17. Kennicutt, R.C., Kent, S.M. 1983, AJ 88, 1094
18. Krongold, Y., Dultzin-Hacyan, D., Marziani, P., de Diego, J.A. 2003, RMxAA, 39, 225
19. Leon, S., Combes, F., Menon, T.K. 1998, A&A, 330, 37
20. Leon, S., Verdes-Montenegro, L. 2003, A&A, 411, 391
21. Leon, S., Verdes-Montenegro, L., Sabater, J., Espada, D., Lisenfeld, U., Ballu, A., Sulentic, J., Verley, S., Bergond, G., Garcia, E. 2008, A&A, 485, 475
22. Lisenfeld, U., Verdes-Montenegro, L., Sulentic, J., Leon, S., Espada, D., Bergond, G., Garcia, E., Sabater, J., Santander-Vela, J., Verley, S. 2007, A&A, 462, 507
23. Matthews, L.D., van Driel, W., Gallagher, J.S., III. 1998, AJ, 116, 1169
24. Miller, C.J., Nichol, R.C., Gómez, P.L., Hopkins, A.M., Bernardi, M. 2003, ApJ, 597, 142

25. Perea, J., del Olmo, A., Verdes-Montenegro, L., Yun, M.S. 1997, ApJ, 490, 166
26. Richter, O.-G., Sancisi, R. 1994, A&A, 290, L9
27. Sabater, J., Leon, S., Verdes-Montenegro, L., Lisenfeld, U., Sulentic, J., Verley, S. 2008, A&A, 486, 73
28. Sulentic, J., Verdes-Montenegro, L., Bergond, G., Lisenfeld, U., Durbala, A., Espada, D., Garcia, E., Leon, S., Sabater, J., Verley, S., Casanova, V., Sota, A. 2006, A&A, 449, 937
29. Swaters, R.A., Schoenmakers, R.H.M., Sancisi, R., van Albada, T.S. 1999, MNRAS, 304, 330
30. Verdes-Montenegro, L., Sulentic, J., Lisenfeld, U., Leon, S., Espada, D., Garcia, E., Sabater, J., Verley, S. 2005, A&A, 436, 443
31. Verdes-Montenegro, L., Yun, M.S., Perea, J., Del Olmo, A., Ho, P.T.P. 1998, ApJ, 497, 89
32. Verdes-Montenegro, L., Yun, M.S., Williams, B.A., Huchtmeier, W.K., Del Olmo, A., Perea, J. 2001, A&A, 377, 812
33. Verley, S., et al. 2007, A&A, 470, 505
34. Verley, S., et al. 2007, A&A, 472, 121
35. Verley, S., Combes, F., Verdes-Montenegro, L., Bergond, G., Leon, S. 2007, A&A, 474, 43
36. Xu, C., Sulentic, J. 1991, ApJ, 374, 407
37. Zasov, A.V., Sulentic, J.W. 1994, ApJ, 430, 179

Diffuse Light and Galaxy Interactions in the Core of Nearby Clusters

Magda Arnaboldi

Abstract The kinematics of the diffuse light in the densest regions of the nearby clusters can be unmasked using the planetary nebulae (PNs) as probes of the stellar motions. The position–velocity diagrams around the brightest cluster galaxies (BCGs) identify the relative contributions from the outer halos and the intracluster light (ICL), defined as the light radiated by the stars floating in the cluster potential. The kinematics of the ICL can then be used to assess the dynamical status of the nearby cluster cores and to infer their formation histories. The cores of the Virgo and Coma are observed to be far from equilibrium, with mergers currently ongoing, while the ICL properties in the Fornax and Hydra clusters show the presence of sub-components being accreted in their cores, but superposed to an otherwise relaxed population of stars. Finally the comparison of the observed ICL properties with those predicted from Λ CDM simulations indicates a qualitative agreement and provides insights on the ICL formation. Both observations and simulations indicate that BCGs halos and ICL are physically distinct components, with the "hotter" ICL dominating at large radial distances from the BCGs halos as the latter become progressively fainter.

1 Diffuse Light in Clusters

Deep imaging of massive clusters of galaxies shows that a population of star exists that fills the space among galaxies in clusters. While the indication of its existence came with the early detection by Zwicky in 1951 for the Coma cluster, it was only with the advent of the wide-field cameras equipped with mosaic CCDs that the properties of the clusters' diffuse light, i.e., morphology, radial distribution, and colors, were studied in a quantitative way [7, 23, 29, 35, 39].

Accurate photometric measurements of the diffuse light are difficult to perform because (1) its features are at extremely faint surface brightness of $<1\%$ of the night

M. Arnaboldi (✉)
European Southern Observatory, Garching, Germany; and Observatory of Turin, INAF, Turin, Italy
e-mail: marnabol@eso.org

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_34,

sky and (2) it is difficult to disentangle the contribution of the extended outer halos of the brightest cluster galaxies (BCGs) in a cluster core from that of the stars free floating in the cluster potential, i.e., the intracluster light component or ICL. Without the kinematic measurements, the division between BCG halos and ICL is somewhat arbitrary.

Since the discovery of free-floating intracluster planetary nebulae (ICPNs) in the Virgo cluster [2], extensive imaging and spectroscopic observations were carried out to determine their projected phase-space distribution, and from it the amount of ICL.

The measurement of the ICL in clusters is relevant from the determination of the baryonic fraction condensed in stars, the star formation efficiency, and the metal enrichment of the hot intracluster medium, especially in the cluster cores. Because of the long dynamical time across the cluster regions, we expect the distribution function $f(\bar{x}, \bar{v})$ of the ICL stars to be different depending on the formation mechanism and its assembly history.

2 PNs as Kinematical Traces

Planetary nebulae (PNs) are the late phase of solar-like stars and in stellar populations older than 2 Gyrs one star every few millions is expected to undergo such phase [10]. Stars in the PN phase can be detected via the relative bright emission in the optical [OIII] line, at λ 5,007 Å, because the nebular shell re-emits 10% of the UV photons emitted by the white dwarf at its center in the [OIII] λ5,007 Å line. When such emission line is detected, the line-of-sight velocity of the PN can be measured via a Gaussian fit.

The number density of PNs traces the light of the parent stellar population. According to the single stellar population theory, the luminosity-specific stellar death rate is independent of the precise star formation history of the associated stellar population [33]. This property is captured in a simple relation such that

$$N_{PN} = \alpha L_{gal} \tag{1}$$

where N_{PN} is the number of all PNs in a stellar population, L_{gal} is the bolometric luminosity of that parent stellar population, and α is the luminosity-specific PN number. Predictions from the stellar evolution theory are further supported by the empirical evidence that the PN number density profiles follow light in late- and early-type galaxies [12, 26] and that the luminosity-specific PN number α stays more or less constant for (B−V) color < 0.8, and then decreases by about a factor 7 for very red (B−V) > 0.8 and old stellar population [10].

3 The Virgo Cluster Core

The presence of diffuse light in the Virgo cluster core is clearly illustrated by the deep image of Mihos et al. [29], reaching $\mu_V = 28$ mag arcsec^{-2}. It shows a variety of features such as streamers, arcs, plumes, and very extended diffuse halos

surrounding the large galaxies in the field at surface brightness level of $\mu_v \approx 26.5$ mag arcsec^{-2}, and in the case of M87, with flattened isophotes (c/a=0.5) out to $\sim 37' = 161$ kpc. The different morphologies suggest that we may be seeing several systems superposed along the line of sight (LOS) to the Virgo core and their dynamical status may be characterized by different kinematics.

For the Virgo cluster, there has been considerable success with a two-step approach of identifying PN candidates with narrowband imaging followed by multiobject spectroscopy. Arnaboldi et al. [2] observed the outer regions of the giant elliptical M86, measuring velocities for 19 objects. While M86's $v_{sys} = -240\,\mathrm{km\,s^{-1}}$, three of these PNs have velocities larger than the mean velocity of the Virgo cluster $\bar{v}_{virgo} = 1,100\,\mathrm{km\,s^{-1}}$, and turned out to be true ICPNs. Subsequently, 23 PNs were detected in a spectroscopic survey[1] with 2dF on the 4 m Anglo-Australian Telescope [3, 20]. The first confirmation based on detecting the [OIII] doublet in a single PN spectrum was made in Arnaboldi et al. [4]. Since then, Arnaboldi et al. ([5], hereafter A04) began a campaign to systematically survey PN candidates in the Virgo cluster using multiobject spectroscopy with the FLAMES/GIRAFFE spectrograph on the VLT. A04 presented the first measurements of the velocity distribution of PNs from three survey fields in the Virgo cluster core and concluded that in two of these fields the light is dominated by the extended halos of the nearby giant elliptical galaxies, while the ICL component dominates the diffuse light in only one field, where a "broad" line-of-sight velocity distribution is measured, and all PNs are true "ICPNs."

The A04 sample was further enlarged with the PN spectra obtained with FLAMES at the ESO VLT by Doherty et al. ([14], hereafter D09), and Fig. 1 shows the overview of all the fields in the Virgo cluster core studied thus far.

3.1 PNs Line-of-Sight Velocity and Projected Phase-Space Distributions

On the bases of the PNs' position and v_{LOS} in the Virgo cluster core, one can compute the projected phase-space diagram by plotting the PNs v_{LOS} vs. radial distance from the center of M87. In this v_{LOS} vs. $radius$ diagram several regions can be identified with very different densities: for projected distances $R < 2,400''$ most of the PNs are strongly clustered around the systemic velocity of M87, $v_{sys} = 1,307\,\mathrm{km\,s^{-1}}$, while at $R > 2,400''$ the PN velocities spread widely over a velocity range more typical for the Virgo cluster. From this intracluster region, we see a string of low PN velocities at $\leq 800\,\mathrm{km\,s^{-1}}$ reaching far into the M87 halo. At projected distance $R < 1,300''$ there are two of these intracluster PNs at $\sim 400\,\mathrm{km\,s^{-1}}$. The remaining PNs are distributed symmetrically around M87's v_{sys} and have mean velocity $1,276 \pm 71\,\mathrm{km\,s^{-1}}$ and velocity dispersion $\sigma = 247\,\mathrm{km\,s^{-1}}$ (A04). At $R \sim 2,000''$, five PNs are tightly clustered around $v_{sys} = 1,307\,\mathrm{km\,s^{-1}}$;

[1] These results were all based on single line identifications, although the second oxygen line was seen with the right ratio in the composite spectrum of 23 PNs observed by Freeman et al. [20].

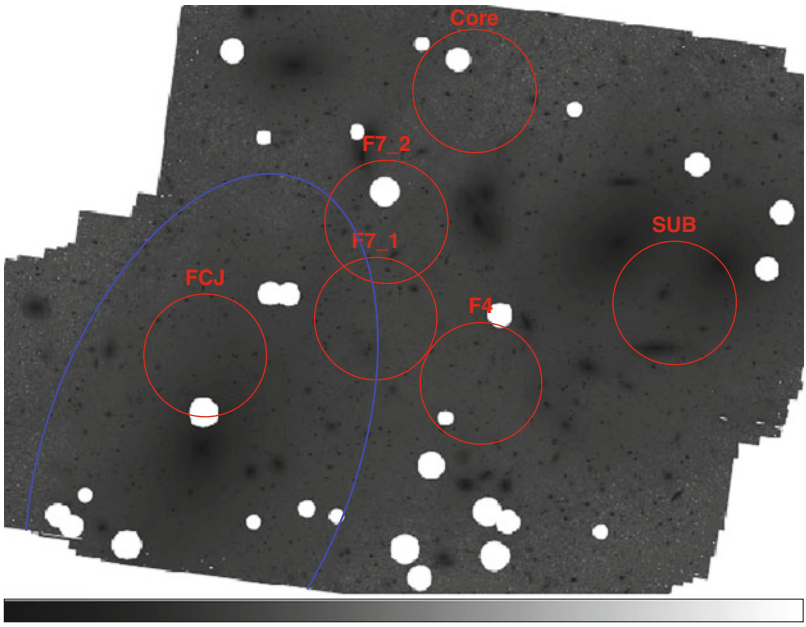


Fig. 1 Deep image of the Virgo cluster core showing the diffuse light distribution [29], with the target fields of [5] and [14] indicated by *red circles*. The *blue ellipse* shows the outer isophotes of the M87 halo according to Kormendy et al. [25]

these have mean velocity $1,297 \pm 35\,\mathrm{km\,s^{-1}}$ and an rms dispersion of 78 km $\mathrm{s^{-1}}$ [14]. At comparable radii there are two additional PNs with velocities of 753 and $634\,\mathrm{km\,s^{-1}}$; compared to the previous five, these are 7σ and 8σ outliers. It is unlikely that one or two of these outliers are part of the same (very asymmetric) distribution as the five PNs clustered around M87's v_{sys}.[2] By contrast, they fit naturally into the stream leading from a distance of $1,300''$ from the M87 center all the way into the ICL.

The PNs bound to the M87 halo are then only those ones which are clustered around the systemic velocity of M87. These are confined to radii $R < 2,400''$. Outside $R = 2,400''$ in Fig. 2 the PNs have larger relative velocities with respect to M87's v_{sys}, with an approximately uniform distribution in the range -300 to $-2,600\,\mathrm{km\,s^{-1}}$. Those in the radial range $2,400'' < R < 3,600''$ are confined to negative velocities with respect to M87, indicating that the ICL component is not phase-mixed yet. These are probably encroaching stars from M86 and other Virgo components. By contrast, the PNs further than $3,600''$ from M87 show a broad distribution of velocities, more characteristic of the cluster as a whole.

[2] The probability of finding five PNs with velocity dispersion less than $80\,\mathrm{km\,s^{-1}}$ from a Gaussian velocity distribution with $\sigma = 250\,\mathrm{km\,s^{-1}}$ around $\bar{v} = 1,300\,\mathrm{km\,s^{-1}}$ as measured at 60 kpc is less than 1%.

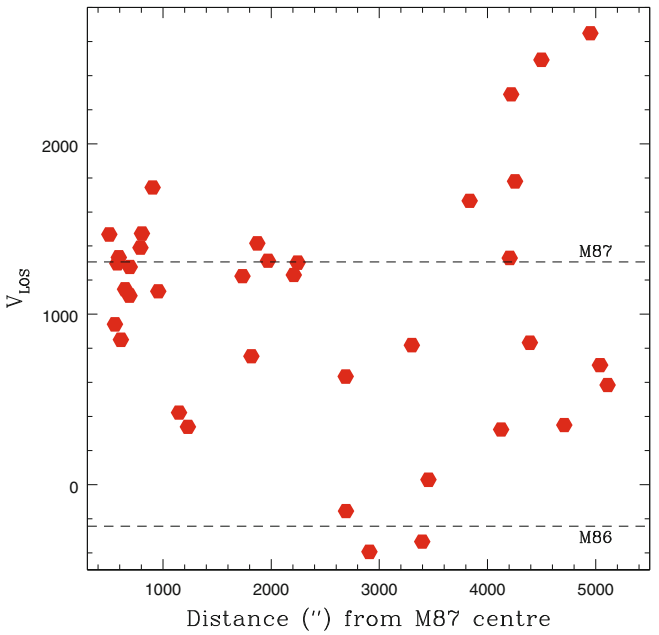


Fig. 2 Distribution of v_{LOS} vs. projected distance from the center of M87 for all spectroscopically confirmed PNs in the Virgo core. From [14]

The ICL PNs show up as approximately flat velocity distribution (VD) besides the peak of velocities from PNs bound to M87. A flat distribution of velocities in addition to the peak near M87's systemic velocity is also seen in the LOSVD of the dwarf spheroidal galaxies in the same region of the Virgo cluster core [8]. However, for the dwarf galaxies the broad velocity distribution extends to significantly more red-shifted velocities, indicating that the dwarfs and ICL PNs kinematics can only partially be related.

In Fig. 3 we plot the PNs' positions and those of the dwarfs' in the cluster core. Those PNs in the densest part of the projected phase-space diagram are spatially associated with the halo of M87, and they are segregated in radii within $R < 161$ kpc, while the ICPNs are scattered across the whole core region, including the M87 halo.

3.2 Dynamical Status of the Virgo Cluster Core

The LOSVD of dwarf spheroidals (dE+dS0) in a 2° radius circular region centered on M87 is very flat and broad, with the peak of the distribution at 1,300 km s^{-1} and a long tail of negative velocities [8]. The LOSVD of the ICPNs now confirms that this asymmetry is also present in the very center of the Virgo core, in a 1° diameter region. The projected phase-space diagram of Fig. 2 shows that velocities

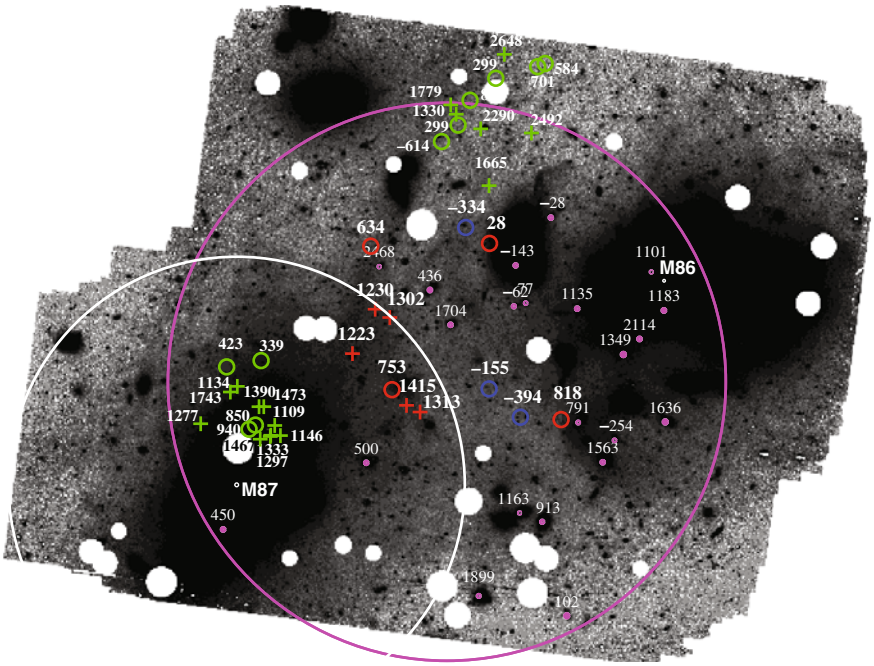


Fig. 3 Deep image of the Virgo cluster core showing the distribution of the intracluster light (from [29]). The spatial distribution of the spectroscopically confirmed PNs is overlaid. The A04 targets are shown in *green*; D09 targets are shown in *red* if red-shifted with respect to Earth and *blue* if blue shifted. Objects with velocities higher than the mean velocity of Virgo (1,100 km s^{-1}) are shown as *crosses* and those with lower velocities shown as *circles*. Dwarf spheroidals are marked as *magenta dots*. The velocities (in km s^{-1}) are labeled for all objects shown. The nominal "edge" of the M 87 halo at ≈160 kpc is indicated with a white circle. The *pink circle* has a 1.5 degree diameter and is centered on the projected midpoint of M87 and M86. North is up and East is to the left. From D09

near the systemic velocity of M86 are seen to about half-way from M86 to M87. The asymmetry and skewness of the LOSVD may arise from the merging of sub-clusters along the LOS [36]. In a merging of two clusters of unequal mass, the LOSVD is highly asymmetric with a long tail on one side and a cutoff on the other side, shortly ($\sim 10^9$ year) before the clusters merge. The observed LOSVDs of the PNs, GCs [13], and (dE+dS0) in the Virgo core can be interpreted as evidence that the two massive sub-clusters in the Virgo core associated with the giant ellipticals M87 and M86 are currently falling toward each other – more or less along the LOS, with M87 falling backward from the front and M86 forward from the back – and will eventually merge, i.e., the entire core of the Virgo cluster must then be out of virial equilibrium and dynamically evolving.

What is the relative distance between M87 and M86? Do their halos already touch each other or are they just before their close pass? PNLF distances [24] and

ground-based surface brightness fluctuation distances indicate that M86 is behind M87 by just under ∼0.15 mag. However, the most recent surface brightness fluctuation measurements find that M87 and M86 are only at very slightly different distances. Within the errors, the distance moduli (M87: 31.18 ± 0.07, M86: 31.13 ± 0.07) are consistent with being either at the same distance or separated by 1–2 Mpc. Unfortunately the evidence from the relative distances of M87/M86 is not conclusive at this stage.

3.3 ICL Large-Scale Distribution in the Virgo Cluster from PNs Narrowband Surveys

Several studies investigated the properties of the ICL in the core of the nearby Virgo cluster by mapping the number density distribution of ICPNs. Expanding on the earlier narrowband imaging work in the Virgo cluster core [1, 3, 17–19, 32], Castro-Rodriguéz et al. [11] completed a survey campaign of the ICL distribution at larger scales, outside the Virgo cluster core. In total, they covered more than 3 $\deg^2$ in Virgo, at 11 different positions in the cluster and at distances between 80 arcmin and some 100 arcmin from the Virgo core. In several of these regions, the ICL is at least two magnitudes fainter than in the core.

The diffuse light observed in the core of a galaxy cluster contains several luminous stellar components that add up along the LOS to the cluster center: the extended faint halos of the brightest galaxies and the ICL contribution. When computing the ICL fraction in the Virgo core, the surface brightness measurements must then be corrected for the fraction of stars bound to the extended galaxy halos. When one selects only true ICPNs, Castro-Rodriguéz et al. [11] measure a surface brightness for the ICL of about $\mu_B = 28.8 - 29.5$ mag arcsec^{-2} in the Virgo core and these surface brightness values are consistent with those inferred from the detection of IC RGB stars [41].

The comprehensive summary of all surface brightness measurements in Fig. 4 based on ICPN number counts indicates that most of the diffuse light is detected in fields located in the core, within a distance of about 80 arcmin ≈ 350 kpc, from M87. Outside the core, the mean surface brightness decreases sharply and the ICL is confined within isolated pointings.

Diffuse stellar light is also measured in sub-structures, around M49 and in the M60/M59 sub-group. The fields F04-2 and F04-6 from Feldmeier et al. [19] are situated in the outer regions of M49, at about 150 kpc from the galaxy center. These fields may contain PNs from the halo of M49 and those associated with the ICL component, which may have formed within this sub-clump of the Virgo cluster. Because the spectroscopic follow-up is not yet available for these PN candidates, we cannot quantify the fraction of light in the M49 halo and ICL for these fields.

The result for the ICL being more centrally concentrated than the galaxies in the Virgo cluster is confirmed generally for clusters with central BCGs. It agrees with

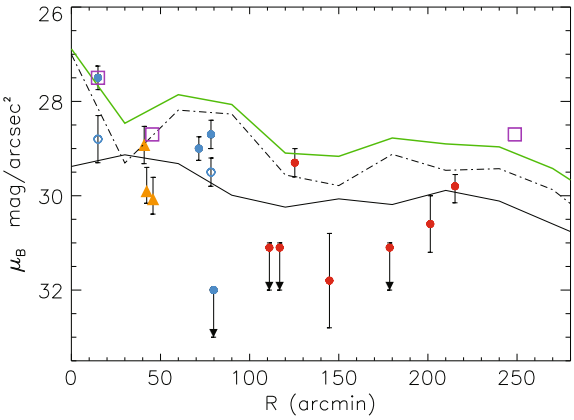


Fig. 4 Surface brightness measurement of diffuse light in the Virgo fields (*points*) compared with the surface brightness profile of the Virgo galaxies averaged in annuli (*lines*); radial distances are computed with respect to M87. The *green dotted line* represents the radial surface brightness profile from light in Virgo galaxies from Binggeli et al. [9]. The *dotted-dashed* and *double dotted-dashed lines* correspond to the surface brightness profile associated with giants and dwarf galaxies, respectively. The *full blue dots* show the surface brightness measurements in the Virgo core fields using ICPNs. The *open circles* indicate the ICL surface brightness computed from true ICPNs, i.e., PNs not bound to galaxy halos. The *triangles* represent the surface brightness of the ICL based on IC RGB star counts (see [41] and the reference therein). The *full red dots* show the surface brightness measurements outside the Virgo core fields using ICPNs. The *arrows* indicate the upper limits for the measurements of PNs at certain field positions. The *magenta open squares* indicate the surface brightness average values μ_B at 15, 50, and 240 arcmin based on the measurements from Feldmeier et al. [19] data; the measurements at 240 arcmin (F04-2 and F04-6) are close to M49. Radial distances are computed with respect to M87. From Castro-Rodriguéz et al. [11]

statistical observations of intermediate redshift clusters [42] and their diffuse light radial profiles.

4 Observing Techniques for the Kinematics of Diffuse Light in Clusters

In this section we present a short overview of the observing techniques that allow studies of the kinematic of the diffuse light in clusters far beyond 15 Mpc and explore distant clusters out to Coma, at 100 Mpc. A number of techniques have been developed recently that allow detection and velocity measurements of PNs in the same observing run. We refer to the work of Méndez et al. [28], which used a sequence of narrowband plus broadband images, followed by the dispersed image of the same field. This technique has been applied to NGC 4967 in Virgo and to NGC 1344 in Fornax. The PN Spectrograph (PN.S) is an instrument on the William Herschel Telescope dedicated to measuring PN velocities in nearby galaxies [16]. The pupil is split into half before being dispersed in opposite directions in twin spectrographs. A combination of the two exposures allows the identification of emission

line objects and their velocity measurements from the separation between positions in two spectral images.

4.1 Counter Dispersed Slitless Imaging Technique

The counter-dispersed slitless technique (CDI) used by McNeil et al. [27] uses only two exposures for each field to obtain positions and velocities of PNs for the cD galaxy NGC 1399, in the Fornax cluster, at D=17 Mpc distance. The field is observed with a dispersed image through a narrowband filter centered at the red-shifted emission of the PN at the systemic velocity of the galaxy being studied. Next the spectrograph is rotated 180° and the same field is exposed again, this time with the dispersion in the opposite direction. As in the PN.S observations, the velocity is a function of the separation between the position of the PN in the two frames. In this way, the slitless technique avoids the two-stage selection and measurement process. Because there are no slits or fibers, the light loss is reduced. The number of measurable PN velocities is not limited by either the number of available fibers or the restrictive geometry of the slit. For a comprehensive presentation of the CDI technique and the calibration procedure to derive relative and absolute velocity and sky position of the emission line sources we refer the reader to Mc Neil et al. [27].

4.2 The Multi-slit Imaging Spectroscopy Technique

For the PNs in distant galaxies $D > 20$ Mpc, the flux in the [OIII] λ 5,007 Å emission becomes of the same order as the sky noise in a 30−40 Å wide filter; therefore we need to deploy a technique which substantially reduced the noise in the background sky to be able to detect them. Observing the entire field through slits, a narrowband filter and a dispersing element significantly reduces the signal from the sky, because we limit it to few Angstroms only. The multi-slit imaging spectroscopy (MSIS) technique pioneered at the 8.2 m Subaru telescope and FOCAS spectrograph [21] involves a grid of slits that are stepped until the entire field has been spectrally imaged through a narrowband filter and a grism. It is a blind technique and spectra are obtained for all the emission line objects that happen to lie beyond the slits. These types of observations detected sample of intracluster PNs in the Hydra [40] and in the Coma cluster [6, 21].

5 The Un-mixed Kinematics of the Intracluster Stars in the Fornax and Hydra Cluster Cores

In this section we provide a concise overview of the most recent results on the un-mixed kinematics of intracluster stars based on PNs velocities in the cores of the

Fornax and Hydra clusters, from the works of McNeil et al. [27] and Ventimiglia et al. (2010, in prep.).

The Fornax cluster and NGC 1399 – Using the CDI technique with the FORS1 spectrograph at the ESO VLT, McNeil et al. [27] discovered 187 PNs around NGC 1399, the cD galaxy in the Fornax cluster. Data were extracted from a mosaic of five fields, which included also the nearby galaxy NGC 1404. Each PN was further classified on the basis of the light contribution from the two galaxies at its position, and the difference between the PN velocity and the systemic velocities of NGC 1399 ($v_{sys} = 1,425\,\mathrm{km\,s^{-1}}$; from NED) and NGC 1404 ($v_{sys} = 1,917\,\mathrm{kms^{-1}}$; from NED). This procedure identified 146 PNs associated with NGC 1399, 23 PNs associated with NGC 1404, while 6 PNs were unassigned. The projected PNs phase-space distribution v_{LOS} vs. radius in the NGC 1399 surveyed regions indicated the presence of a kinematic component at low velocity $v_{mean} = 800\,\mathrm{km\,s^{-1}}$ also, in addition to the PN population of NGC 1399 and NGC 1404. A total of 12 PNs are associated with the low-velocity component, and they are scattered across the NGC 1399 light distribution, with a concentration in the north-central part. The most recent globular cluster work also shows a sample of low-velocity objects at similar velocity [37] which independently support the reality of this sub-structure.

The presence of a velocity sub-component in the PNs sample superposed on a PN population bound to the NGC 1399 halo indicates the presence of a heterogeneous population including stars left over from previous accretion, merger, or tidal stripping events; these stars are still in the process of phase mixing.

The Hydra cluster and NGC 3311 – Ventimiglia et al. (2010, in prep.) performed MSIS observations with FORS2 at the ESO VLT of the core region of the Hydra cluster, centered on its cD galaxy NGC 3311. They detected a total of 56 PNs in a single field of $100 \times 100\,\mathrm{kpc}^2$ and analyzed their velocity field. The PNs LOSVD in this region follows a multi-peaked distribution; see Fig. 5: in addition to a broad symmetric component centered at the systemic velocity of the cluster ($v_{hydra} = 3,683\,\mathrm{km\,s^{-1}}$), two narrow peaks are detected at $1,800$ and $5,000\,\mathrm{kms^{-1}}$. These secondary peaks unmask the presence of un-mixed stellar populations in the Hydra core.

The spatial distribution of the PNs associated with the narrow velocity sub-component at $5,000\,\mathrm{km\,s^{-1}}$ is superposed and concentrated on an excess of light in the northeast quadrant of NGC 3311, as detected from the two-dimensional light decomposition of the NGC 3311/NGC 3309 B band photometry. On the same sky position we detect a group of dwarf galaxies with $v_{LOS} \approx 5,000\,\mathrm{km\,s^{-1}}$. Deep long-slit spectra obtained at the position of the dwarf galaxy HCC26 situated at the center of the light excess show absorption line features from both HCC26 and the light excess which are consistent with v_{LOS} of about $5{,}000\,\mathrm{km\,s^{-1}}$ (Ventimiglia et al. 2010). We conclude that the PNs in the $5,000\,\mathrm{km\,s^{-1}}$ sub-component, the dwarfs galaxies and the light excess in the northeast quadrant of NGC 3311 occupy the same region of the phase space and are physically associated.

Also in the case of the relaxed Hydra-I cluster, PN kinematics, photometry, and deep absorption spectra support the evidence for an accretion event whose stars are being added to both the cluster core and the halo of NGC 3311.

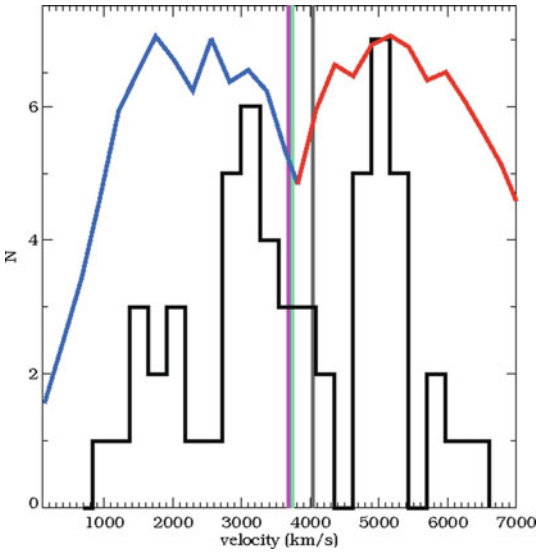


Fig. 5 The PNs LOSVD in the core of the Hydra cluster obtained from the MSIS observations by Ventimiglia et al. (2010). The LOSVD is shown by the *black line*, while the *blue and red curves* indicate the measured transmission curves of the narrowband filters used to cover the Hydra velocities; the normalization is arbitrary. The *magenta, green, and gray vertical lines* indicate the systemic velocity of the Hydra-I cluster, NGC 3311, and NGC 3309, respectively

6 The Ongoing Sub-cluster Merger in the Coma Cluster Core

The Coma cluster is the richest and most compact of the nearby clusters, yet there is growing evidence that its formation is still ongoing. A sensitive probe of this evolution is the dynamics of intracluster stars, which are unbound from galaxies while the cluster forms, according to cosmological simulations. With the MSIS technique Gerhard et al. [21] detected and measured the v_{LOS} of 37 ICPNs associated with the diffuse stellar population of stars in the Coma cluster core, at 100 Mpc distance. These are the most distance single stars whose spectra have been acquired in addition to cosmological supernovae stars. Gerhard et al. [22] detected clear velocity sub-structures within a 6 arcmin diameter field centered at $\alpha(J2000) = 12:59:41.8$; $\delta(J2000) = 27:53:25.4$, nearly coincident with the field observed by Bernstein et al. [7] and ∼5 arcmin away from the cD galaxy NGC 4874. A sub-structure is present at $\sim 5,000\,\mathrm{km\,s^{-1}}$, probably from in-fall of a galaxy group, while the main intracluster stellar component is centered around $\sim 6,500\,\mathrm{km\,s^{-1}}$, $\sim 700\ \mathrm{km\ s^{-1}}$ offset from the nearby cD galaxy NGC 4874 ($v_{sys} = 7,224\ \mathrm{km\ s^{-1}}$; from NED). The kinematics and the elongated morphology of the intracluster stars [38] show that the cluster core is in a highly dynamically evolving state. In combination with galaxy redshift and X-ray data, this argues strongly that the cluster is currently in the midst of a sub-cluster merger. The NGC 4889 sub-cluster is likely to have fallen into Coma from the eastern A2199 filament, in a

direction nearly in the plane of the sky, meeting the NGC 4874 sub-cluster arriving from the west. The two inner sub-cluster cores are presently beyond their first and second close passages, during which the elongated distribution of diffuse light has been created. Gerhard et al. [22] also predict the kinematic signature expected in this scenario and argue that the extended western X-ray arc recently discovered traces the arc shock generated by the collision between the two sub-cluster gas halos.

7 Cosmological Simulations and ICL

The predicted spatially averaged radial distribution of ICL from recent high-resolution hydrodynamical simulations of cluster formation in Λ CDM universe [30] is in broad agreement with the observed radial profiles for the ICL in clusters. Furthermore, predictions from these simulations that the largest portion of the ICL is formed during the assembly of the most luminous cluster galaxies [31, 34] are supported by the observed ongoing mergers in Coma and Virgo cores and by the presence of ICL around sub-clusters/groups.

Quantitative analysis of the ICL kinematic from cosmological simulations is ongoing (Coccato et al. 2010, in prep.). Studies of the galaxy halo and ICL particles in cosmological simulations [15] further support the physical distinction between the central BCG and the ICL component in clusters.

8 Summary and Conclusions

The kinematics of the ICL provide unique information to assess the dynamical status of the nearby cluster cores:

- In the Virgo cluster, M87 and the M86/M84 sub-clusters are approaching along the LOS and they are currently before their first close passage.
- In the Coma cluster, the elongated morphology and kinematics of the ICL, the galaxy morphology and X-ray data are consistent with the merging of NGC 4889 and NGC 4874 along a binary orbit, the two galaxies being currently observed beyond their first and second close passages.
- The Hydra and Fornax clusters show evidence for un-mixed stellar components coming from accretion/tidal stripping events in their cores.

The superposition of the various kinematic components underscores the complexity of velocity measurements in cluster cores and the necessity of using discrete tracers to detect these components. In all cases the kinematical data indicate that the galaxy halos and ICL are discrete components; and the former do not blend continuously in the latter. The evidence for merging (in Coma) and accretion indicates that the build up of the ICL is an ongoing process.

Acknowledgments I wish to acknowledge my long-standing and productive collaboration with Ken Freeman, which lead to some 40 published refereed papers and tens of observing nights carried out together at several telescopes around the world (2.3 m SSO, AAT, CFHT, VLT, and Subaru to name a few). Dear Ken, I look forward to our future collaborations!

I also thank David Block for the invitation to participate at the conference to honor Ken Freeman and all his hard work, dedication, and endurance which made the Ken Freeman Conference in the Namibia desert and its proceedings come true.

Finally I wish to thank my collaborators Ortwin Gerhard, Lodovico Coccato, Payel Das, Michelle Doherty, Emily McNeil, Nieves Castro-Rodriguéz, and Giulia Ventimiglia.

References

1. Aguerri, M., et al. 2005, AJ, 457, 771
2. Arnaboldi, M., et al. 1996, ApJ, 472, 145
3. Arnaboldi, M., et al. 2002, AJ, 123, 760
4. Arnaboldi, M., et al. 2003, AJ, 125, 514
5. Arnaboldi, M., et al. 2004, ApJ, 614, L33 (A04)
6. Arnaboldi, M., et al. 2007, PASJ, 59, 419
7. Bernstein, G.M., et al. 1995, AJ, 110, 1507
8. Binggeli, B., Popescu, C.C., Tammann, G.A. 1993, A&AS, 98, 275
9. Binggeli, B., Tammann G.A., Sandage, A. 1987, AJ, 94, 251
10. Buzzoni, A., Arnaboldi, M., Corradi, R.L.M. 2006, MNRAS, 368, 877
11. Castro-Rodriguéz, N., et al. 2009, A&A, 507, 621
12. Coccato, L., et al. 2009, MNRAS, 394, 1249
13. Côté, P., et al. 2001, ApJ, 559, 828
14. Doherty, M., et al. 2009, A&A, 502, 771 (D09)
15. Dolag, K., Murante, G., Borgani, S. 2010, MNRAS, 405, 1544
16. Douglas, N., et al. 2002, PASP, 114, 1234
17. Feldmeier, J.J., et al. 1998, ApJ, 503, 109
18. Feldmeier, J.J., et al. 2003, ApJS, 145, 65
19. Feldmeier, J.J., et al. 2004, ApJ, 615, 196
20. Freeman, K.C., et al. 2000, ASP Conf Ser, ASP, San Francisco, CA, vol. 197, p. 389
21. Gerhard, O., et al. 2005, ApJ, 621, L93
22. Gerhard, O., et al. 2007, A&A, 468, 815
23. Gregg, M.D., West, M., 1998, Nature, 396, 549
24. Jacoby, G., et al. 1990, ApJ, 356, 332
25. Kormendy, J., et al. 2009, ApJS, 182, 216
26. Kimberly, H.A, et al. 2008, ApJ, 683, 630
27. McNeil, E.K., Arnaboldi, M., et al. 2010, A&A, 518, 44
28. Méndez, R.H., et al. 2001, ApJ, 563, 135
29. Mihos, J.C., et al. 2005, ApJ, 631, L41
30. Murante, G., et al. 2004, ApJ, 607, L83
31. Murante, G., et al. 2007, MNRAS, 377, 2
32. Okamura, S., et al. 2002, PASJ, 54, 883
33. Renzini, A., Buzzoni, A. 1986, In: Spectral Evolution of Galaxies, eds. Chiosi C., Renzini A., Reidel, Dordrecht, p. 195
34. Rudick, C.S., et al. 2006, ApJ, 648, 936
35. Rudick, C.S., et al. 2010, ApJ, 720, 569
36. Schindler, S., Boehringer, H. 1993, A&A, 269, 83
37. Schuberth, Y., et al. 2010, A&A, 513, 52
38. Thuan, T.X., Kormendy, J. 1977, PASP, 89, 466
39. Uson, J.M., et al. 1991, ApJ, 369, 46
40. Ventimiglia, G., Arnaboldi, M., Gerhard, O. 2008, AN, 329, 1057
41. Williams, B.F., et al. 2007, ApJ, 654, 835
42. Zibetti, S., et al. 2005, MNRAS, 358, 949

Feedback in Star and Galaxy Formation

Joseph Silk

Abstract I review some of the outstanding questions that pertain to feedback in star and in galaxy formation. I address the initial stellar mass function, the efficiency of star formation, the galactic star formation rate, and the role of supernovae and supermassive black holes.

1 Introduction

Here are some outstanding questions that pertain to feedback in star and in galaxy formation. Can we predict the initial stellar mass function? Can we account for the efficiency of star formation? Can we account for the galactic star formation rate? Do we understand supermassive black hole feedback? The answer to the questions in all cases is no. But we should not despair. The missing link that yields a common thread to these questions is the need for a robust theory of star formation. I will argue that feedback in its diverse manifestations helps to partially resolve these issues. However, much still remains to be done, both observationally and theoretically.

2 In the Beginning . . .

Star formation theory begins with the founder of the theory of gravitation. Isaac Newton realized in a letter he wrote to a clergyman on December 10, 1692, that fragmentation and subsequent star formation were inevitable in an infinite and initially homogeneous cloud. Gravity operated irreversibly and inevitably in accumulating matter around density fluctuations.

J. Silk (✉)
Department of Physics, Beecroft Institute for Cosmology and Particle Astrophysics, University of Oxford, Oxford OX1 3RH, UK
e-mail: j.silk1@physics.ox.ac.uk

D.L. Block et al. (eds.), *Galaxies and their Masks*,

DOI 10.1007/978-1-4419-7317-7_35,

> If the matter was evenly disposed throughout an infinite space, it could never convene into one mass; but some of it would convene into one mass and some into another, so as to make an infinite number of great masses, scattered at great distances from one to another throughout all that infinite space. And thus might the sun and fixed stars be formed, supposing the matter were of a lucid nature.

Newton's insight was remarkable. However, he could not understand how gravity could differentiate between luminous bodies, or stars, and opaque bodies, or planets.

> How the sun alone should be changed into a shining body whilst all the planets continue opaque, or all they be changed into opaque ones whilst he remains unchanged, I do not think explicable by mere natural causes, but am forced to ascribe it to the counsel and contrivance of a voluntary Agent.

An intensely religious man, Newton gave up in despair at this point and appealed to a higher entity to come to the rescue.

James Jeans was not one to share this opinion, however. He developed fragmentation into quantitative physics. To him,

> From the intrinsic evidence of his creation, the Great Architect of the Universe now begins to appear as a pure mathematician.

In 1902, he developed the theory of gravitational fragmentation which is now central to our understanding of star formation.

> We have found that as Newton first conjectured, a chaotic mass of gas of approximately uniform density and of very great extent would be dynamically unstable: nuclei would tend to form in it, around which the whole matter would eventually condense. All celestial bodies originate by a process of fragmentation of nebulae out of chaos, of stars out of nebulae, of planets out of stars and satellites out of planets.
>
> James Jeans

But Jeans did not solve the challenge posed by Newton of why stars as opposed to planets. The astronomer who faced this challenge was Arthur Eddington, who developed the theory of self-gravitating polytropic spheres in order to model stars. Simple stability considerations led him to realize that stars occupied a relatively narrow mass range. He showed in 1926 that star formation was inevitable:

> Imagine a physicist calculating on a cloud-bound planet and ending with the dramatic conclusion, "What 'happens' is the stars."
>
> Arthur Eddington

3 Star Formation

The Jeans mass sets the scale of fragmentation. It is defined to be the mass within a sphere of diameter the Jeans length, approximately the distance a sound wave crosses in a free fall time, and is proportional to $T^{3/2}\rho^{-1/2}$. At low densities, interstellar clouds radiate freely and are isothermal. During the isothermal phase of contraction, the Jeans mass decreases. Eventually the cloud becomes self-shielding and the ensuing contraction is approximately adiabatic once the optical depth is large. The Jeans mass increases in this phase, which is the precursor to the phase of

Kelvin–Helholtz contraction onto the stellar main sequence. The minimum opacity-limited Jeans mass is the fragmentation scale. It can be shown, quite insensitively to metallicity or dust content, to be $\alpha_g^{-3/2}$ solar masses, where $\alpha_g = Gm_p^2/e^2$ is the so-called gravitational fine structure constant. This gives a minimum fragment mass of $\sim 0.003\,\mathrm{M}_\odot$, a result that is found in essentially all numerical simulations of current epoch star formation. The dependence on temperature is approximately as $T^{1/4}$, and yields $\sim 0.01\,\mathrm{M}_\odot$ for primordial abundances, appropriate to Population III. In general, fragmentation theory applied to a collapsing interstellar cloud implies that the minimum fragment mass is too small to be a star. Additional physics is needed (Fig. 1).

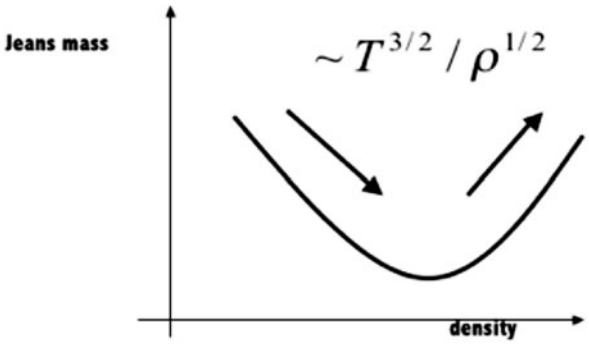


Fig. 1 The Jeans mass (defined to be the mass contained in a sphere of diameter the Jeans length, $\pi v_s/\sqrt{G\rho}$, where v_s is the sound speed in a uniform gas of temperature T), as a function of increasing density ρ. Initially the collapse is isothermal and the Jeans mass decreases. This phase is identified with the process of gravitational fragmentation. Once opacity is important at high density, the collapse is approximately adiabatic and the Jeans mass increases

A key addition is the accretion of cold gas. In the case of a singular isothermal sphere, accretion onto the core occurs at a rate v_s^3/G. In nearby cold molecular clouds, at $T \sim 10\,\mathrm{K}$, the inferred accretion rate is $\sim 10^{-6}\,\mathrm{M}_\odot$/year and yields solar mass protostars on a timescale of order the Kelvin–Helmholtz time. However, in the case of the first stars, the presence of trace H_2 as a coolant means that the temperature $T \sim 10^3\,\mathrm{K}$. Consequently, the accretion rate is $\sim 10^{-3}\,\mathrm{M}_\odot$/year, and one concludes that Population III stars, accreting over $10^5 - 10^6$ year, had characteristic masses of $\sim 10^2 - 10^3\,\mathrm{M}_\odot$. Again, numerical simulations confirm this result.

However, fragmentation and accretion do not suffice to reproduce the initial mass function of stars. A third process must be added, namely feedback, to halt the accretion, otherwise low-mass stars would not form at present. In general, protostellar feedback halts collapse by tapping stellar gravitational energy via releasing magnetic energy. This simultaneously resolves the angular momentum problem, in that there is of the order of two orders of magnitude too much specific angular momentum in cloud cores to form stars directly.

3.1 Feedback: Interstellar Clouds

One requires magnetic feedback to account for the turbulence observed in cloud cores. It simultaneously results in inefficient star formation: were cores to collapse on a free fall time, one would have excessive star formation. Protostellar outflows are ubiquitous and provide momentum input by interactions of jets with the magnetized ISM. This suffices to prolong cloud longevity. In the case of massive clouds, OB stars provide feedback that ultimately disrupts the clouds. One observes over a wide range of molecular cloud masses that the star formation efficiency, defined to be star formation rate divided by gas mass and multiplied by cloud-free fall time, is approximately 2% [7]. In this way one can arrive at a star formation rate for the Milky Way Galaxy that is comparable to what is observed globally. In fact the Milky Way converts about 2% of its molecular gas content (approximately $3.10^9\,M_\odot$) into stars over cloud lifetimes of typically $\sim 10^7$ year.

4 Feedback: Disk Galaxies

It is at first sight rather remarkable that star formation in disk galaxies, both near and far, can be described by a simple law, with star formation efficiency (SFE) being the controlling parameter:

$$SFE = SFR \times ROTATION\ TIME/GAS\ MASS = Constant$$

The motivation comes from the gravitational instability of cold gas-rich disks, which provides the scaling, although the normalization depends on feedback physics. For the global law, in terms of star formation rate and gas mass per unit area, supernova regulation provides the observed efficiency of about 2% which fits essentially all local star-forming galaxies. One finds from simple momentum conservation that

$$SFE = \frac{\sigma_{gas} v_{cool} m*_{SN}}{E_{SN}^{initial}} \approx 0.02$$

This is a crude estimator of the efficiency of supernova momentum input into the interstellar medium but it reproduces the observed global normalization of the star formation law. The fit applies not only globally but also to star formation complexes in individual galaxies such as M51 and also to starburst galaxies. This law is known as the Schmidt–Kennicutt law, and its application reveals that molecular gas is the controlling gas ingredient, and that in the outer parts of galaxies, where the molecular fraction is reduced due to the ambient UV radiation field and lower surface density, the star formation rate per unit gas mass also declines.

For disk instabilities to result in cloud formation, followed by cloud agglomeration and consequent star formation, one also needs to maintain a cold disk by accretion of cold gas. There is ample evidence of a supply of cold gas, for example, in the M33 group. Other spiral galaxies show extensive reservoirs of HI in their

outer regions. Recent data extend the Schmidt–Kennicutt law to $z \sim 2$. Remarkably an efficiency of 2.5% fits low- and high-redshift star-forming galaxies, with $SFE \propto SFR \times \Omega$, and starburst galaxies also lie on this relation [3]. There is a tendency for ultraluminous starbursts at $z \sim 2$ to have somewhat higher SFE.

5 Luminosity Function of Galaxies

Theory provides the mass function of dark halos. Observation yields the luminosity function of galaxies, usually fit by a Schechter function. Comparison of the two is at first sight disconcerting. One can calculate the M/L ratio for the two functions to overlap at one point, for a mass M^* corresponding to L_*. Define $t_{cool} = \frac{\frac{3}{2}nkT}{\Lambda(T)n^2}$ and $t_{dyn} = \frac{3}{\sqrt{32\pi G\rho}}$. For star formation to occur, cooling is essential and the condition $t_{cool} < t_{dyn}$ guarantees cooling in an inhomogeneous galactic halo where gas clouds collide at the virial velocity. One finds that

$$M^*_{cool} = \alpha^3 \alpha_g^{-2} \frac{m_p}{m_e} \frac{t_{cool}}{t_{dyn}} T^{1+2\beta}$$

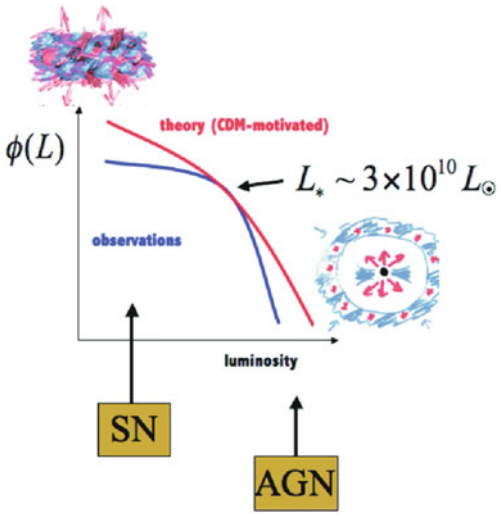


Fig. 2 The galaxy luminosity function: observation versus theory. The luminosity function is described by the Schechter function, an analytic fit to the data on nearby galaxies, whose shape is characterized by a typical luminosity L_*, above which the number density declines exponentially, and a power-law tail proportional to $L^{-\alpha}$, where $\alpha \approx 1$. The mass function is given by the Press–Schechter expression for the nonlinear evolution of self-gravitating spherical overdensities in an expanding universe and is parametrized by a typical mass M_* and low-mass tail proportional to M^{-2}. One can identify a characteristic mass-to-luminosity at one point where the two functions overlap, M_*/L_*

For a cooling function $\Lambda(T) \propto T^{\beta}$, over the relevant temperature range ($10^5 - 10^7$ K), one can take $\beta \approx -1/2$ for a low-metallicity plasma [4]. The result is that one finds a characteristic galactic halo mass, in fundamental constants, to be of order $10^{12}\,M_\odot$. The inferred value of the mass-to-light ratio M/L is similar to that observed for L_* galaxies. This is a success for theory: dissipation provides a key ingredient in understanding the stellar masses of galaxies, at least for the "typical" galaxy. The characteristic galactic mass is understood by the requirement that cooling within a dynamical time is a necessary condition for efficient star formation. However, the theory greatly overestimates galaxy numbers at low and high masses. Feedback is needed to address this problem (Fig. 2).

6 Feedback in Low-Mass Galaxies

Reionization gives an inevitable feedback for the lowest mass dwarfs. An abrupt increase of the sound speed to $\sim 10 - 20\,\mathrm{km/s}$ at $z \sim 10$ means that dwarfs of mass $10^6 - 10^7\,M_\odot$ which have not yet collapsed and fragmented into stars will be disrupted. However more massive dwarfs are unaffected, as are the high σ peaks that develop into early collapsing, but rare, low mass dwarfs. The accepted solution for gas disruption and dispersal in intermediate mass and massive dwarfs ($10^8 - 10^{10}\,M_\odot$) is by supernova feedback. Most gas is ejected by the first generations of supernovae for systems with escape velocity $\lesssim 50\,\mathrm{km/s}$, leaving dim stellar remnants behind. This yields an acceptable fit to the low-mass end of the galaxy luminosity function for the classical dwarfs.

One recent issue has emerged, however. The discovery of ultrafaint, low-mass dwarfs in the MWG halo may be considered as an indication, and even verification, of a prediction of the supernova feedback hypothesis. The only worry is that a model tuned to fit more massive dwarfs such as the Magellanic Clouds is ineffective at dwarf disruption: the overabundance problem remains at the faint end of the luminosity function [6]. Models which fit the massive dwarfs have excessively efficient feedback and overproduce the numbers of ultrafaint dwarfs. This problem seems to be common to all semi-analytic galaxy formation models (SAMs). Conversely, models tuned to the ultrafaint dwarfs cannot account for the Magellanic Clouds. Perhaps the loophole lies in small number statistics: with only two Magellanic Clouds one cannot draw overly far reaching conclusions. A similar comment applies to the M31 halo where there are also just two relatively massive dwarfs.

7 Feedback in Massive Galaxies

Supernovae cannot eject significant amounts of gas from massive galaxies. Baryons continue to be accreted over a Hubble time and the stellar mass grows. The consequences are that massive galaxies are overproduced in the models and that the

massive galaxies are too blue. Moreover the baryon fraction is typically only of order half of the primordial baryon fraction.

A clue as to a solution for these dilemmas comes from the accepted explanation of the Magorrian relation, which relates supermassive black hole mass to spheroid velocity dispersion. This requires collusion between black hole growth and the initial gas content of the galaxy when the old stellar spheroid formed. One conventionally appeals to outflows from the central black hole that deliver momentum to the protogalactic gas. When the black hole is sufficiently massive, the Eddington luminosity is high enough that residual gas is ejected. An estimate of the available momentum supply comes from equating the Eddington momentum with self-gravity on circumgalactic gas shells,

$$L_{Edd}/c = GMM_{gas}/r^2$$

Blowout occurs and star formation terminates when the SMBH–σ relation saturates. This occurs for $M_{BH} \propto \sigma^4$, the observed slope, and gives, at least in order of magnitude, the correct normalization of the relation.

There is also a role for AGN feedback at late epochs, when the AGN radio mode heats halo gas, inhibits cooling, resolves the galaxy luminosity function bright end problem, and accounts for the red colours of massive early-type galaxies. AGN feedback in the radio-quiet mode may also account for the suppression in numbers of intermediate mass and satellite galaxies. Feedback from AGN in the host galaxies preheats the halo gas that otherwise would be captured by satellites.

However, reality may be not quite so simple. A more detailed examination suggests that negative feedback in momentum-driven winds by supermassive black holes falls short of explaining the observed $M_{BH} - \sigma$ correlation by a factor of a few [9]. Moreover comparison of baryonic fractions with bulge-to-disk ratios in nearby galaxies demonstrates that AGN alone do not eject significant amounts of baryons [1]. Something else seems to be needed.

7.1 The AGN–Star Formation Connection

The most plausible addition to the physics is inclusion of star formation, induced and enhanced by the SMBH outflows. If AGN-driven outflows trigger star formation, the star formation rate is boosted by a factor t_{dyn}/t_{jet}, and the outflow momentum is amplified by supernovae [8]. Consequently, the momentum supplied to the gas is boosted by the combination of AGN and star formation. There is extensive evidence, recently compiled by Netzer, that demonstrates the intimate connection of AGN luminosity and star formation rate over a wide dynamic range. Of course the causal direction is uncertain and indeed the phenomena could be mutually self-regulating. To go beyond phenomenology, many details need to be refined, the most pressing perhaps being the nature of the black hole growth. However, there are examples of jet-induced global star formation, as seen locally in Minkowski's object, and jet-induced CO formation (and excitation) at high redshift. CO is a prerequisite

for star formation and has been detected in large amounts in the host galaxies of high-redshift quasars.

7.2 Modes of Star Formation

Incorporation of a positive role by AGN for star formation in extreme environments leads one to argue that a case can be made for two distinct feedback-regulated modes of star formation: at low redshift via supernovae and without AGN and at high redshift with triggering by AGN playing a central role. One would expect a transition between these two modes as the AGN duty cycle becomes shorter beyond $z \sim 1$. Indeed a recent compilation [5] of the specific star formation rate (SSR or star formation rate per unit stellar mass) to $z \sim 7$ in the GOODS field suggests that the star formation timescale (or 1/SSR) goes from the MWG value of $\sim$ 3Gyr at low redshift to $\sim$ 0.5Gyr at $z \gtrsim 2$.

There are two other transitions in this redshift range that may be relevant. At high redshift, major mergers between galaxies are common. Indeed the high-redshift ULIRGs are invariably undergoing major gas-rich mergers. Theory suggests that at low redshift, gas accretion by cold streams is important and that the cold streams are invariably clumpy and essentially indistinguishable from minor mergers of gas-rich dwarfs. In terms of the cosmic star formation history, normal star-forming galaxies dominate at low redshift ($z \lesssim 2$) whereas ULIRGs dominate at high redshift ($z \gtrsim 2$).

If the disk formation mode is distinct from the spheroid formation mode, then SMBH might be expected to show some reflection of alternative growth histories. So-called pseudobulges form from secular instability of disks and contain smaller SMBH than do the more massive bulges that may have formed via major gas-rich mergers. It is interesting that SMBH in pseudobulges lies low on the Magorrian relation [Kormendy and Tremaine, Private Communication], possibly reflecting the different black hole formation histories and the associated distinct star formation modes. Recent data on $z \sim 6$ quasars suggest that the most massive black holes indeed lie high on the black hole/dynamical mass relation.

Much work needs to be done to see whether allowance for two modes of star formation can help resolve some of the outstanding problems in galaxy formation. Perhaps the greatest challenge in any combination of cold stream/minor merger/major merger scenario for gas delivery to drive both star formation and SMBH feeding is that $\sim$ 16% of nearby galactic disks are bulgeless. In addition to the many uncertainties in star formation theory (and I have not addressed one of the key issues, that of the IMF), there remains the nature of black hole growth. Whether the black holes grow by gas accretion, in which case feedback may play a role in angular momentum transfer [2], or by mergers, or by an appropriate combination, remains unresolved (Fig. 3).

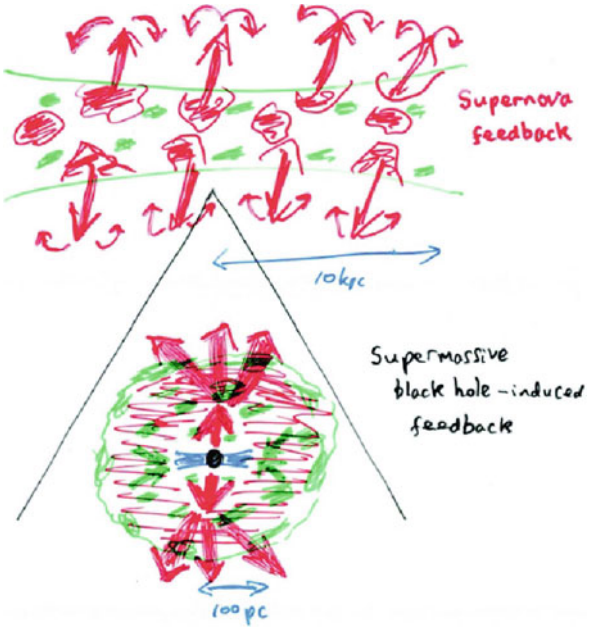


Fig. 3 The case for two modes of galaxy formation. Supernovae drive turbulence and fountains in star-forming disk galaxies and are responsible for the low global star formation efficiency. Supermassive black holes quench star formation, accounting for the redness of elliptical galaxies, and may also play a role in triggering star formation in starbursts, thereby enhancing both the star formation efficiency and the specific star formation rate

8 Conclusions

In summary, I have argued that cold gas flows via filaments/minor mergers lead to disk and bulge formation. Supernovae drive turbulence and fountains in star-forming disk galaxies and are responsible for the low global star formation efficiency.

Major mergers along with hot gas infall followed by cooling form massive spheroids at high efficiency. The role of SMBH is to quench star formation at early epochs, thereby accounting for the redness of elliptical galaxies, to heat intracluster gas at late epochs, thereby preventing gas cooling and late star formation. Of course this process cannot be completely efficient and indeed 30% of nearby ellipticals have modest amounts of ongoing star formation. More speculatively, SMBH may also play a role in triggering star formation in starbursts, especially in ULIRGs. Positive SMBH feedback can enhance both the star formation efficiency and the specific star formation rate.

The origin of SMBH remains a mystery and must certainly play a key role in ascertaining the detailed nature of SMBH feedback. Improved resolution in theory

and observation is needed. The great projects of the future, including the ELTs, JWST, and LSST, will surely play key roles in this endeavour.

Acknowledgments I have benefitted enormously and learnt much from my interactions with Ken Freeman. He is one of those rare individuals who is genuinely interested in whatever crazy idea you might throw at him and which he will take apart with forensic discipline. He reduces such notions to their bare essentials, which rarely might contain a hint for a future project that he alone can discern and design. Thanks, Ken, and happy birthday!

References

1. Anderson, M., Bregman, J. 2010, ApJ, 714, 320
2. Antonuccio-Deloglu, V., Silk, J. 2010, MNRAS, 405, 1303
3. Genzel, R., et al. 2010, MNRAS, in press, arXiv1003.5180
4. Gnat, O., Sternberg, A. 2007, ApJS, 168, 213
5. Gonzalez, V., et al. 2010, ApJ, 713, 115
6. Koposov, S., et al. 2009, ApJ, 696, 2179
7. Krumholz, M., Tan, J. 2007, ApJ, 654, 304
8. Silk, J., Norman, C. 2009, ApJ, 700, 262
9. Silk, J. and Nusser, A. 2010, ApJ, in press, arXiv1004.0857

When Bad Masks Turn Good

Roberto G. Abraham

Abstract In keeping with the spirit of a meeting on 'masks,' this talk presents two short stories on the theme of dust. In the first, dust plays the familiar role of the evil obscurer, the enemy to be defeated by the cunning observer in order to allow a key future technology (adaptive optics) to be exploited fully by heroic astronomers. In the second story, dust itself emerges as the improbable hero, in the form of a circumstellar debris disks. I will present evidence of a puzzling near-infrared excess in the continuum of high-redshift galaxies and will argue that the seemingly improbable origin of this IR excess is a population of young circumstellar disks formed around high-mass stars in distant galaxies. Assuming circumstellar disks extend down to lower masses, as they do in our own Galaxy, the excess emission presents us with an exciting opportunity to measure the formation rate of planetary systems in distant galaxies at cosmic epochs before our own solar system formed.

1 Bad Dust – The Struggle of Extragalactic Adaptive Optics

Our understanding of the high-redshift universe has been revolutionized by deep fields, several of which have been extensively surveyed at all accessible wavelengths. See Fig. 1 for an up-to-date summary of existing deep fields. These fields have been used to study galaxy formation and evolution out to very high redshifts. Because galaxies at such high redshifts are typically < 1 arcsec in size, kinematical investigations require adaptive optics (AO) spectroscopy. The promise of such observations has been held out as an exciting next step for many years (see, for example, the many references in van Breugel & Bland-Hawthorn [28]). Unfortunately, it is now clear that only very limited AO observations are ever going to be undertaken in any existing deep field. Why not?

R.G. Abraham (✉)
Department of Astronomy and Astrophysics, University of Toronto, Toronto, ON L6M 2N5, Canada
e-mail: abraham@astro.utoronto.ca

D.L. Block et al. (eds.), *Galaxies and their Masks*,

DOI 10.1007/978-1-4419-7317-7_36,

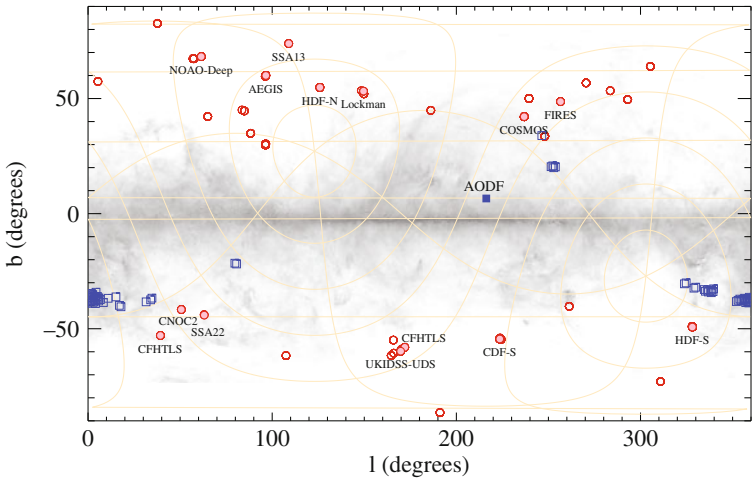


Fig. 1 Galactic dust map with the positions of existing deep fields shown in *red*. Note how all of these fields lie away from the galactic plane, with most concentrated at the poles. Existing deep fields lie in regions where extinction is low, but at latitudes where there are not enough stars for efficient AO. In contrast, the field with the best properties for adaptive optics (and described in detail in Damjanov et al. [7]) is labeled 'AODF' (for adaptive optics deep field). This field lies very close to the plane of Galaxy (where the density of 13–16 mag stars is high) but still has very low reddening coefficient ($E(B-V) < 0.1$). The AODF corresponds to a small $\sim 1\,\mathrm{deg}^2$ hole in the the filamentary structure. The hole itself is invisible at this resolution, but it is described in detail in the text. Other fields explored by Damjanov et al. are shown in *blue*. (This figure was kindly generated by Jarle Brinchmann and is based on data available here: http://www.strw.leidenuniv.nl/~jarle/Surveys/DeepFields/)

The fundamental problem preventing us from exploiting the Deep Fields shown in Fig. 1 with adaptive optics is that even with a laser guide star AO system, one still needs a reasonably proximate natural guide star for the tip-tilt correction. As a result, all existing deep fields are barely accessible to AO. This is because two of their main selection criteria have been that they contain as few bright stars as possible to avoid saturation in long exposures and that they lie in regions of low galactic extinction. In other words, to avoid the *bad dust*. Thus all existing deep fields are at high galactic latitudes, where the density of suitable guide stars is at a minimum. For example, Davies et al. [8] report that only 1% of the Lyman break galaxy sample of Mannucci [19] are accessible to the VLT laser guide star system (after sensible line flux and luminosity cuts). The situation is similar with Gemini: again, only about 1% of the galaxies in the Gemini Deep Deep Survey were targetable with the existing AO system. The Keck LGS system has more forgiving tip-tilt constraints, but the accessible sky fraction in deep fields remains small. Even with the upcoming Gemini Multi-Conjugate AO system (MCAO), the H-band sky coverage at the galactic poles will only be around 15% for a *single* natural guide star and the system will have very marked improvement in mean Strehl ratio in fields where more than the minimum number of natural guide stars are available (see

http://www.gemini.edu/sciops/instruments/adaptiveOptics/MCAOPerformance.html). This is because the geometry of the guide stars on the sky impacts the uniformity of Strehl [13], so a constellation of multiple laser beacons performs much more effectively when augmented by multiple natural guide stars. We have modeled this effect with real catalogs and show below that there are huge benefits for next-generation AO from having multiple guide stars.

Do astronomers really have to put up with this? Perhaps not. One way forward has been suggested by Damjanov et al. [7], who propose that 'AO-friendly' patches on the sky be developed in order to allow efficient AO-based extragalactic observations to be undertaken in the future. These authors identify a number of patches on the sky which have the rare combination of both high stellar surface density ($\geq$ 0.5 arcmin^{-2}) and low extinction ($E(B-V) \leq 0.1$). The basic strategy for identifying these fields is illustrated in Fig. 2. One simply divides an all-sky map of stellar counts (in a suitable magnitude range optimized for AO guide star usefulness) by an extinction map. The corresponding maxima on the resulting map correspond to potential sites for excellent AO performance on extragalactic targets, which we define as sites having enough natural guide stars to let the telescopes really perform to their design limits, but without having very much obscuring dust in the field. There are quite a few maxima in Fig. 2 . . . are they all good? Sadly, no. Some simply correspond to patches of sky with abnormally low extinction without many suitable

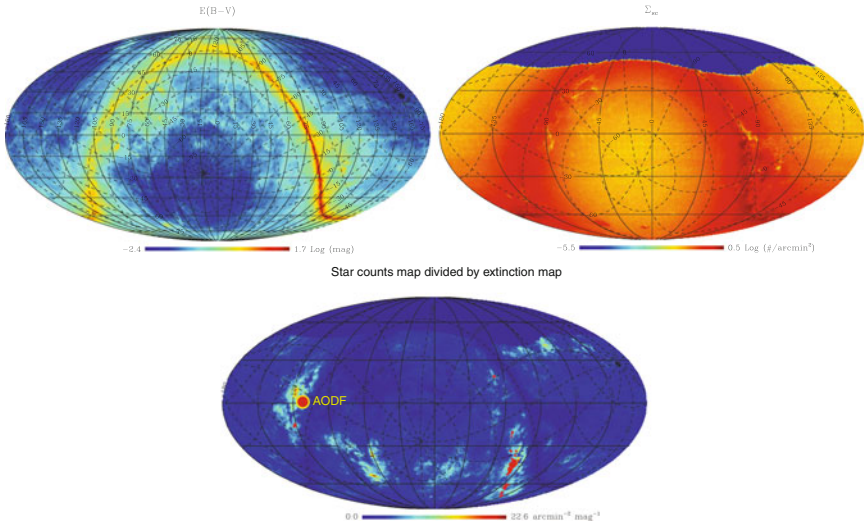


Fig. 2 *(Top left:)* All sky map of extinction, scaled logarithmically. The *solid line grid* corresponds to the celestial coordinate system with RA in degrees increasing to the left. Zero degrees lies at the center of the figure. A galactic coordinate system is overploted with *dashed lines*. *(Top right:)* The corresponding map of star count surface density for stars in the USNO UCAC2 catalog. *(Bottom:)* A map constructed by multiplying the map at the *top left* by the inverse of the map at the *top right*. Maxima in this figure correspond to potentially interesting locations for undertaking extragalactic adaptive optics observations. The location of the preferred field is shown with a *large solid red circle* and labeled "AODF." See text for details

guide stars, and (annoyingly) some correspond to regions near the celestial poles that do not make for practical fields accessible from both hemispheres. So it turns out one has to quite carefully explore the properties of all the maxima individually, which is done in Damjanov et al. [7]. An example of the sorts of field-to-field variation seen is given in Fig. 3, which compares the properties of various maxima in Fig. 2 to the corresponding properties in various existing deep fields. After doing this sort of analysis ad nauseam it turns out that one location in the sky stands out as as being particularly good. Thus a 1 deg^2 area of sky is centered at RA 7 h 24 m 3 s, Dec -1° 27′ 15″. This equatorial field is well-placed for both hemispheres, has abundant natural guide stars, few very bright stars (light from whom can scatter in optics and reduce signal to noise), has low−absolute extinction and has low variance in extinction throughout the field. We are somewhat grandiosely calling this field the adaptive optics deep field and have shown its location in both Figs. 1 and 2.

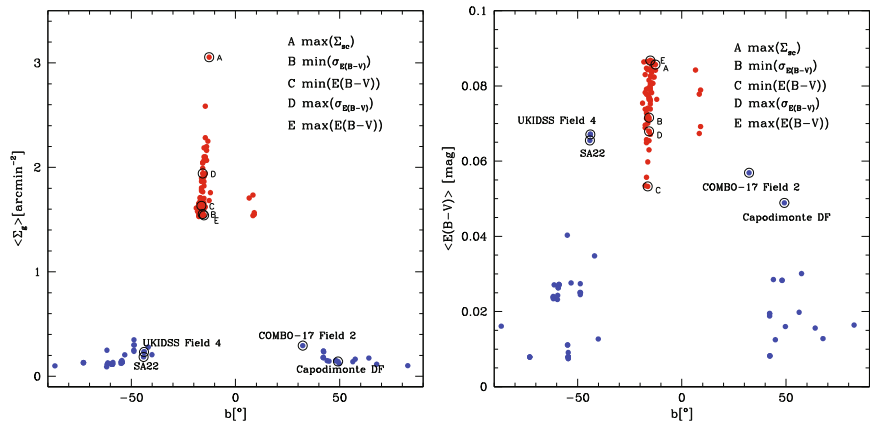


Fig. 3 Reddening coefficient $E(B-V)$ as a function of the star counts surface density Σ_{sc} for hundred and sixty five 1 deg^2 fields with $\Sigma_{sc} > 0.5$ arcmin^{-2} and $E(B-V) \leq 0.1$. The fields are *color-coded* based on their equatorial coordinates. The fields flagged with *open circles* have the highest star counts surface density or lowest standard deviation of the reddening coefficient. *Colored arrows* point at the representative fields for each of the three sightlines. (Figure taken from [7])

Some CFHT Director's Discretionary Time has been invested in discovering whether the properties of the sky at RA 7 h 24 min 3s, Dec -1° 27′ 15″ really are as salutary as claimed. Figure 4 shows a 'true color' image constructed from two 10 min snapshot in g′ and z′ of the field, obtained with CFHT's MegaCam instrument. This image is essentially a sanity check. The image shows the 1° field-of-view (made out of five dithered exposures) and it is apparent here that bright stars are indeed sparse. In fact, Fig. 5 shows that very bright stars in this field are sufficiently sparse that only about 3% of the field is contaminated by bright star halos. Figure 6 shows the field is rich in field galaxies, even in a 10 min integration almost directly through the plane of the Milky Way! At a more quantitative level Fig. 7 shows that

Fig. 4 A 'true color' image constructed from two 10 min snapshot in g′ and z′ of the adaptive optics deep field, obtained with CFHT's MegaCam instrument to check on the degree of stellar contamination in the data. Data were obtained in March 2010 through thin clouds to evaluate the distribution and color of the brightest stars in the field. The *image* shows the 1° field of view (made out of five dithered exposures) and it is apparent here that bright stars are indeed sparse. The most prominent one being in the *center* (mag. 9.9)

the number counts of faint field galaxies in z-band is exactly as expected from the corresponding number counts in the CFHT Legacy Survey (which, of course, was conducted at high galactic latitude to avoid dust contamination).

The results presented so far show that the 1 deg^2 patch of sky at RA 7 h 24 min 3 s, Dec $-1°\ 27'\ 15''$ has about 10 times the stellar density of existing deep fields coupled with extinction values at the high end of those in existing deep fields. But does this really make such a big difference to adaptive optics systems? How this translates into practical performance benefits is presented in Fig. 8, in which a multi-conjugate adaptive optics simulator has been used to compute the distribution of a figure of merit (a combination of peak Strehl ratio and uniformity in Strehl; see figure caption of Fig. 8) in one of the proposed fields, as well as in an existing deep field (COMBO-17). Simulations are presented for both an 8 m telescope with a next-generation AO system (i.e., multi-conjugate adaptive optics) and a 30 m ELT. The improvement in efficiency obtained by undertaking AO in the preferred field is astonishing. In fact, the improvement in ELT performance is rather surprising, given the standard assumption that such telescopes will be able to use much fainter natural guide stars. The explanation for the improvement with ELTs lies in the fact that the figure of merit weights not only the peak Strehl, but also the uniformity

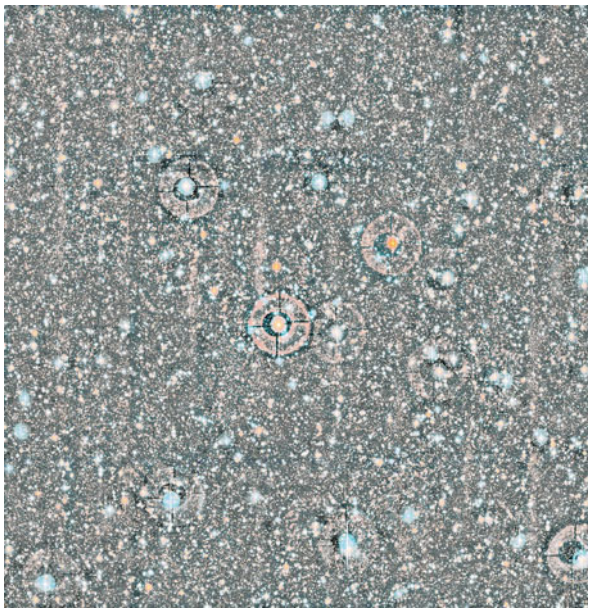


Fig. 5 A highly contrast stretched version of the previous figure used to explore the fraction of the image contaminated by halos due to internal reflections in the MegaPrime optics. There are only five stars that cause a potential problem. Each halo covers a disk of < 3 arcmin radius, resulting in 28 arcmin^2 affected by each bright star, leading to ~ 120 arcmin^2 over the whole field affected which is to be compared to the MegaCam field of view of $\sim 3{,}300$ arcmin^2. Contamination is less than 4%, which is deemed a negligible loss

Fig. 6 A zoom-in of a small portion of Fig. 4. As expected, the image is dense with field galaxies, even on short (10 min) exposures

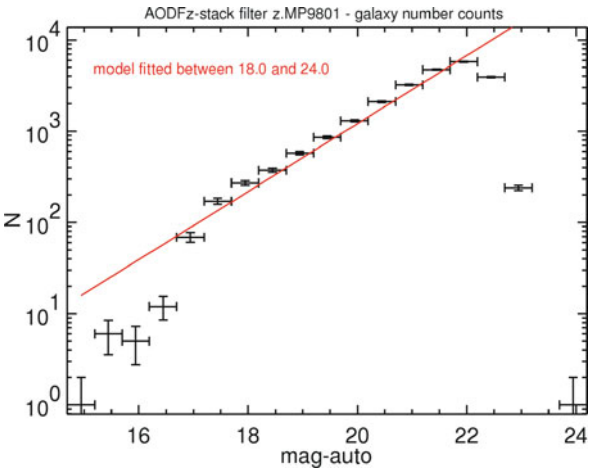


Fig. 7 Galaxy number counts in the adaptive optics deep field as function of magnitude in the z-band, based on the short 10 min integrations shown in the previous figures. The *red line* shows the expected counts based on the CFHT Legacy Survey (obtained with the same MegaCam filters), obtained at high galactic latitude. The galaxy counts are completely nominal, even though the field is nearly in the plane of the galaxy

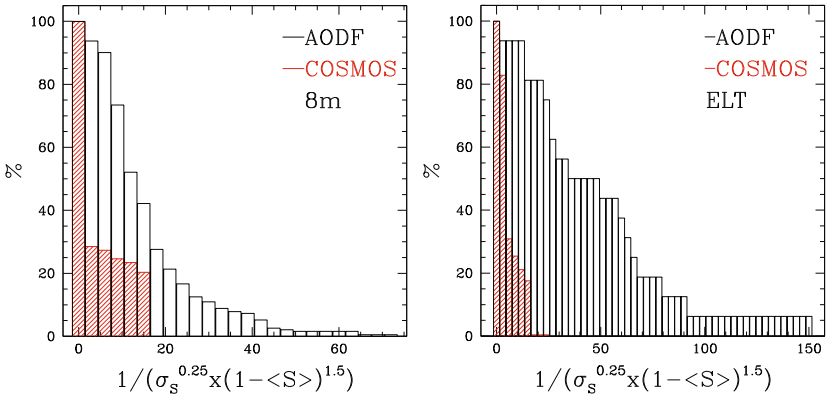


Fig. 8 A comparison of the performance of a multi-conjugate adaptive optics system in the proposed field and the COSMOS field, based on a figure of merit which combines peak Strehl and uniformity of Strehl. The histograms are computed from 192 random $2' \times 2'$ pointings in each field. Cumulative distributions are shown, so each bin shows the percentage of fields in which the figure of merit is larger than the specified value. Histograms were computed using a search algorithm to choose optimal natural guide stars which are then fed into the Gemini multi-conjugate adaptive optics (MCAO) simulator. (*Left*) Distributions for an 8 m telescope with MCAO parameters identical to Gemini GEMS used in our proposed field (*solid black line*) and the COSMOS field (*hatched red line*). (*Right*) The corresponding distribution for a 30 m ELT in the COSMOS field. Note that, unlike the histogram on the left, this is only intended to show qualitative trends (obviously, we do not have a real MCAO system on a 30 m telescope to model), but it does make the point that on ELTs the natural guide star problem will remain nearly as severe as for 8–10 m class telescope. See text for details

of the Strehl. As noted in the Science Justification Strehl uniformity comes from having multiple natural guide beacons (the minimum desirable is three). With a 30 m ELT the sky coverage over which a single suitably bright guide star is available is high ($>$ 50% even at the poles) but the area over which three bright natural guide stars exist to feed an multi-conjugate adaptive system is still low near the poles.

The next big question is, even if we defeat the bad dust and focus efforts in doing extragalactic adaptive optics in this preferred field, how big a win is it? I think this depends largely upon your degree of faith in how important AO will be to the future of large telescopes. AO is a key system feature of all proposed future 20–30 m class telescopes (TMT/E-ELT/GMT). Although the natural guide star problem is less severe for these extremely large telescopes (ELTs), our analysis shows that the problem remains significant, particularly if multiple natural guide stars are needed (see below). These ELTs are designed to operate synergistically with ALMA and JWST. For example, many of the key projects described in the JWST Design Reference Mission rely on either extreme depth or serendipitous lines-of-sight (e.g., strong gravitational lensing). *If such JWST observations are to be synergistic with ground-based AO follow-up, they cannot be undertaken efficiently in any existing deep field*. It would be regrettable if only 1–10% of rare targets discovered with JWST in a deep field could be followed-up with a ground-based IFU. It is becoming clear that existing and planned AO systems are set to enable transformative high-redshift science, but they will do so only in the regions of the sky in which they are effective. And, as noted above, *no existing deep field is suitable for efficient AO*. So I think we need a new deep field.

2 Good Dust – A Surprise at High Redshifts

In the second half of this chapter I would like to switch gears completely in order to talk about 'good dust.' What I mean by this is that I want to step away from the common bias (shown in the previous section) of viewing dust as some sort of mask preventing us from studying what we really want to study, and instead focus on viewing the dust as the extraordinarily rich source of information that it really is. In particular, I would like to explore the idea that a natural goal for future spectral synthesis modeling will be to extend such models to incorporate contributions from dust. Of course I am aware of the many difficulties in modeling dust and recognize that the uncertain physics of dust production and destruction makes modeling dust in any detail a huge challenge. However, there certainly is considerable scope for incorporating dust into models in a number of fairly simple ways, and we really ought to be doing this already, if only for the sake of consistency. And, ultimately, I am driven to conclude that we need to improve dust models because the clear trajectory of future observations is toward longer wavelengths, so we need to get with the program to better handle the flood of data that is soon going to be coming our way.

As an example of the sort of basic modeling I think we should be doing already, let me note that most models treat dust extinction in the UV and dust emission in the infrared independently. There are a few notable exceptions, such as the *Sunrise* models of Jonsson [15] and the work presented in Riffel et al. [23] using the *STARLIGHT* code [2, 5, 20]. But on the whole most modeling work does not use the mid-infrared properties of galaxies to inform the choice of the UV extinction curve being used, despite the energetic connection between these components. At a very basic level, which is all I am really advocating we consider here, there is really no excuse for this. We have known for nearly 20 years (since the classic paper of [10]) that the big grains responsible for the linear rise in the near-UV spectral slope are also responsible for the thermal far-IR emission and that the polycyclic aromatic hydrocarbon molecules responsible far-UV non-linear UV extinction are also responsible for the mid-IR emission lines and also that whatever mysterious grains are responsible for the 2,175 Å bump in the UV extinction curve also correspond to an additional component in the mid-IR emissivity peaking at around 60 μm. Knowing this, there is no reason why the mid-IR properties of galaxies ought not be used to place some basic constraints on the choice of UV extinction curve, at least at the simple level of putting limits on the far-UV slope and on the strength of the 2,175 Å bump.

I believe that incorporating dust into spectral synthesis models can yield some interesting surprises. A nice example of this is given in an investigation led by my student Erin Mentuch [21], who looked at the UV/Visible/NIR/Mid-IR colors of 88 $0.5 < z < 2$ galaxies taken from the Gemini Deep Deep Survey [1]. Detailed modeling of these colors (see Fig. 9) shows clear evidence for a near-infrared excess at around 3 μm, which at the redshifts of these galaxies is seen in the IRAC [5.8 μm] and [8.0 μm] bands. Figure 10 shows that this excess can be modeled as an additional SED component consisting of a modified 850 K graybody augmented with a mid-IR PAH emission template spectrum, as suggested by da Cunha et al. [6]. The luminosity of the excess SED component is correlated with the star-formation rate of the galaxy, so the excess shows some promise as an extinction-free star formation tracer. Hints that the excess is correlated with star-formation activity and morphology can be seen in HST imaging data (e.g., the examples shown in Fig. 10). But the main interest of the excess lies in the interpretation of its origin. The five best candidates for excess emission are

1. AGN.
2. The high-redshift counterpart to the interstellar cirrus emission seen in our own Galaxy [4, 12].
3. Reflection nebulae [24–26].
4. Post-AGB stars/planetary nebulae [9, 22].
5. Protostellar/protoplanetary disks in massive star-forming regions [14, 16–18].

Mentuch et al. [21] come down firmly in favor of the last candidate, in effect attributing the excess light to the collective emission from the thousands of flared circumstellar disks around massive stars in these galaxies at high redshifts. AGN can largely be ruled out on the basis of IRAC color–color diagrams for the galaxies, though a much more convincing argument against them will be given below. Cir-

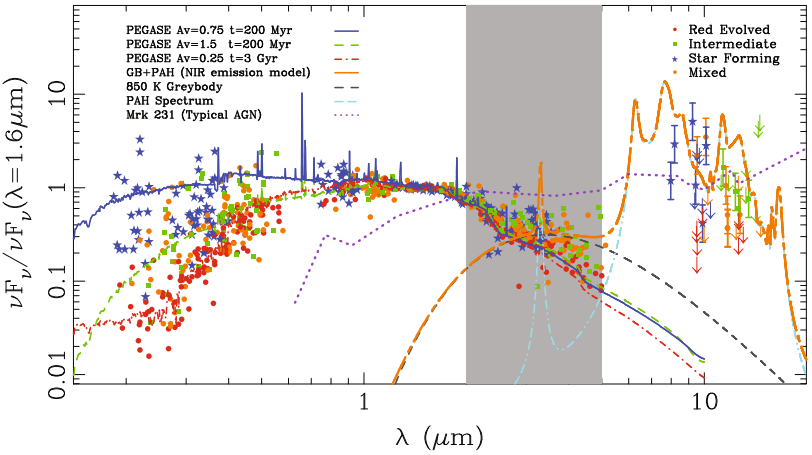


Fig. 9 Figure taken from Mentuch et al. [21] showing VIzK$_s$+IRAC and MIPS photometry from the galaxies in the Gemini Deep Deep Survey, normalized at $\lambda = 1.6\,\mu$m with the 2–5 μm region of NIR excess highlighted in gray. Plot symbols are keyed to galaxy spectral type, as shown in the legend. The *blue solid curve* is a stellar component model from PEGASE.2 for a typical star-forming galaxy (t=200 Myr, A$_v$=0.75). A dusty star-forming galaxy (t=200 Myr, A$_v$=1.5) and evolved galaxy (t=3 Gyr; A$_v$=0.25) are also plotted. The *orange dash-long-dashed line* represents the NIR emission component model (GB+PAH) from da Cunha et al. [6] that is supplemented in the SED modeling. An 850 K graybody emitter (*gray dashed line*) contributes most of the emission of this component at 2–5 μm, although the 3.3 μm PAH feature (the *cyan dot-dashed curve*) has some contribution ($< 20\,\%$ in a given IRAC band). For reference, an SED of an AGN (Mrk 231; *purple dotted curve*) is plotted. For pedagogical purposes, the GB+PAH and AGN curves have been normalized to the average MIPS flux of the sample. For those objects with MIPS detections, the AGN model is not a good representation of the photometry

rus, reflection nebulae, and post-AGB stars can be ruled out on the basis of simple scaling relations which show the predicted contributions from these objects are more than an order of magnitude too low to explain the excess. So, in a sense, circumstellar disks are the only candidate that remains standing after the others are eliminated.

A simple model showing that the interpretation of the NIR excess as the collective emission from flared circumstellar disks is easy to concoct (Fig. 11), starting from published flared disk models. The example shown in Fig. 11 is based on the model of Dullemond et al. [11], who calculate the NIR excess fluxes due to circumstellar flared disks to be $L_{\rm NIR} = (310, 8, 0.1, 0.0035)\ L_\odot$, for stars of spectral types B2, A0, G2 and M2, luminosities of $L = (1000, 500, 1, 0.5)\ L_\odot$ and masses of $M = (10, 4, 1, 0.4)\ M_\odot$. Using the results of their Table 1, the NIR excess due to the flared disk is related to the mass of the star by

$$L_{\rm NIR}(M) = 0.0865\, L_\odot \left(\frac{M}{M_\odot}\right)^{3.48} \quad (1)$$

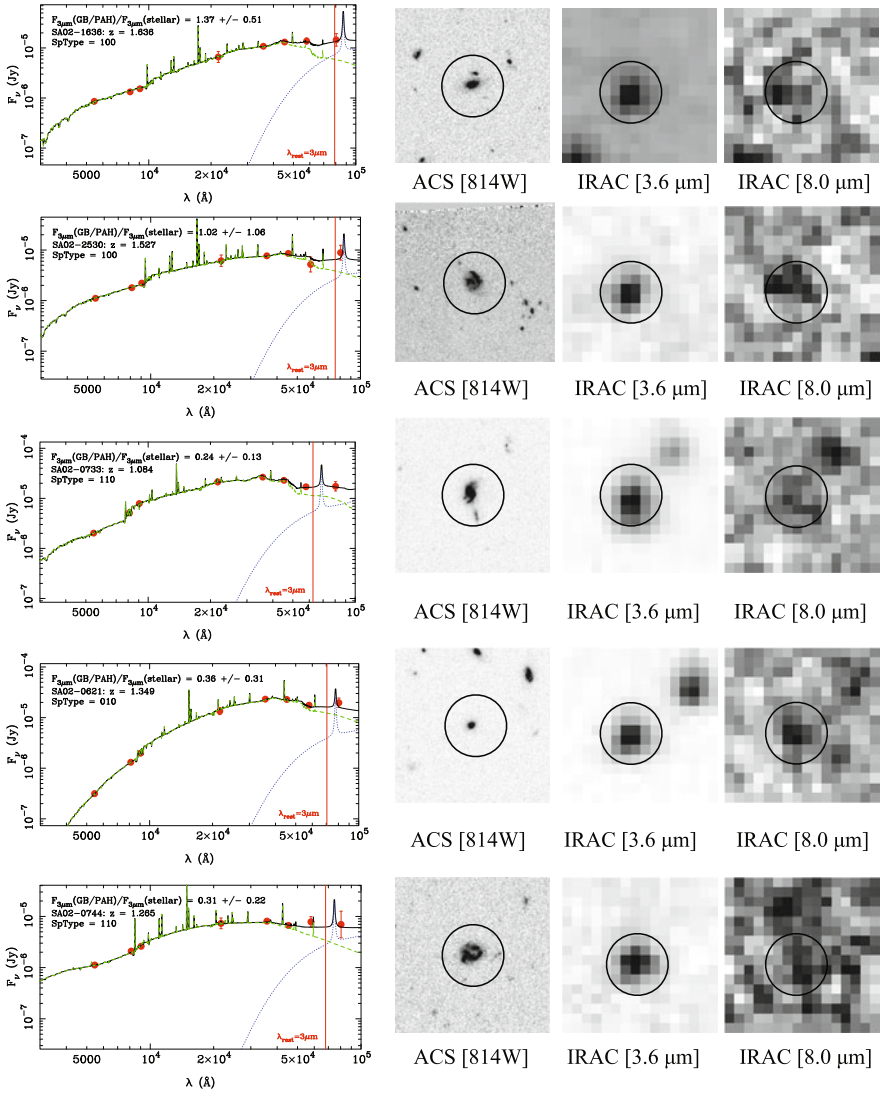


Fig. 10 Figure taken from Mentuch et al. [21] showing spectral energy distribution fit examples for objects (band fluxes shown as *red dots*) with SFRs $> 1\,M_{\odot}$/year. The SED (*solid black line*) is composed of both a stellar component (*green dashed*) and an NIR emission component consisting a 850 K graybody emission and a PAH spectrum (*blue dotted line*). About 10 arcsec postage stamps from *Hubble's* ACS [814W], IRAC [3.6 μm] and [8.0 μm] are shown for each object. The spectral type of the galaxy is given in the *top left corner* showing the excess is seen in star-forming, intermediate, and mixed population galaxies. For reference, $\lambda_{\rm rest} = 3\,\mu$m is plotted as a *red vertical line*. The *black circles* show the 4 arcsec apertures used to derive our photometry

We can integrate this over the IMF to figure out the total luminosity excess in the NIR for a given amount of mass formed during a period of star formation. Using the BG03 IMF [3], we find the NIR excess due to circumstellar disks is $(L/M)_{\mathrm{NIR,disks}} = 500\,L_{\odot}/M_{\odot}$. While this seems like a high light-to-mass ratio, it is in fact very typical for very young ~ 1 Myr star-forming regions, which output high IR luminosities.

One can then easily estimate the total NIR excess luminosity density at 3 μm for a galaxy forming stars at a rate of SFR, with only one free parameter, namely the disk lifetime, t_{excess}, which is thought to be around 1 Myr (so in practice that is what we fix it at):

$$L_{\mathrm{NIR}}(\mathrm{SFR})_{\mathrm{disk}} = 500\,\mathrm{L}_{\odot}/\mathrm{M}_{\odot}\left(\frac{\mathrm{SFR}}{\mathrm{M}_{\odot}/\mathrm{year}}\right)\left(\frac{\mathrm{t_{excess}}}{\mathrm{year}}\right) \tag{2}$$

Notice how we have 'closed the circle' and are using the UV properties of galaxies to infer IR properties, in the manner advocated at the start of this section, because the star-formation rates injected into this equation are obtained from fits to the panchromatic spectral energy distributions of galaxies.

As Fig. 11 shows, this simple flared disk model does a surprisingly credible job of explaining the excess emission. With essentially no 'tuning,' the simple model goes straight through the data points. With the exception of AGN, all other candidates for the excess emission require unrealistic abundances or extremely high duty cycle star

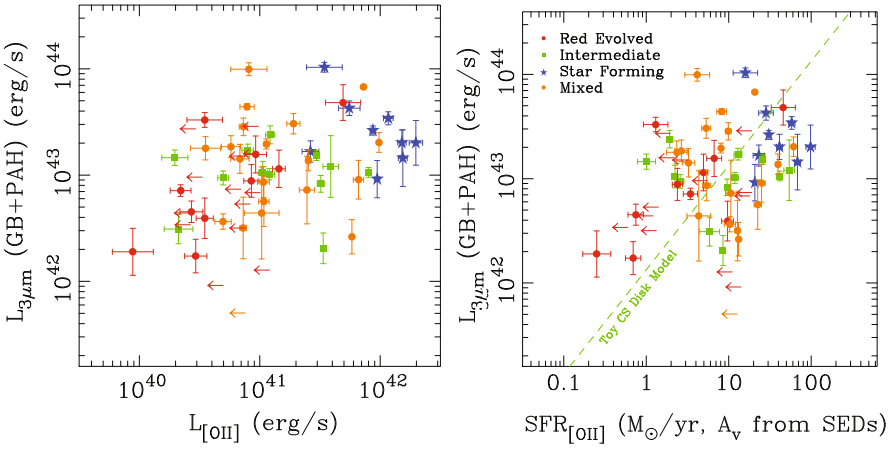


Fig. 11 A figure (from [21]) showing the luminosity of the 3 μm excess (modeled as a gray-body+PAH component) is correlated ($r = 0.49 \pm 0.06$) with the [OII] line luminosity (*left panel*) and the star-formation rate (*right panel*). The *dashed green line* is the luminosity expected from an NIR excess due to circumstellar disks (that last for 1 Myr) in massive star-forming regions, assuming a simple model that scales linearly with the SFR ([21] for details). The total luminosity of this model matches the galaxy's total luminosity in the NIR. Galaxies are indicated by spectral types as star-forming (*blue stars*), evolved (*red circles*), intermediate (*green squares*), and mixed (*orange circles*)

formation to even get in the ballpark. Of course, the Dullemond et al. [11] model is certainly not unique and one does not wish to push the agreement shown in Fig. 11 too far. However, the figure does indicate that the circumstellar disk interpretation is at least plausible.

How firmly can one rule out AGN as being the simple explanation for the NIR excess? Mentuch et al. [21] argues that 11% of the objects from the selected GDDS sample fall in the portion of the Stern et al. [27] IRAC color–color diagram appropriate to AGN and most of these are right on the selection boundary. Nonetheless, it would be better to rule out AGN more directly. It turns out that this is actually fairly straightforward to do at low redshifts using data from the SINGS survey, which has resolved UV–IR imagery for a sample of nearby objects. Figure 12 is one example galaxy from this sample, NGC6946. The top row shows the construction of a $J - [3.6]$ color map and it is straightforward to identify the spatial locations of sites of NIR excess on this map and enquire whether the excess is occurring in an active nucleus. In this case the answer is certainly 'no,' and it turns out that is true for most galaxies in the SINGS sample (Fig. 13).

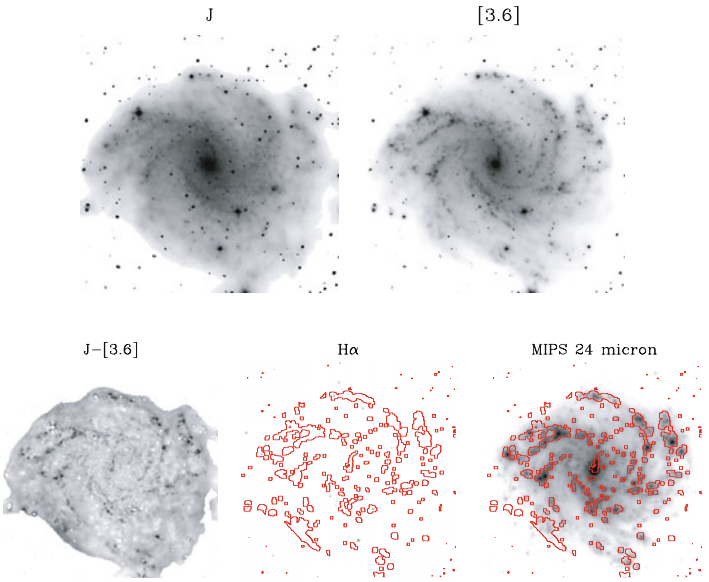


Fig. 12 An analysis of the sites of excess 3 μm emission in NGC6946 based on data from the SINGS sample. The *top row* shows the the J-band image (*left*), [3.6] μm image (*middle*), and $J - [3.6]$ color map (*right*). Regions of excess [3.6] emission from this color map are thresholded to determine where on the galaxy the regions are located. In this case it is clear that the regions of excess emission are located near the galaxy's nucleus. Excess regions are shown as *red contours* on subsequent plots. In the *bottom row* regions of excess [3.6] emission are shown superposed on Hα and 24 μm maps. By determining the fraction of Hα and 24 μm emission within the *red contours* one can estimate the fraction of excess [3.6] emission that originates in sites of star formation

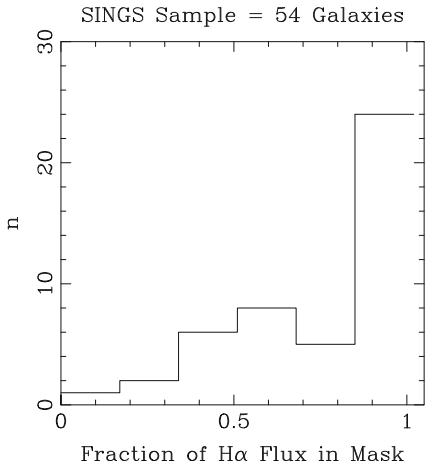


Fig. 13 A histogram showing the distribution of the fraction of a galaxy's total Hα flux contained within sites of excess [3.6] μm emission for 54 galaxies in the SINGS sample. It is clear that in the vast majority of cases the [3.6] μm emission originates in sites of active star formation

We conclude that the most likely explanation for the 2–5μm excess seen in Fig. 9 is the contribution from thousands of flared circumstellar disks around massive young stellar objects seen in the integrated light of these high-redshift galaxies. It seems natural to suppose that the presence of circumstellar disks around massive stars at high redshifts would also imply the presence of disks around less massive stars and of course it is around these less massive systems that we would expect planets to form. Therefore this 3μm excess might present us with an opportunity to probe the formation of planets (as seen in their total integrated light) at cosmic epochs even before our own solar system formed. Perhaps the most interesting follow-up measurement from a cosmological standpoint would be the measurement of something like the cosmic evolution of the volume-averaged planet formation rate density.

References

1. Abraham, R.G., et al. 2004, AJ, 127, 2455 (Paper I)
2. Asari, N.V., Cid Fernandes, R., Stasińska, G., Torres-Papaqui, J.P., Mateus, A., Sodré, L., Schoenell, W., Gomes, J.M. 2007, MNRAS, 381, 263
3. Baldry, I.K., Glazebrook, K. 2003, ApJ, 593, 258
4. Bernard, J.P., Boulanger, F., Desert, F.X., Giard, M., Helou, G., Puget, J.L. 1994, A&A, 291, L5
5. Cid Fernandes, R., Asari, N.V., Sodré, L., Stasińska, G., Mateus, A., Torres-Papaqui, J.P., Schoenell, W. 2007, MNRAS, 375, L16
6. da Cunha, E., Charlot, S., Elbaz, D. 2008, MNRAS, 388, 1595
7. Damjanov, I., et al. 2010, PASP, submitted
8. Davies, R., et al. 2007, Messenger, 131, 7

9. de Ruyter, S., van Winckel, H., Maas, T., Lloyd Evans, T., Waters, L.B.F.M., Dejonghe, H. 2006, A&A, 448, 641
10. Désert, F.-X., Boulanger, F., Puget, J.L. 1990, A&A, 237, 215
11. Dullemond, C.P., Dominik, C., Natta, A. 2001, ApJ, 560, 957
12. Flagey, N., Boulanger, F., Verstraete, L., Miville Deschênes, M.A., Noriega Crespo, A., Reach, W.T. 2006, A&A, 453, 969
13. Flicker, R., Rigaut, F. 2001, In: Beyond Conventional Adaptive Optics, ESO Conf Workshop Proc., 58, 377
14. Haisch, K.E., Jr., Lada, E.A., Lada, C.J. 2000, AJ, 120, 1396
15. Jonsson, P. 2006, MNRAS, 372, 2
16. Longmore, S.N., Maercker, M., Ramstedt, S., Burton, M.G. 2007, MNRAS, 380, 1497
17. Maercker, M., Burton, M.G. 2005, A&A, 438, 663
18. Maercker, M., Burton, M.G., Wright, C.M. 2006, A&A, 450, 253
19. Mannucci F. 2007, Astronomy with Laser Guide Star Adaptive Optics, http://www.mpiahd.mpg.de/PARSEC/Ring2007/TalksPostersPDF/Tuesday/Highz FilippoMannucci.pdf
20. Mateus, A., Sodré, L., Cid Fernandes, R., Stasińska, G., Schoenell, W., Gomes, J.M. 2006, MNRAS, 370, 721
21. Mentuch, E., et al. 2009, ApJ, 706, 1020
22. Phillips, J.P., Ramos-Larios, G. 2005, MNRAS, 364, 849
23. Riffel, R., Pastoriza, M.G., Rodríguez-Ardila, A., Bonatto, C. 2009, MNRAS, 400, 273
24. Sellgren, K. 1984, ApJ, 277, 623
25. Sellgren, K., Werner, M.W., Dinerstein, H.L. 1983, ApJ, 271, L13
26. Sellgren, K., Werner, M.W., Allamandola, L.J. 1996, ApJS, 102, 369
27. Stern, D., et al. 2005, ApJ, 631, 163
28. van Breugel, W., Bland-Hawthorn, J. 2000, Proceedings from ASP Conference, Imaging the Universe in Three Dimensions, ASP, San Francisco, CA, vol. 195

Spitzer's View of Galaxies in the High-Redshift Universe

Giovanni G. Fazio

Abstract One of the most important observations made by the Spitzer Space Telescope has been the detection of luminous galaxies back to the era of reionization ($z \sim 8$), when the universe was less than 700 million years old. The key advance made by Spitzer imaging is the ability, for the first time, to sample the redshifted rest-frame visible light of these galaxies. When combined with broadband multi-wavelength data, Spitzer observations can be fit to stellar population synthesis models to determine the spectral energy distribution of these galaxies and to constrain their stellar masses and ages and their star formation histories. As a result, there is evidence that most of the stellar mass of these galaxies formed at even higher redshifts ($z > 10$ to 12) and that a significant number of galaxies should exist in this region. Searches for galaxies at $z \sim 9$ to 10 continue. Spitzer observations of massive lensing clusters have also played a pivotal role in this study. The first IRAC detection of a $z > 6$ galaxy came from such observations. Since most of these results were obtained with Spitzer/IRAC 3.6/4.5 μm bands, the Spitzer Warm Mission, when combined with future HST/WFC3 observations, will provide a unique opportunity to obtain the first complete census of the assembly of stellar mass as a function of cosmic time back to the era of reionization, yielding unique information on galaxy formation in the early universe.

1 Introduction

An outstanding challenge to modern astrophysics is to understand the earliest stages of how galaxies formed and evolved. It is important to observe the high-redshift universe to (1) discover and study the properties of the first galaxies and to understand their role in re-ionization of the early universe; (2) determine what physical processes govern star formation in high-z galaxies and to understand how stellar mass is assembled; (3) determine the presence and nature of the first stars (Population III; metal free) and how they constrained galaxy evolution.

G.G. Fazio (✉)
Harvard Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
e-mail: gfazio@cfa.harvard.edu

D.L. Block et al. (eds.), *Galaxies and their Masks*,

DOI 10.1007/978-1-4419-7317-7_37,

Current ideas suggest that the period of the Dark Ages is the epoch during which gas began to collapse and form the first stars, which then began to ionize the universe. This process of ionization started after a few 100 Myr ($z \sim 10$–12) of the Big Bang and was completed by ~ 700 Myr ($z \sim 6$). The challenges of observing this process of galaxy formation are particularly difficult. These first galaxies are small in size, extremely faint and must be observed primarily at infrared wavelengths due to the cosmological redshift. However, recent observations obtained with the Spitzer Space Telescope, when combined with multi-wavelength data from the Hubble Space Telescope and large ground-based telescopes, have achieved major advances in our understanding of how galaxies assembled in the early universe.

2 The Spitzer Space Telescope

The Spitzer Space Telescope, launched on 25 August 2003, is producing a significantly new view of the universe at infrared wavelengths (Fig. 1). Spitzer is the fourth and final element in NASA's Great Observatory series [24]. It consists of an 85-cm telescope and three cryogenically cooled instruments capable of imaging from 3 to 180 μm wavelength and spectroscopy from 5 to 40 μm wavelength. Combining the intrinsic sensitivity achievable with a cryogenic telescope in space with the high sensitivity of modern, large-area infrared detector arrays, Spitzer is providing the astronomical community with huge gains in capability for exploring the infrared universe. Primary among Spitzer's scientific objectives is the study of the formation and evolution of galaxies in the early universe, understanding energy sources in ultraluminous galaxies, the study of star formation and evolution, observations of exoplanets and their atmospheres, determining the structure and evolution of planetary disks around nearby stars, and exploring the nature and distribution of brown dwarfs.

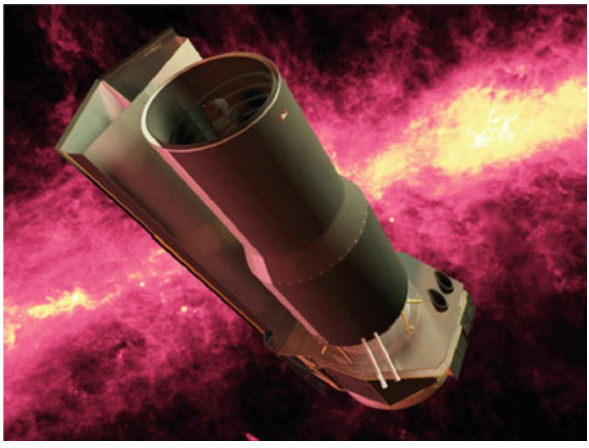


Fig. 1 The Spitzer Space Telescope NASA/JPL

Of particular interest in the study of the z >5 universe is one of the three focal-plane instruments, the Infrared Array Camera (IRAC; [10]). IRAC is an imaging camera operating at four broadband wavelength bands (called channels) centered at 3.6, 4.5, 5.8, and 8 μm. The short wavelength detectors are InSb devices, while the two longer wavelength detectors are Si:As devices. The instrument observes two fields of view (FOVs) of 5.12×5.12 arcmin simultaneously, with their centers about 6.6 arcmin apart, leaving a gap of about 1.5 arcmin between the FOVs. Each FOV is seen simultaneously by two detectors (one InSb and one Si:As detector), fed through lenses and filters and through a beamsplitter. Each array has 256×256 pixels, with the same plate scale (1.2 arcsec per pixel). The IRAC low-background sensitivity at 3.6 and 4.5 μm is 0.52 and 1.0 μJy (5 σ, 1 h integration), corresponding to 24.6 and 23.8 mag AB, respectively.

The superfluid helium, which was used to cool Spitzers mirror to 5.5 K and instruments to 1.4 K, was exhausted in May 2009, ending the cryogenic phase of Spitzer's operation. However, during the Spitzer Warm Mission, the instrument base plate remained at 27 K by radiation cooling, permitting IRAC bands at 3.6 and 4.5 μm to continue to operate with the same sensitivity as in the cryogenic phase. The Spitzer Warm Mission can continue to operate until 2014.

3 IRAC Observations of the Early Universe

One of the original scientific objectives of IRAC was to understand the formation and evolution of normal galaxies to redshifts z >3 by means of deep surveys in the four IRAC bands. At the time (early-1990s) this limit was selected because it was apparently beyond the peak in the known space density of luminous quasars. Specifically, the IRAC sensitivity requirement was set by the need to determine the spectral energy distribution (SED) of a L^* galaxy at z = 3. The four IRAC filters were originally designed to determine the photometric redshift of galaxies, based on the 1.6 μm peak (due to the H- opacity minimum in the atmospheres of late type stars) and the 2.3 μm CO absorption feature [19, 25]. As shown in Fig. 2, the 1.6 and 2.3 μm features appear early in galaxy spectra and are persistent with galaxy age. It is also important to note from Fig. 2 that, for the first time at high redshift, Spitzer/IRAC offered access to the rest-frame visible light of galaxies, longward of the Balmer break.

In summary, Spitzer/IRAC provides a unique perspective for observations of the early universe: (1) the IRAC filters (3–9 μm) probe, for the first time, the rest-frame optical and near-IR light at high redshifts; (2) these observations, when compared to rest-frame ultraviolet observations, are less affected by dust extinction; (3) IRAC observations provide the first view of light from longer lived stars that dominate the stellar mass of high-redshift galaxies; (4) from these observations the unique properties of a galaxy's stellar mass and age can be determined; and (4) they enable a new perspective on the study of early galaxy formation and evolution. In actual practice, using IRAC color–color diagrams based on the four IRAC bands only (e.g.,

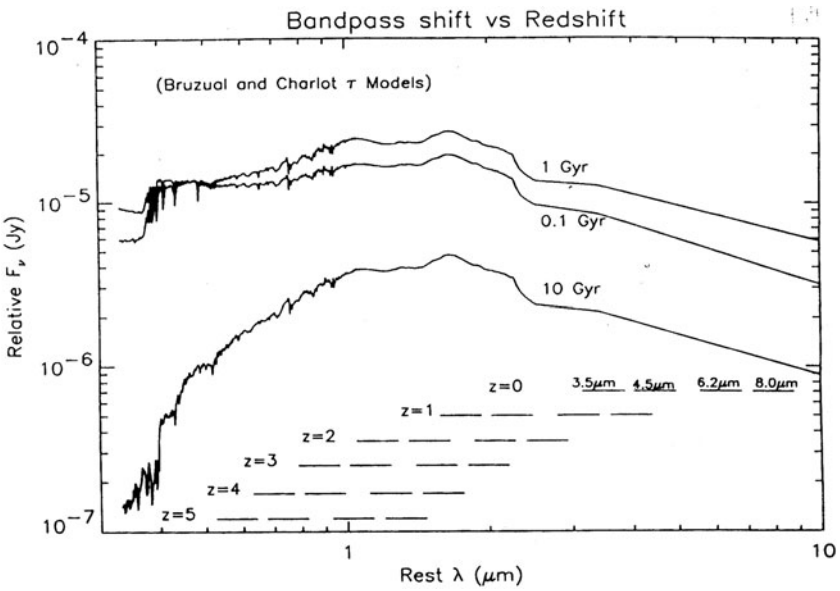


Fig. 2 Predicted spectra of galaxies aged 0.1, 1, and 10 Gyr old [3]. Note the early appearance and persistence of the features at 1.6 and 2.3 μm

[3.6]−[4.5] vs. [5.8]−[8.0]), to identify high-z galaxies is practical only to z ∼ 2–3. Beyond this redshift range the sensitivity of the 4.5 and 8.0 μm bands are insufficient to detect galaxies. Therefore, other techniques, based on optical and near-IR wavelengths, have to be used to select high-redshift galaxies.

4 Identifying High-Redshift Galaxies

One method used to identify high-redshift galaxies in a multi-wavelength survey is the Lyman Break Galaxy (LBG) technique, which was first developed by Steidel et al. [21, 22]. This method uses the Lyman break at 912 Å and the dimming between 912 and 1,216 Å due to the Lyman-alpha forest in a galaxy spectrum to identify its redshift. Since the observed wavelength is (1+z) times the rest-frame wavelength, absence of a signal in the wavelength band shortward of the Lyman break can be used to measure z. Thus a B-band dropout indicates z = 4; V-band z = 5; i-band z = 6; z-band z = 7; Y-band z = 8, and J-band z = 10, etc. The second technique for identifying high-redshift galaxies is use the Lyman-alpha Emitter (LAE) technique. This method uses a narrow band filter (100–200 Å wide) to detect the Lyman-alpha emission line (1,216 Å) from atomic hydrogen in a galaxy spectrum. The filter is positioned in a wavelength band between the night sky emission lines to reduce the background emission. A sky survey is then carried out in both the narrow-band filter and a wide-band filter that includes the narrow-band wavelength. Comparison of the

images taken with the two filters permits identification of the Lyman-alpha emission galaxy. For example, a narrow-band filter at 9,730 Å can identify a Lyman-alpha emission galaxy at z ∼ 7 [18].

5 IRAC Observations of Galaxies at z = 5 to 7 (1.2–0.78 Gyr after Big Bang)

Using the Lyman Break Technique, Eyles et al. [9] presented evidence for one of the first HST and Spitzer/IRAC detections of galaxies at z ∼ 6. Using the HST Great Observatories Origins Deep Survey (GOODS) in B-, V-, I-, and z-bands and the IRAC bands and applying the I-band dropout technique, they identified four z ∼ 6 galaxies, two of which had IRAC detections in the 3.6 and 4.5 μm bands (Fig. 3). The redshift was later spectroscopically confirmed using the Keck telescope. From the model fit it was determined that the galaxy SBM03 #1 was massive (2.3×10^{10} solar masses). The Spitzer/IRAC photometry revealed a significant Balmer discontinuity which indicated that the galaxy contained a mature stellar population with ages ∼ 400 Myr old, which implied formation of redshift of z >7. Without the IRAC mid-infrared photometry to very faint magnitudes, the stellar ages and masses of z >5 galaxies are poorly constrained. These observations indicated that galaxies with stellar masses >20% of those of a present-day L* galaxy had already

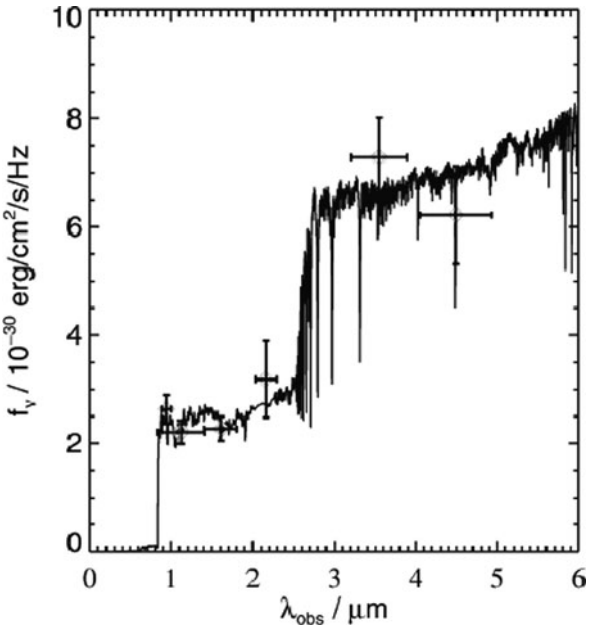


Fig. 3 Observed spectrum of the z = 5.83 galaxy SBM03#1 fit with a Bruzual and Charlot [4] model [9]

assembled within the first Gyr after the Big Bang. The average star formation rate deduced was 5–30 solar masses per year. Similar results were reported by Yan et al. [26] and Dow-Hygelund et al. [7].

The evolution of LBGs between z = 4 to 6 by Stark et al. [20], using HST/ACS and Spitzer/IRAC/MIPS observations in GOODS, showed that no strong evolution of stellar masses and stellar ages of galaxies of fixed UV luminosity existed between z = 6 to 4 and that this was consistent with the drop in the UV luminosity function between z = 4 to 6. They also presented evidence that the stellar mass function of UV luminous galaxies grows significantly from z = 6 to 4.

6 Observations of Galaxies at z >7 (<800 Myr After the Big Bang)

Labbé et al. [15] extended the galaxy mass estimates to z = 7 to 8 using very deep (25–27 mag AB) IRAC observations of six z-band dropout candidates found in GOODS observations of the Hubble Ultra-Deep Field (HUDF). Two of the candidates were clearly detected by IRAC. This redshift region is of particular importance since it is in the reionization era of the Big Bang model. Again, the IRAC observations allow, for the first time, the ability to constrain the rest-frame optical colors, stellar masses, and ages of the highest redshift galaxies. Labbé et al. [15], fitting stellar population models to the spectral energy distributions, determined photometric redshifts in the range 6.7–7.4, stellar masses (1–10) $\times 10^9$ solar masses, stellar ages 50–200 Myr, star formation rates up to $\sim$ 25 solar masses per year, and low reddening. The z = 7 galaxies appear to be much less massive and evolved than the galaxies at z = 2 to 3, but similar to the masses measured at z = 5 to 6. Again the indication is that these z = 7 galaxies formed at z >8, during the era of comic reionization. However, the star formation rate density derived for these objects from their stellar masses and ages is more than three times too small to reionize the universe. These results imply that low-mass galaxies beyond the current detection limit were primarily responsible for reionization. Gonzalez et al. [11] reported similar results from a more robust sample of 11 z = 7 galaxies (z-band dropouts) detected by IRAC. They also concluded that the specific star formation rate at a given galaxy mass value is a constant as a function of redshift from z = 7 to z = 2 but drops suddenly at z = 2 and that the stellar mass density decreases as a function of redshift (Fig. 4).

IRAC observations of sub-L* galaxies in the region between z $\sim$ 7 and 8 by Labbé et al. [17] yielded important insights into the earliest phases of galaxy evolution. These results yielded relatively high stellar ages ($\sim$ 300 Myr) for these galaxies and demonstrated that the specific star formation rate does not depend on stellar mass. They also observed a significant contribution of low-luminosity galaxies (0.06 L* at z = 3) to the stellar mass density at z $\sim$ 7 and concluded that the stellar mass seen in these galaxies could provide a substantial contribution to the reionization of the universe. The evolution of galaxies in the redshift range z = 7 to 8 based on recent HST/WFC3 and Spitzer/IRAC observations (see [16]) in the Hubble Ultra

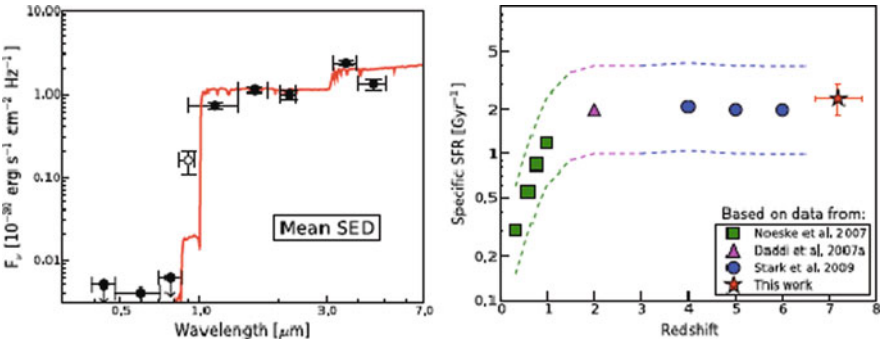


Fig. 4 *Left*: Average SED derived from the 11 z = 7 sources detected by Gonzalez et al. [11]. *Right*: The specific star formation rate derived from the z = 7 galaxies combined with previous results at a constant galaxy mass of 5×10^9 solar masses

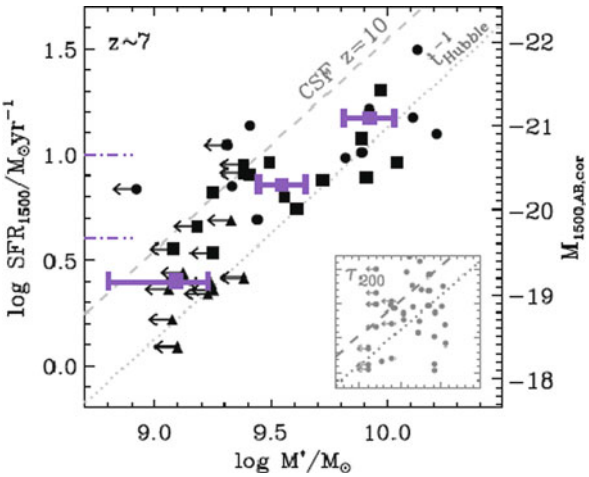


Fig. 5 UV-derived, dust-corrected star formation rate vs. stellar mass for z ∼ 7 galaxies. Data from NICMOS (*circles*), WFC3/IR ERS (*squares*), and WFC3/IR HUDF sample (*triangles*). The *large bars* show the average stellar mass in bins of star formation rate centered on log SFR = 0.4, 0.8, and 1.2 [16]

Deep Field (HUDF09) yielded 21 z ∼ 7 (z-band dropouts) and 3 z ∼ 8 (Y-band dropouts). Galaxies with z ∼ 7 have high ages (>300 Myr) and their derived stellar masses correlate well with the star formation rate (Fig. 5).

The first robust Spitzer/IRAC detection of a Y-band dropout galaxy at z ∼ 8 (650 Myr after the Big Bang) was also reported by Labbé et al. [16]. The magnitude of the Balmer break was indicative of an evolved stellar population. This galaxy had a similar age (∼ 300 Myr) to the z ∼ 7 galaxies, suggesting modest to no evolution between z = 7 and z = 8. These results also implied that initial star formation occurred at z ∼ 12. Labbé et al. [17] also showed that the stellar mass density

decreased as $(1 + z)^{-6}$ between z = 4 to 8 (Fig. 6). The search for z = 10 galaxies (500 Myr after the Big Bang) by Bouwens et al. [1], using observations with the HST/WFC3/IR in the HUDF, found three J-band dropouts, visible in H-band only. Stacking of the IRAC data yielded no detection.

7 The Search for Lensed Galaxies at z ∼ 7

The first high-z galaxy detected by IRAC was the result of a search for lensed high redshift galaxies in massive galaxy clusters [8]. A z ∼ 7 galaxy was detected lensed by the cluster Abell 2218. The gravitationally lensed objects, with a magnification of 25, were discovered by Kneib et al. [14] in HST images. By fitting a Bruzual–Charlot model to the HST and Spitzer/IRAC data the photometric redshift measured was z = 6.6 to 6.8. A significant Balmer break gave a stellar age of 100–400 Myr, which implied the stars formed at 9 <z <12. The star formation rate was ∼ 2.5 solar masses per year and the stellar mass of the galaxy was ∼ 10^9 solar masses. Bradley et al. [2], using HST, found a lensed galaxy at z ∼ 7.6 in the massive galaxy cluster Abell 1689, which was also detected by IRAC. They observed that the magnification was ∼ 9.3 and that star formation was occurring in compact knots of ∼ 300 pc.

8 A Search for the First Stars (∼ 200 Myr after the Big Bang)

Kashlinsky et al. [13] reported measurements, using IRAC, of diffuse flux fluctuations in the cosmic infrared background radiation. These fluctuations were detected in deep IRAC surveys after removing foreground stars and galaxies and noting that the measured anisotropies exceeded the instrument noise and local foreground radiation fluctuations. The amplitude of the fluctuations was ∼ 0.1–0.3 nW/m^{2}sr at 3.6–8 μm wavelength on scales of ∼ 1 arcmin. Kashlinsky [12] and Kashlinsky et al. [13] interpreted these results as evidence for a first light component (Population III stars) at z >8. However, the redshift of these fluctuations was unknown. Cooray et al. [6] and Chary et al. [5], using HST/ACS observations, concluded the effect could be due to undetected dwarf galaxies at z ∼ 1–3. Thompson et al. [23] using HST/NICMOS near-IR data reported only upper limits to any fluctuations and concluded their results were inconsistent with a Population III star interpretation. Future, more extensive observations are planned with the Spitzer Warm Mission to try to resolve these conflicts.

9 Future Observations to Search for High-Redshift Galaxies

The searches thus far for high redshift galaxies have been limited to rather small areas (<0.1 deg^2) or with a small number of massive galaxy clusters. Two major Spitzer Warm Mission Exploration Programs and a HST Cosmology Survey

Multi-Cycle Treasury Program will soon provide much larger survey areas and more sensitive observations to search for high z galaxies as well as a larger number of massive lensed clusters. However, the ultimate study of the early universe will be carried out by the James Webb Space Telescope (JWST), which will launched in 2014.

The Spitzer Extended Deep Survey (SEDS; G. Fazio, PI) will provide a unique opportunity to obtain the first complete census of the assembly of stellar mass and black holes as a function of cosmic time back to the era of reionization, yielding unique information on galaxy formation in the early universe. The survey will also measure galaxy clustering over a wide redshift range, which will provide the critical link between galaxies and their dark matter halos and critical tests of models of early star formation. SEDS will achieve these goals by tracing the stellar mass growth in mass selected samples of galaxies via their broadband spectral energy distributions. The baseline proposal is an unbiased survey with 12 h/pointing at 3.6 and 4.5 μm over five well-studied fields of 0.90 deg^2 total. The survey expects to find (a) >10,000 galaxies at z = 4 to 6 (including ~1,000 galaxies at z = 6), reaching galaxies down to ~ 5×10^9 solar masses at z = 6, necessary to robustly measure M* at that redshift, i.e., the galaxies that dominate the global stellar mass density, and (b) >100 massive galaxies at z = 7, which will firmly anchor the high mass end of the early galaxy populations and provide targets bright enough for future spectroscopic follow up with 20–30 m telescopes, JWST, and ALMA. The proposed five field deep survey will enable several secondary science objectives. These include (1) galaxy evolution in the redshift range z ~ 1–4, (2) AGN variability, and (3) measurement of the cosmic infrared background spatial fluctuations.

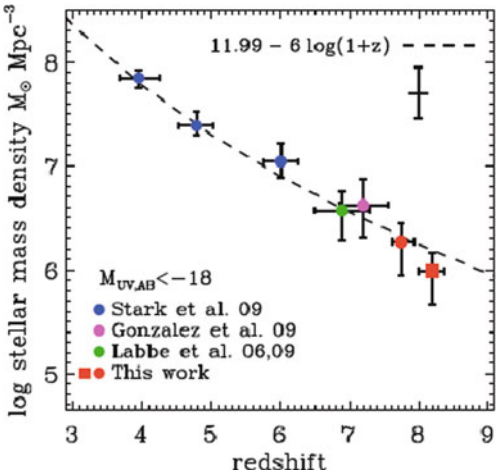


Fig. 6 Evolution of the integrated stellar mass density from z = 4 to 8 [17]. The data at z = 4–6 are from Stark et al. [20]; the data at z = 7 are from Gonzalez et al. [11] and Labbé et al. [15, 16]; and the data at z >7 from Labbé et al. ([16], labeled "This work" in the figure). The *floating error bar* indicates the expected cosmic variance for the z ~ 8 samples

Massive clusters of galaxies are now recognized as very effective "cosmic telescopes". Because of the gravitational lensing effect, they can amplify significantly the background sources (by factors of a few tens) thereby bringing into view faint sources that would otherwise be unobservable. Note that in the background-limited case, which is applicable to IRAC observations, a factor of 20–30 gravitational amplification translates into increasing the integration time by a factor of 400–900. Because of this tremendous gain in sensitivity, IRAC imaging of lensing clusters will permit JWST depth ($\sim$10 nJy) to be achieved with Spitzer. Despite this great possibility, however, the full potential of the lensing cluster technique has not yet been realized due to the small number of clusters that have well constrained accurate mass models. During the Spitzer Warm Mission Exploration Program, E. Egami (PI) will conduct an IRAC imaging survey of 47 massive lensing clusters (5 h/band, 2 bands) for which accurate mass models have been developed through many years of intensive imaging/spectroscopic campaigns with HST, Keck, and VLT telescopes. This is the first time when such a large, statistical sample of such well-characterized clusters will be systematically employed to probe the high redshift universe. This IRAC survey is a key component of a more comprehensive program, which includes HST/WFC3 and Herschel observations in 2010. Scientifically, IRAC data will be used to (1) characterize z >6 galaxies (expecting $\sim$ 50 z $\sim$ 7 $\sim$ 8 galaxy detections), (2) support future Herschel and ALMA surveys, and (3) search for z >6 supernovae.

The HST Cosmology Multi-Cycle Treasury Proposal survey (S. Faber, PI; H. Ferguson, Co-PI) will document the first third of galactic evolution from z = 8 to 1.5 and test for evolution in the properties of Type Ia supernovae to z $\sim$ 2 by imaging more than 2,50,000 galaxies with HST WFC3/IR and ACS. Five premier multi-wavelength regions were selected from within the Spitzer SEDS survey, providing complementary IRAC data down to 26.5 AB mag, and a unique resource for stellar masses at high redshifts. The use of five widely separated fields mitigates cosmic variance and yields statistically robust samples of galaxies down to 10^9 $M_\odot$ out to z $\sim$ 8. The program incorporates a two-tiered strategy using a "Wide" component (2 orbits deep over $\sim$0.2 deg^2) and a "Deep" component (12 orbits deep over $\sim$0.04 deg^2). Combining these with ultra-deep imaging from the Cycle 17 HUDF09 program yields a three-tiered strategy for efficient sampling of both rare/bright and faint/common objects.

One of the key science goals of the JWST is to study the first light sources and the era of reionization of the universe. It will identify the first luminous sources to form and will determine the ionization history of the early universe. JWST is a large (6.6-m) cold (<50 K), infrared-optimized space telescope, which will orbit around the second Earth–Sun Lagrange point. The observatory will have four instruments: a near-infrared camera (NIRCam), a near-infrared multi-object spectrograph (NIRSpec), and a tunable filter imager (TFI) which will cover the wavelength range from 0.6 to 5.0 μm, and a mid-infrared camera/spectrometer (MIRI) which will cover the wavelength range from 5.0 to 29 μm.

10 Summary

Detecting the most distance galaxies known in the universe, back to the era of reionization ($z \sim 8$) has been one of most remarkable achievements of Spitzer Space Telescope (85-cm mirror). Dramatic new results on this topic have been achieved over the last year.

When combined with deep, broadband multi-wavelength data, Spitzer observations can be fit to stellar population synthesis models to determine, for the first time, the spectral energy distribution of these galaxies and to constrain their stellar masses and ages and their star formation histories. These results have enabled a new perspective on studies of early galaxy formation: (1) massive galaxies ($\sim 10^{10}$ $M_{\odot}$) with stellar ages ~200–300 Myr existed in early universe ($z \sim 6$ to 8); (2) the stars in these galaxies formed several hundred years earlier ($z \sim 8$ to 12); (3) comprehensive measures of the stellar mass density exist over $4 < z < 8$; (4) Stellar mass correlates well with SFR; (5) evidence exists for constant SFR from $z \sim 10$ to 12; (6) there exists a substantial contribution of low-luminosity galaxies at $z=7$ that could provide a significant fraction of the energy to ionize the universe.

The Spitzer Warm Mission/HST will provide a unique opportunity to obtain the first complete census of the assembly of stellar mass as a function of cosmic time back into the era of reionization, yielding unique information on galaxy formation in the early Universe.

References

1. Bouwens, R.J., et al. 2009, arXiv:0912.4263
2. Bradley, L.D., et al. 2008, ApJ, 678, 647
3. Bruzual, G., Charlot, S. 1993, ApJ, 405, 538
4. Bruzual, G., Charlot, S. 2003, MNRAS 344, 1000
5. Chary, R., et al. 2008, ApJ, 681, 53
6. Cooray, A., et al. 2007, ApJ, 659, L91
7. Dow-Hygelund, C.C., et al. 2005, ApJ, 630, L137
8. Egami, E., et al. 2004, ApJ, 618, L5
9. Eyles, L.P., et al. 2005, MNRAS 364, 443
10. Fazio, G.G., et al. 2004, ApJS, 154, 10
11. Gonzalez, V., et al. 2010, ApJ, 713, 115
12. Kashlinsky, A., et al. 2005, Nature, 438, 45
13. Kneib, J.P., et al. 2004, ApJ, 607, 697
14. Kashlinsky, A. 2005, PhR, 409, 361
15. Labbé, I., et al. 2006, ApJ, 649, L67
16. Labbé, I., et al. 2009, arXiv:0911.1356
17. Labbé, I., et al. 2010, ApJ, 708, L26
18. Ota, K., et al. 2008, ApJ, 677, 12
19. Simpson, C., Eisenhardt, P. 1999, PASP 97, 451
20. Stark, D.P., et al. 2009, ApJ, 697, 1493
21. Steidel, C., et al. 1996, ApJ, 462, L17

22. Steidel, C., et al. 1999, ApJ, 519, 1
23. Thompson, R., et al. 2007, ApJ, 666, 658
24. Werner, M., et al. 2004, ApJS 154, 1
25. Wright, E. L., Eisenhardt, P., Fazio, G. 1993, BAAS 268, 93
26. Yan, H., et al. 2005, ApJ 634, 109

Bandshifting and Other Masks of the Clumpy Populations in High-Redshift Galaxies

Bruce G. Elmegreen and Debra Meloy Elmegreen

Abstract Galaxies observed at high redshift give a biased view of their structures, clump masses, clump ages, and surface densities as a result of bandshifting, angular resolution limitations, and surface brightness dimming. These biases produce trends with redshift, but they are not likely to affect our impression that young galaxies are intrinsically more clumpy than modern galaxies. The clumps probably originate as gravitational instabilities in gas-rich, turbulent disks. The origin of the high gas fractions is not known: they could result from mergers or from rapid cosmological infall.

1 Introduction

High redshifts show galaxies in their formative state, but there are many selection effects that can lead to difficulties in interpretation. Considering the theme of this conference, we refer to these effects as "cosmological masks." Three types of masks will be considered here:

- Bandshifting of the restframe FUV for a high-z galaxy into the optical passband of a telescope. This effect does two things. It shows star formation regions much more prominently than they appear in visible images of local galaxies because the FUV from these regions is relatively more luminous compared to the visible background from older stars. Second, it emphasizes younger regions in SED fits, making selected regions seem younger than the average and mixed populations appear younger than their true age span. Both of these effects occur for local galaxies too, since younger regions are bluer in both cases, but they can place a bias on observable high-redshift morphologies and they can make observable star formation rates higher than the average.
- Image magnification and surface brightness dimming. These two effects go together. Because of the curvature of the universe, a kiloparsec length stops

B.G. Elmegreen (✉)
IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA
e-mail: bge@us.ibm.com

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_38,

decreasing in angular size at about $z = 1$ and stays about constant until $z \sim 5$. A galaxy 10 kpc in size subtends about 1 arcsec for $z = 1$–5, making the study of galactic morphology possible with the Hubble Space Telescope (HST), which has a FWHM resolution of $\sim$ 0.09 arcsec in optical bands. Compensating for this enhanced viewing capability is a strong decrease in surface brightness, as $1/(1+z)^4$ (see [7]).

- The age of the universe gets younger, so the underlying disk and bulge get younger and more irregular with increasing redshift. This means that bulges begin to look like giant star-forming regions at $z \sim 3$ and they may even be slightly off-center.

Several studies have modified local galaxy images to make them appear as they might at high redshift. Barden et al. [3] shifted SDSS galaxies to $z = 0.15$, 0.5, and 1. They found a significant loss of faint structure, especially at the periphery where tidal features might be. Overzier et al. [26–28] studied local galaxies they called "Lyman Break Analogs," which are "supercompact UV-luminous galaxies" [24] observed with GALEX, having $L_{\rm FUV} > 10^{10.3}\ L_\odot$ and $I_{\rm FUV} > 10^9\ L_\odot$ kpc^{-2} at $z < 0.3$. These objects are rare with a space density of $\sim 10^{-6}$ Mpc^{-3}. Overzier et al. [28] redshifted LBAs to $z = 2$, 3, and 4, and found that small clumps blend together and faint peripheral tidal features disappear with increasing z. They also found that asymmetry decreases at higher z, which means that faint peripheral asymmetric features can get lost. Their conclusion was that high-z clumpy galaxies could be mergers even though there is no direct evidence for tidal features yet. Petty et al. [29] studied Gini coefficients (a measure of smoothness), M_{20} (a measure of central concentration) and the Sérsic index n (a measure of radial profile shape related to central concentration) for a sample of local galaxies redshifted to $z = 1.5$ and $z = 4$. They found that all the redshifted galaxies got smoother (smaller Gini) with increasing redshift and some got more centrally concentrated (more negative M_{20}) with increasing redshift. All of these studies indicate that bandshifting, angular resolution degradation out to $z \sim 1$, and surface brightness dimming combine to hide the faint, small, red structures in galaxies at high z.

Figure 1 shows the redshift progression of four important quantities: the angular size of 10 kpc, the observed surface brightness for a fixed intrinsic surface brightness, the restframe wavelength for the ACS i_{775} filter, and the age of the universe. The ΛCDM parameters in Spergel et al. [30] are assumed. At $z < 0.5$, objects of a fixed physical size get smaller in angle with increasing distance, but starting at $z \sim 1$, they increase in angle because of gravitational magnification in the universe. The surface brightness in a fixed bandpass decreases rapidly, as $1/(1+z)^4$. Distant galaxies with the size and luminosity of the Milky Way can be large in angle but extremely faint in surface brightness – only those with the brightest star formation will be observable as extended objects.

The restframe wavelength of the HST ACS filter i_{775} (7,750 Å) decreases to the wavelength of the Palomar Observatory IIIa-J plates (Blue) by $z \sim 1$, which means that within $z \sim 1$ we are still viewing galaxies in their restframe optical bands. This is convenient because several major HST surveys (GOODS, GEMS, COSMOS)

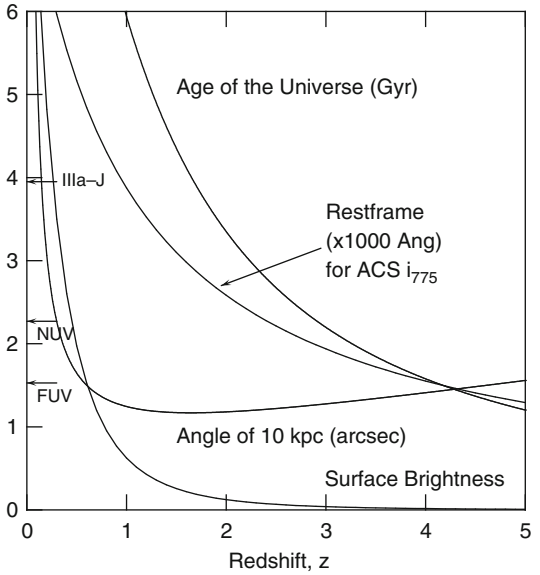


Fig. 1 Variation of surface brightness, angular size of a 10 kpc region, restframe wavelength for the ACS i_{775} band, and age of the universe, all versus redshift, from the standard ΛCDM cosmology. The scale for each uses the ordinate numbers with different units, except for surface brightness, which is simply $10/(1+z)^4$

cover a redshift range out to $z \sim 1$ or 1.5. Then observed galaxy morphologies can be compared with Hubble classifications based on optical images. Beyond $z \sim 2$, the i_{775} band begins to show the restframe ultraviolet. At $z = 2.41$, i_{775} shows the GALEX NUV band, which has a restframe 2,271 Å, and at $z = 4.07$, i_{775} shows the GALEX FUV band at 1,528 Å.

Putting these effects together, we see that at $z \sim 3$, galaxies 10 kpc in size subtend an angle of 1.28 arcsec, so they are resolved by HST. Their surface brightnesses are dimmer than in local galaxies by a factor $(1+z)^{-4} = 0.004$, so only the most star-bursting can be seen. In the i_{775} band, they are showing us their restframe NUV distribution and the entire age of the universe for them is only 2.2 Gyr, which means that the oldest A-type stars are evolving off the main sequence. These same stars would be 11.4 Gyr old today, and mostly located in the halo, thick disk, old globular clusters, and bulge. Population I stars in modern galaxy disks have not formed yet.

2 Clumpy Galaxies

We can get an idea of what modern galaxies would look like at $z \sim 3$ from the GALEX archives, using the NUV and FUV images obtained with this satellite. For example, a collection of GALEX images is at
`http://www.galex.caltech.edu/media/glx2006-02r_img01.html`.

Most of the main features of today's galaxies can be seen there: exponential disks, spiral arms and bars, lots of small star-forming regions, and a diverse proportion of disks and bulges. In fact real galaxies at $z \sim 2$–4 rarely look like this. Figure 2 shows a sample of types from the Hubble Ultra Deep Field (UDF) imaged in i_{775} band. These types come from a study of all $\sim 1,000$ galaxies in the UDF that are larger than 10 pixels in diameter. The numbers of each type are indicated on the left. The bottom two rows are normal types; spirals dominate with 313. Chain galaxies [11] and clump clusters [17] together contain the same number, 313, while the double galaxies and tadpole shapes contribute another 248 to the clumpy morphologies. The most peculiar are the chains and clump clusters, which typically have neither central concentration at i_{775} nor exponential disks. Their star-forming regions are also extremely large and bright compared to the underlying disks.

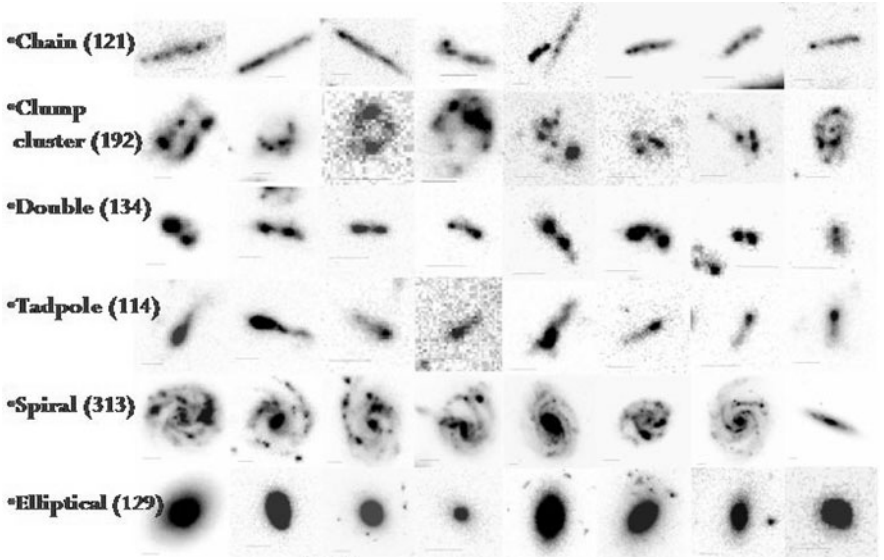


Fig. 2 Basic morphologies of galaxies in the UDF, names on the left with number indicating the numbers of each type. These are from a sample that includes all galaxies with diameters larger than 10 pixels. Eight examples of each type are shown (from [18])

Deep HST fields indicate that faint galaxies are predominantly clumpy (e.g., [1]). Photometric redshifts and SED fits to UDF galaxies suggest clumpy types dominate beyond $z \sim 1.5$ [10]. Spirals and ellipticals could also extend to high z, but the most inactive and intrinsically red of them would be difficult to see with the ACS camera. As a result, we do not know how prevalent normal-looking spirals are beyond $z \sim 2$. Observations in the restframe optical bands require IR cameras, and their resolution is relatively poor, making it difficult to see the star formation and spiral arm structure required for good classifications. Spiral types at high z are usually inferred only from their disk-like or extended structures. They could be different from local spirals in terms of clumpiness and symmetry.

We identified four types of disk galaxies in the GOODS survey [21]. Figure 3 shows two galaxies of each type. The length bar at the top left represents 2 kpc for the Wolf et al. [32] photometric redshift given in the lower left and the galaxy name from COMBO17 is given in the lower right. Many more examples are in Elmegreen et al. [21]. On the left are normal spirals with long stellar density waves. Second from the left are somewhat normal flocculent spirals, which means they have no prominent stellar density waves but spiral-like patches of star formation throughout. Second from the right are clumpy galaxies with relatively large clumps and an underlying red disk and on the right are clumpy galaxies without an underlying red disk. Note that the clumpy galaxies have mostly round clumps, not spiral-like clumps like we see in the flocculents. We also studied these basic disk types in the UDF [20], but mostly the results from GOODS will be discussed here.

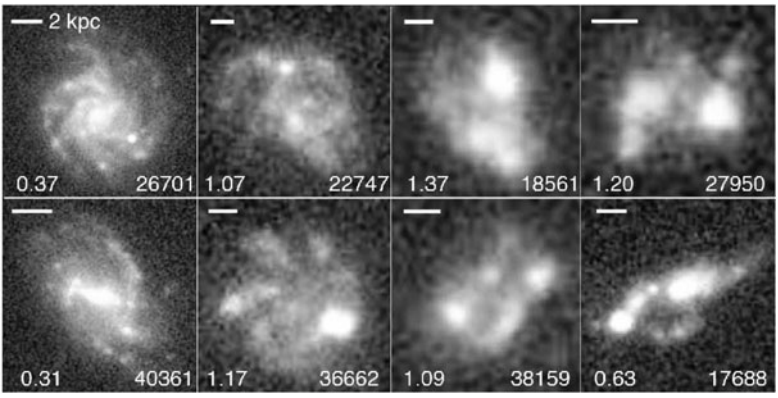


Fig. 3 Morphologies of disk galaxies in GOODS. The Wolf et al. [32] redshifts are in the *lower left* and the COMBO17 *numbers* are in the *lower right*. Galaxy types include normal density-wave spirals, flocculent spirals, clump clusters with *red underlying disks*, and clump clusters with *blue or no evident underlying disks*, in order from *left to right*. Two examples are shown for each type. Color images are in Elmegreen et al. [21]

The nature of the clumpy types can be ascertained from a variety of observations. The fundamental question we have been addressing is what is the nature and origin of the clumps themselves. We believe that the clumps are star-forming regions in a gas rich and turbulent disk, rather than individual galaxies, each with their own dark matter halos, in the process of merging. This does not mean that the whole galaxies cannot be merging or interacting with something else. There is in fact little or no evidence that whole clumpy galaxies are merging, but the usual evidence for this, such as peripheral tidal tails and bridges, could be too faint to see at high redshift, as discussed above.

An important result that has direct bearing on both the question of clump origin and the question of whole galaxy merging is the statistical distribution of chain

galaxies and so-called clump-cluster galaxies, which are the most clumpy types (right two pairs of panels in Fig. 3). On a distribution of chains and clump-clusters with respect to the ratio of minor to major axes, the chains occupy the region with low minor to major axis ratio and the clump clusters occupy the region with high minor to major axis ratio. This is obvious because of the way the galaxies are defined, but more important is that the two types blend together to give a nearly constant frequency in this distribution function, which is the case when a disk with fixed thickness is viewed in random orientations [14]. Also, the clumps in both types have the same range for colors and magnitudes and the whole galaxies have the same range for colors and magnitudes. Thus we conclude that chains are edge-on clump clusters. The importance of this conclusion is that the chains are fairly straight and the clumps in the chains deviate from the average midplane by only a small amount, a fraction of an ACS pixel on average or less than 100 pc [15]. The chains do not look warped as if they were interacting and the clump positions in the chains are not chaotic as if they came in from outside in a type of merger. The clumps look like star formation regions that formed in a disk where they are currently seen. For the face-on chains, i.e., the clump clusters, this means we are generally viewing a whole disk that has fragmented into ~ 5 giant clumps of star formation. In older versions of such fragmented disks, the clump stars have smoothed out to make the red component between the clumps (for cases like those second from the right in Fig. 3), and in younger versions of fragmented disks, most of the light is from what may be the first generation of major star formation.

A second important result is that the clumps are probably not simply the band-shifted artifacts of normal, restframe UV patches. There are two reasons for thinking this. Clumpy galaxies observed at $z > 1$ or so have clumpy structures, unlike local galaxies observed with GALEX: there are fewer clumps and each clump is more luminous in the high-redshift cases. Second, the GOODS survey has clumpy types at such low redshifts that we are actually viewing the restframe B-band with the ACS i_{775} or z_{850} filters. We know what B-band images of spiral and flocculent galaxies look like and they are different from clump clusters, i.e., clump clusters have no exponential disk light profile along radial strips, they often have no bulge or central concentration at all, even in the IR [20], they have fewer and larger clumps than normal spirals and the clumps do not have spiral-like shapes.

Third, clump clusters cover the same redshift range in the GOODS images as other galaxies that are clearly interacting [21]. There are M51-type interactions, one-arm tidal arm interactions, antenna–galaxy type interactions, ring–galaxy interactions, and so on, in the GOODS images [19]. There are also bent chains that look like interacting chains with tight clump confinement to a circular arc [16]. The clump clusters and straight chains are not like this. Evidently, tidal and other features indicative of interactions and mergers are present in GOODS images, even though surface brightness dimming would decrease their visibilities as discussed above. These features must therefore be brighter than local tidal features to be seen even with this dimming. Perhaps this greater brightness is a result of a greater gas fraction, which would promote higher star formation rates in tidal features.

3 Local Analogs and Galaxy Interactions

Several groups have searched for local analogs to clump cluster galaxies. Casini and Heidmann [8, 9] and Maehara et al. [25] discovered local "Clumpy Irregular Galaxies" of normal size. Examples are Markarian 296, Mrk 325, Mrk 7, Mrk 8 (from Casini and Heidmann [8, 9]), and Kiso UV excess galaxies 1618+378, 1624+404, 1626+413, and Mrk 297 from Maehara et al. [25]. Both the Kiso and the Markarian galaxies are unusually bright in the ultraviolet, so UV excess is evidently a good way to find these types, both locally and at high redshift. Garland et al. [23] studied "Luminous Compact Blue Galaxies," which are small, high-luminosity, high-surface brightness, and blue. The Garland et al. galaxies are also gas rich (in CO and HI), as are high-redshift clump clusters [12, 31]. The local analogs are rotating like some clump clusters too [22, 33], although both types also have distorted velocities. These distortions led Garland et al. [23] to suggest that both the local and the high redshift clumpy galaxies are interacting. Overzier et al. [26–28] studied "Lyman Break Analogs," as discussed above. Most Lyman Break Analogs are centrally concentrated, unlike many clumps clusters at high redshift (e.g., Fig. 2), but they all have highly clumpy star formation. Overzier et al. suggest that many or most of the local analogs are interacting, and by inference, that clump clusters may be interacting too.

The issue of interactions has two facets. First, it is possible that the clumpiest galaxies in the local universe are all interacting, but the clumpiest galaxies at high redshift are interacting only in some cases and doing something else in other cases. The main point is that highly clumpy structure seems to require violent instabilities as a star-formation trigger in the gas disk. Such instabilities need high gas fractions and fast turbulent motions as pre-conditions [5, 14]. Also, star formation rates are often so large that the large mass of gas that must be present has to accrete quickly, or else it will be consumed before it builds up. For an instability model, the star formation rate should be comparable to the accretion rate, as star formation is regulated by a threshold column density in the Toomre Q parameter. A typical star formation rate of $30\,M_\odot$ year^{-1} corresponds to a gas accretion rate of $10^{10}\,M_\odot$ per 300 Myr rotation time for a 10 kpc radius and a Milky Way mass. This is almost a whole disk mass of accretion in one or two rotations. Such a high gas accretion rate is possible for a merger at any epoch that has such large galaxies, but at high redshift there are also purely cosmological inflows of gas at about this same accretion rate [2, 13]. Note that cosmological inflow in this context does not mean zero metallicity, because the intergalactic gas could be contaminated by earlier star formation. Cold flow accretion differs from gas-rich mergers primarily in the clumpiness of the gravitational potential: cold flow accretion should be smoother than merger-driven accretion.

This diversity in accretion morphology leads to the second facet in our discussion of interactions relevant to clump clusters. Clumpy accretion can perturb an existing stellar disk to a large extent. Even minor mergers can put a high fraction of the existing disk stars into a spheroid, bulge, or halo. If the associated gas accretion were then to make a clump cluster phase, the clumps could form in a spheroid-rich environment, and this could make them spiral-like, rather than round [4]. Also,

merger-driven accretion tends to make thick disks with flares, while disk thickening by internal clump stirring makes thick disks with a radially independent scale height [6]. Thick disk components of spiral galaxies today look more like the latter [6, 34].

4 Selection Effects for Clump Properties

The trends shown in Fig. 1 have implications for the properties of clumps that are observed in clumpy galaxies. The clumps do not actually have to be resolved to be seen, and most likely many are not, but they do have to be separated from each other to be recognized. This requires a separation of $\sim$ 5 or more pixels in the ACS, considering the point spread function is about 3 pixels FWHM. Surface brightness dimming affects the clumps we see even more, because we would not notice extended clumps that are below the surface brightness threshold of a survey. We would also not notice whole galaxies that are below the surface brightness threshold. Our surveys are constrained by surface brightness dimming because we choose only galaxies that are large enough in angle to reveal the main star formation structure. We are therefore not magnitude-limited in the usual sense, but surface brightness limited. This surface brightness limit, combined with the self-imposed minimum pixel size for our study of star formation (10 pixels diameter) leads to a faintest total magnitude, but this is not like a magnitude-limited survey because our faintest magnitude depends on galaxy size.

What we find as a result of these limitations is that the average measured mass of a clump increases with redshift out to $z \sim 1.2$ in our GOODS survey, but the average clump mass per unit galaxy luminosity is approximately invariant (Fig. 4, from [21]). The average clump mass M is large, between 10^6 and $10^{8.5}\,\mathrm{M}_\odot$ depending on z, and the average clump mass divided by the galaxy luminosity is also large, with $M/10^{-0.4B_{\rm rest}} \sim 0.1 - 1$ for $z = 0.1$–1.2. In local galaxies, a giant star complex may contain $10^5\,\mathrm{M}_\odot$ in a galaxy with $B_{\rm rest} \sim -20.3$ mag. Then locally, $M/10^{-0.4B_{\rm rest}} \sim 10^{-3}$.

Possible reasons for the clump mass increase with z might be (1) clumps blend as physical resolution gets worse, (2) the Jeans length increases with increasing turbulent speed at increasing z, (3) the absolute clump surface brightness increases at the detection threshold, and (4) the average galaxy luminosity increases with increasing cosmological volume.

Some of these possibilities do not seem likely, given the constant ratio of $M/10^{-0.4B_{\rm rest}}$. This constant ratio means that clump blending is not causing the increase in clump mass with z because then the clump mass observed within a given resolved galaxy would increase, but the galaxy luminosity would not. That is, the galaxies are resolved so the clump separation has to be resolved too for a constant $M/10^{-0.4B_{\rm rest}}$ ratio. Second, the increase in turbulence with z cannot explain the constant $M/10^{-0.4B_{\rm rest}}$: as for clump blending, increased turbulence for a given galaxy luminosity would increase the clump mass for that galaxy and therefore increase $M/10^{-0.4B_{\rm rest}}$. Third, both clumps and galaxies have surface brightness

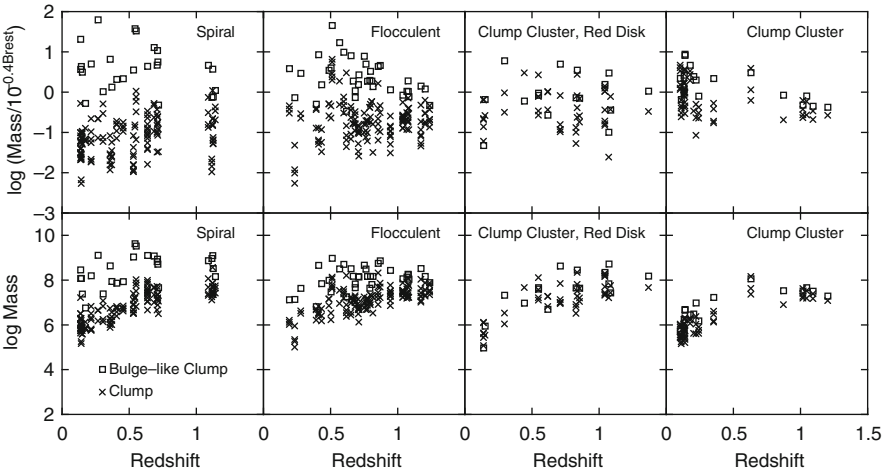


Fig. 4 Clump mass is shown versus redshift in the *bottom panels*, and clump mass per unit galaxy luminosity is shown in the *top panels*. The clump mass increases with z because surface brightness dimming causes the galaxy mass in our survey to increase with z and clump mass scales with galaxy mass

limits for detection. If clump mass increases but $M/10^{-0.4B_{\rm rest}}$ does not, then the clump surface brightness has to increase in inverse proportion to the relative clump area for a given galaxy size. This is not observed, as clump clusters have similar morphologies with increasing redshift.

The best explanation for the increase in M with constant $M/10^{-0.4B_{\rm rest}}$ is that the selected galaxy luminosity and mass increase with z. More massive galaxies have more massive clumps. Increasing galaxy luminosity is probably the actual selection effect at work: we tend to observe galaxies close to the surface brightness detection limit (there are relatively few high above that limit) and this limit gets intrinsically brighter with z. Thus the only galaxies we can observe at high z are the more massive galaxies, which tend to contain more massive clumps. The implication is that there are many fainter galaxies, possibly equally large in physical size and also with star-forming clumps, that we have not observed yet. If most of today's galaxies start in a clumpy phase, then all but the most massive and highest surface brightness of these could be missing from existing deep surveys.

A second trend in the observations is a decrease in average clump age with redshift [21], something like age $\propto (1+z)^{-4}$ (see Fig. 5). This is not the result of a younger universe because the trend would be different. It is more likely the result of bandshifting for stellar populations with mixed ages. For main sequence stars, the stellar lifetime is roughly proportional to the stellar surface temperature to the negative 4 power. The stellar surface temperature, in turn, is proportional to the inverse of the wavelength at peak emission. For bandshifting, restframe wavelength decreases as $(1+z)^{-1}$. Thus mean detectable population age decreases with increasing redshift as $(1+z)^{-4}$. This is about the observed trend in the figure. The

interclump age (Fig. 5, right hand side) is more constant than the clump age with z because of a selection effect: the "interclump" is defined to be whatever avoids star-forming regions. When we pick the interclump regions, we choose the reddest parts of the galaxies.

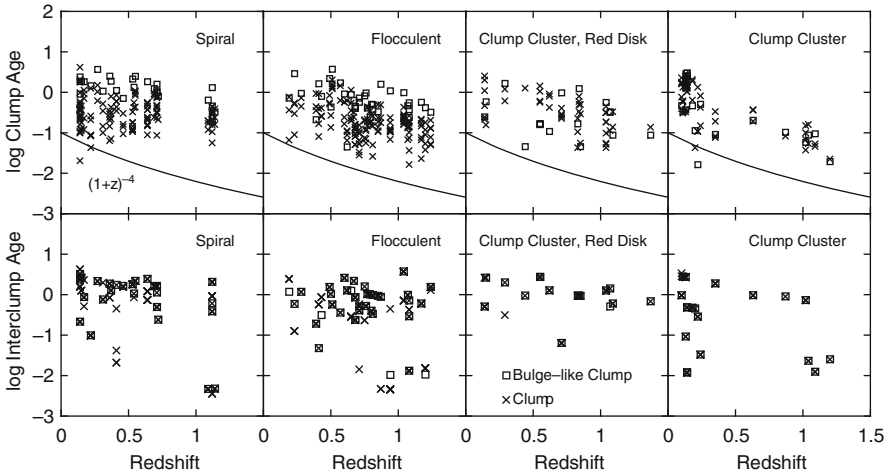


Fig. 5 Interclump age is shown versus redshift in the *bottom panels* and clump age is shown in the *top panels*. The clump age decreases slightly with z because younger populations are emphasized when the restframe wavelength is shorter. A curve representing the trend $(1+z)^{-4}$ is shown

Combining these two trends for mass and age, we get a third trend for the average star formation rate in a clump. Average star formation rate is defined to be the clump mass divided by the age. This increases sharply with z because the clump mass increases and the clump age decreases. The observed trend in star formation rate per clump is about as $(1+z)^8$. This comes about as follows: the clump mass is proportional to the galaxy luminosity (see above), which is approximately the surface brightness limit times the physical area limit. The surface brightness limit goes as $(1+z)^4$ and the area limit goes as z^2 for small z and barely separated clumps. Only the first arises if the clumps are well separated. At the same time, clump age decreases as $(1+z)^{-4}$. The ratio gives the observed $(1+z)^8$ dependence if the clumps are well separated and a stronger redshift dependence if they blend.

The product of the age and the dynamical rate, $(G\rho)^{1/2}$, is more constant with z than the star formation rate [21]. This is because the clump density is M/R^3 and $R \sim (M/\Sigma)^{1/2}$ for clump mass M, size R, and surface density Σ. The product of age and dynamical rate is therefore $\propto (1+z)^{-2}z^{-1/2}$ if resolvable size $\sim z$. This function has a weaker change with z than the previous function, $(1+z)^8$.

Finally, there is a redshift trend with clump and interclump mass surface density, in just the manner expected from selection effects, i.e., both get denser as $(1+z)^4$ because the apparent, uncorrected, surface brightness stays constant at about the detection threshold.

5 Conclusions

Selection effects for high z-band shifting enhance UV-bright star forming clumps, and de-enhance red bulges and red underlying exponential disks. These effects are not enough to explain the unusually clumpy, high-z morphology because this morphology is present even at moderate redshifts where the instruments view optical bands. Clumps in moderate to high-redshift galaxies are more massive than star-forming clumps are in local galaxies, when compared with their galaxy's luminosities. This increased relative mass presumably arises because the ISM turbulent speed is high compared to the rotation speed and because the disk gas fraction is high. The same high relative star formation masses occur in local Dwarf Irregulars and in Luminous Compact Blue Galaxies, which also have relatively high turbulent speeds and gas masses.

Bandshifting favors measurement of lower ages at higher z in mixed populations because UV and blue light emphasizes younger stars. Stellar main sequence age scales approximately with the inverse fourth power of its surface temperature and the dominant wavelength for emission from a region scales inversely with temperature. Thus the dominant age observed for a mixed population scales approximately with $(1 + z)^{-4}$.

Surface brightness dimming causes a selection effect in which galaxies with normal or low star formation rates per unit area are not included in intermediate-to-high redshift surveys. This dimming contributes to the observation that clump masses are large in an absolute sense because the only galaxies that can be observed with resolvable sizes are also intrinsically bright and massive.

Surface brightness dimming combined with bandshifting that favors low ages affect the derived star formation rates approximately as $(1 + z)^8$. This is a biased result from two selection effects. It is not that the derived star formation rates for individual clumps are wrong, but only that we observe only the rates that are above a z-dependent threshold.

Angular resolution affects the definition of a clump, but not the clump mass if individual clumps are well separated from each other. This appears to be the case for the most clumpy galaxies. Resolution also produces no obvious trends from selection bias in the range $z = 1$–5 because the point spread function measured in kiloparsecs is about constant.

Even with these selection effects, it is clear that we are observing a new morphology of galaxies at approximately $z > 0.5$. These galaxies are almost entirely clumpy with clump masses of $\sim 10^7 - 10^8\,\mathrm{M}_\odot$, clump ages of ~ 100 Myr, clump sizes of ~ 1 kpc or less, depending on resolution, and clump surface densities in a range from 10 to $100\,\mathrm{M}_\odot\ \mathrm{pc}^{-2}$. The surface densities are like those in whole galaxy disks today. The star formation rates in clumps are approximately 0.1–1 $\mathrm{M}_\odot$ year^{-1}. The galaxies have little central disk concentration, although there are bulges seen in the instrument IR in between 1/3 and 1/2 of the galaxies. They have weak or no exponential disks, but some have underlying red disks that are approximately exponential. There is no symmetry in these galaxies and no obvious global spiral waves.

References

1. Abraham, R.G., Tanvir, N.R., Santiago, B.X., Ellis, R.S., Glazebrook, K., van den Bergh, S. 1996, MNRAS, 279, L47
2. Agertz, O., Teyssier, R., Moore, B. 2009, MNRAS, 397, L64
3. Barden, M., Jahnke, K., Häussler, B. 2008, ApJS, 175, 105
4. Bournaud, F., Elmegreen, B.G. 2009, ApJ, 694, L158
5. Bournaud, F., Elmegreen, B.G., Elmegreen, D.M. 2007, ApJ, 670, 237
6. Bournaud, F., Elmegreen, B.G., Martig, M. 2009, ApJ, 707, L1
7. Carroll, S.M., Press, W.H., Turner, E.L. 1992, ARAA, 30, 499
8. Casini, C., Heidmann, J. 1976, A&A, 47, 371
9. Casini, C., Heidmann, J. 1976, A&ASS, 24, 473
10. Conselice, C.J., Blackburne, J.A., Papovich, C. 2005, ApJ, 620, 564
11. Cowie, L., Hu, E., Songaila, A. 1995, AJ, 110, 1576
12. Daddi, E., et al. 2010, ApJ, 713, 686
13. Dekel, A., Sari, R., Ceverino, D. 2009, ApJ, 703, 785
14. Elmegreen, B.G., Elmegreen, D.M. 2005, ApJ, 627, 632
15. Elmegreen, B.G., Elmegreen, D.M. 2006, ApJ, 650, 644
16. Elmegreen, D.M., Elmegreen, B.G. 2006, ApJ, 651, 676
17. Elmegreen, D.M., Elmegreen, B.G., Hirst, A.C. 2004, ApJ, 604, L21
18. Elmegreen, D.M., Elmegreen, B.G., Rubin, D.S., Schaffer, M.A. 2005, ApJ, 631, 85
19. Elmegreen, D.M., Elmegreen, B.G., Ferguson, T., Mullan, B. 2007, ApJ, 663, 734
20. Elmegreen, B.G., Elmegreen, D.M., Fernandez, M.X., Lemonias, J.J. 2009, ApJ, 692, 12
21. Elmegreen, D.M., Elmegreen, B.G., Marcus, M.T., Shahinyan, K., Yau, A., Petersen, M. 2009, ApJ, 701, 306
22. Förster Schreiber, N.M. et al. 2009, ApJ, 706, 1364
23. Garland, C.A., Pisano, D.J., Williams, J.P., Guzmán, R., Castander, F.J., Sage, L.J. 2007, ApJ, 671, 310
24. Heckman, T.M., et al. 2005, ApJ, 619, L35
25. Maehara, H., Hamabe, M., Bottinelli, L., Gouguenheim, L., Heidmann, J., Takase, B. 1988, PASJ, 40, 47
26. Overzier, R.A. et al., 2008, ApJ, 677, 37
27. Overzier, R.A. et al., 2009, ApJ, 706, 203
28. Overzier, R.A., Heckman, T.M., Schiminovich, D., Basu-Zych, A., Goncalves, T., Martin, D.C., Rich, R.M. 2010, ApJ, 710, 979
29. Petty, S.M., de Mello, D.F., Gallagher, J.S., III, Gardner, J.P., Lotz, J.M., Mountain, C.M., Smith, L.J. 2009, AJ, 138, 362
30. Spergel, D.N. et al., 2003, ApJS, 148, 175
31. Tacconi, L.J., et al. 2010, Nature 463, 781
32. Wolf, C., Hildebrandt, H., Taylor, E.N., Meisenheimer, K. 2008, A&A, 492, 933
33. Yang, Y., et al. 2008, IMAGES. I. A&A, 477, 789
34. Yoachim, P., Dalcanton, J.J. 2006, AJ, 131, 226

Supernovae, Dust, and Cosmology

Brian P. Schmidt and Joerg Fischera

Abstract Supernova of type Ia, exploding white dwarfs near the Chandrasekhar limit of 1.4 solar masses, have gained tremendous importance due to their utility as cosmological distance indicators. Observations of these events have revealed the unexpected and astonishing discovery of an accelerating Universe. The interpretation and the use of type Ia supernovae as standard candles is, however, not without problems. One of the largest corrections and also largest uncertainty is the correction for dust attenuation caused by dust grains possibly located in the interstellar medium of the host galaxy. An additional, and largely neglected problem arises from light scattered on dust grains which are distributed in interstellar dust clouds in the line of sight between observer and supernova. Based on a model of self-gravitating clouds we discuss the contribution of the scattered light on the supernova light curves and will show how the distance estimates to the supernova events are affected using standard estimation tools. It is shown based on simple statistical considerations that dust scattered light might cause a serious problem if a significant fraction of the supernova events are interspersed within the in-homogenous dusty medium.

1 Introduction

Type Ia Supernovae have played a seminal role in the rise of the Concordance Cosmology Paradigm – the Universe is geometrically flat, and dominated by a Cosmological Constant (Λ) and Cold Dark Matter (CDM). Type Ia supernovae are characterized observationally by a largely homogenous photometric and spectroscopic evolution, which includes a rise to maximum light of 20 days, a strong silicon line near maximum light, a lack of hydrogen, and strong iron lines at late times. In total, these observations are consistent with the explosions of white dwarf

B.P. Schmidt (✉)
RSAA, Mount Stromlo Observatory, The Australian National University, Weston Creek, ACT 2611, Australia
e-mail: brian@mso.anu.edu.au

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7_39,

stars approaching the Chandrasekhar limit, although no direct proof of this has yet been found. Although the observed properties of SN Ia are relatively uniform, the class shows subtle variations in luminosity, shape of the light curve, and spectral signatures which are correlated with each other. The correlation of the shape of the light curve with the luminosity [17, 18] has transformed SN Ia into powerful tools for measuring cosmological distances. This empirical relation provides distances precise to better than 7% and the objects themselves are bright enough to be seen to redshifts greater than $z > 1.5$ with current instrumentation.

Since the 1930s, cosmologists have settled on a model of the Universe based around a homogenous, isotropic universe governed by the laws of general relativity. This model predicts the brightness of objects, as a function of their redshift, is determined solely by the amount of different species of matter (each with an equation of state) contained in the Universe. The abundance of each species of matter is compared to the critical density (ratio being given by Ω), with the equation of state (w) describing how the form of matter's density scales with redshift ($\rho \propto (1+z)^{3+3w}$).

When the first sample of well-measured SN Ia distances were extended to $z > 0.3$ in 1998, the objects were found to be, on average, fainter than expected for a Universe made up of normal gravitating material. The large inferred distances at a given redshift indicated the Universe was accelerating in its rate of expansion, something which can be accommodated by the standard model only if the Cosmos is dominated by a form of matter (called Dark Energy) with an equation of state near $w = -1$ (see review by Perlmutter and Schmidt [16]). At present, cosmological measurements from SN Ia, the CMB, large scale structure, and the Hubble constant are all consistent and provide increasingly precise constraints on the degree to which the cosmos is flat, the relatively abundances of baryons, cold dark matter, and dark energy, and the dark energy's equation of state. These measurements remain consistent with a Universe which is geometrically flat, is made up of 73% Dark Energy (with $w = -1$), 23% Dark Matter, and 4% baryons, a combination of parameters suggested in 1998 to explain the then available observations.

While the acceleration is attributed to the Universe being dominated by a form of Dark Energy whose negative pressure is repulsive under gravity, there is little understanding of the nature of the Dark Energy. Trying to determine if Dark Energy's equation of state varies from that of a cosmological constant (which has $w = -1$ at all times), or explain why it doesn't, is considered one of modern astronomy and physics' largest challenges. SN Ia continue to take a leading role in this science area, still providing the best measurements of Dark Energy's equation of state.

Although a leading method to measure cosmological distances today, SN Ia have their limitations. There are a substantial number of uncertainties, of which the correction for dust attenuation is not only the largest, but also the most uncertain. For example, it is not understood why the empirically derived extinction curve of SN Ia observations seems to be much steeper relative to the mean extinction in the Milky Way. Instead of an absolute to relative extinction $R_B = A_B/E(B-V) \approx 4$ [7] the data are more consistent with $R_V = 2$ which points to a larger ratio between the extinction in B and V.

As data sets have increased in size and precision, the problems with dust (and the associated difficulties in modelling SN Ia colours) have become acute. In 2007, Conley et al. [3] showed that the Hubble Bubble, seen in the analysis of nearby SN Ia data by Jha et al. [10] goes away if, instead of treating dust as having the canonical $R_B \approx 4$, choosing $R_B = 2$, as suggested by some SN Ia studies. Kessler et al. [11] find, depending on exactly how one treats dust, using the MLCS2k2 model [10] or SALT2 [8] model, the derived values for the equation of state parameter, w, can vary by ± 0.2.

To make progress with SN Ia to further refine our Cosmological measurement, it is necessary to consider systematic effects that could be neglected previously due to their relative size. Such an effect might be caused by dust scattered light from grains in interstellar clouds which are located in the host galaxy in the line of sight between observer and supernova. Previous work has found that dust scattered light can lead to a steeper extinction curve if the dust is associated with the supernova or is located close to it [20] (as the extinction curve becomes more close to the absorption curve which is intrinsically steeper). However, a reddening curve as steep as indicated could not be explained. This work explores the contribution of the scattered light to the light curves and provides an estimate of the expected systematic effect on the distance estimate. While such an effect will not make the accelerating Universe go away, it potentially could be a significant limitation to the efforts to constrain Dark Energy's properties via SN Ia in the future.

2 Dust-Scattered Light from SN

2.1 The Cloud Model

For the clouds we assumed that they are self-gravitating isothermal entities in pressure equilibrium with the surrounding medium. For the pressure in the surrounding medium we assumed $p_{\text{ext}}/k = 2 \times 10^4$ K/cm^3 which results from the magnetic, the turbulent, and the thermal pressure [2]. As in the work from Curry and McKee [4] the pressure caused by cosmic rays is neglected as the clouds appear optically thin to this kind of radiation. The density profile is for given pressure entirely determined by the extinction value through the cloud centre. Clouds with a small extinction value are rather flat while clouds with higher extinction values show a steep density profile at the outskirts [see, e.g. [6]]. Clouds which are critical to gravitational collapse are characterized by an extinction $A_V = 8$ mag consistent with values observed for Giant Molecular Clouds [5] and have an over-pressure or over-density in the cloud centre in respect to the edge of the cloud of 14.04. In this work we consider only clouds with extinction A_V in the range from 0.1 to 3 mag with the focus on clouds with A_V less than 1 mag. Those clouds are characterized with a rather flat density profile. We considered the gas in the clouds to be nonmolecular but still non-ionized. In case of isothermal self-gravitating spheres the radius (see Table 1) for given extinction is proportional to the cloud temperature for which we assumed either 50 or 100 K.

Table 1 Assumed cloud sizes for $T_{\rm cl} = 100$ K

$[A_V]_{\rm cl}$ [mag]	0.10	0.175	0.25	0.50	1.00	1.50	2.00	3.00
$R_{\rm cl}$ [pc]	0.17	0.30	0.43	0.83	1.44	1.92	2.19	2.44

2.2 The Dust Model

The contribution of the dust scattered light to the SN light curves depends critical on the optical properties of the dust grains for which we adopted the values derived by Weingartner and Draine [21]. They are by construction consistent with the observed mean properties of dust in the diffuse phase of the Milky Way, the extinction curve, the dust albedo, the g-factor which is the cosine weighted scattering angle of non-polarized light, and the depletion factors of key elements assumed to be condensed in grain species as silicates, graphites, and PAH molecules. The model is furthermore in broad agreement with the observed IR emission [12] as observed with the DIRBE experiment. In the mean the dust extinction in the optical wavelength regime is mostly caused by dust scattering and less by dust absorption. This is reflected by a rather high dust albedo close to 0.6 as determined by Lillie and Witt [13] and Morgan et al. [14]. In addition the light is strongly scattered in forward direction which gives a rather large value for the g-factor around 0.6. Both quantities, the high dust albedo and the high g-factor, lead to a boost of the dust scattered light from grains located *in front* of the SN event.

2.3 The Light Echo

Unless the scattering occurs exactly in direction towards the observer the scattered light shows a certain delay-time t relative to the dust-attenuated light. As consequence the observer sees not only the attenuated light of the intrinsic SN-spectrum L_λ but also the scattered light which has been emitted earlier on. The scattered light is therefore the convolution

$$L_\lambda^{\rm sca}(t) = \int {\rm dt}'\, L_\lambda(t - t')\, K_\lambda(t') \tag{1}$$

where $K_\lambda(t)$, the Kernel of the integral, is the scattered emission of a light flash and where $L_\lambda(t)$ is the light emitted at time t. To solve this problem we consider solely single scattered light. This approximation is quite accurate in case of low column densities where the probability for multiple scattering events becomes negligible.

The distance estimate is preliminarily based on measurements of the SN lightcurves around the peak in the luminosity curve (in B-band reached after ~ 18 days) up to 40 days after the explosion. The scattered emission becomes critical when the emission declines as a weaker signal is mixed with scattered light from the emission peak. To affect the luminosity curve the delay-time of the scattered light needs therefore to be quite short. This fact gives a strong restriction to the location

of the dust. Unless the dust is associated with the supernova or the supernova event occurs inside a dust cloud, the dust needs to be in front of the SN. For simplicity we assumed that the clouds are positioned exactly in the line of sight between observer and supernova.

In general the single scattered photons with the same delay time t originate from scattering centres located on the surface of an ellipsoid with the SN and the observer in the two focus points. If the distance to the observer is much larger than the distance $R(t, \vartheta)$ of the scattering centres to the SN – as is easily fulfilled if the SN occurs in an external galaxy – the shape is well described by a paraboloid with the supernova in the focus point:

$$R(t, \vartheta) = \frac{c\, t}{1 + \cos \vartheta} \tag{2}$$

Here, ϑ is the angle under which the light has been emitted at the location of the source and where $\vartheta = 0°$ refers to a direction lying behind the SN and $\vartheta = 180°$ to a direction towards the observer. c is the velocity of light.

In our axial symmetric approximation the scattered light of a light flash observed between t and $t + \mathrm{d}t$ is given by the integral over all scattering events lying on the corresponding paraboloid:

$$K_\lambda(t)\mathrm{d}t = c\,\mathrm{d}t \int \frac{\mathrm{d}R}{2R} n_\mathrm{H}(R(t))\, C_\mathrm{ext}(\lambda)\omega_\lambda \Phi_\lambda(\vartheta_\mathrm{sca}(R(t)))\, e^{-\tau_1(\lambda)}\, e^{-\tau_2(\lambda)} \tag{3}$$

where $C_\mathrm{ext}(\lambda)$ is the extinction cross section per hydrogen atom, ω_λ the dust albedo, and n_H the hydrogen density. $\tau_1(\lambda)$ and $\tau_2(\lambda)$ are the extinction values from the SN event to the scattering location and from the scattering location to the observer. $\Phi_\lambda(\vartheta_\mathrm{sca})$ is the phase function, the probability that a photon is scattered in an angle ϑ_sca away from the initial direction. We used the smooth phase function from Henyey and Greenstein [9] which depends apart from the scattering angle solely on the g-factor. For the time-varying SN spectrum we have taken a template which has been derived from a large sample of SN spectra [15].

The main parameter which, apart from the extinction and the dust temperature, determines the scattered emission is the distance D_cl of the clouds. For convenience we measured the distance in units of the cloud radius R_cl. We assumed five distances from $D_\mathrm{cl} = 1$ to $D_\mathrm{cl} = 32$ equally spaced in $\Delta \ln D_\mathrm{cl} = \ln 2$. The closest position of a supernova event to a cloud is therefore a supernova that explodes exactly behind a cloud.

The total flux of our model is given by

$$L_\lambda^\mathrm{model}(t, \tau_\lambda, D_\mathrm{cl}) = L_\lambda(t)\, e^{-\tau_\lambda} + L_\lambda^\mathrm{sca}(t, \tau_\lambda, D_\mathrm{cl}) \tag{4}$$

where τ_λ is the extinction through the cloud centre. The fluxes are transfered to astronomical units using

$$m^{\rm model}_{\lambda_{ref}} - V_0 = -2.5 \log \left\{ \int {\rm d}\lambda \, R_{\lambda_{ref}}(\lambda) \, L^{\rm model}_{\lambda}(t, \tau_\lambda, D_{\rm cl}) \right\} \tag{5}$$

where $R_{\lambda_{ref}}(\lambda)$ is the filter response function and λ_{ref} the corresponding reference wavelength. The magnitudes are derived using the Bessel filters and are based on the Vega system. The template of the SN spectra is normalized to the maximum magnitude in the optical V-band filter so that we have in case of an un-attenuated SN spectrum at maximum $m_V - V_0 = 0$ mag.

Figure 1 shows for a cloud extinction $A_V = 0.5$ mag the derived scattering light curves in B-band. The curves are compared with the attenuated light without any contribution of scattered emission. Early on the scattering light curves show a strong increase in flux which reflects the steep increase seen in SN light curve. The scattered flux reaches its maximum several days after the SN light. At late times the scattered light is dominated by the scattered emission from the peak of the SN luminosity curves as long as the corresponding scattering surface crosses the clouds interior. The scattering lightcurves for clouds located relatively close to the SN event are therefore rather flat while they decrease very early on for more distant clouds.

To achieve an accuracy of 1% in the obtained luminosity curves up to 40 days after outburst, the scattered emission is critical if the clouds are as far as 10 cloud radii away from the SN.

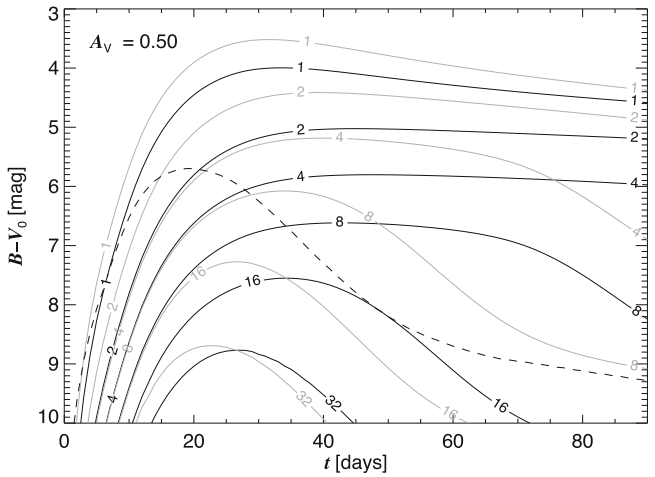


Fig. 1 Scattering light curves in the Bessel B-band for various distances $D_{\rm cl}$ from the cloud centre to the SN event. The curves are labelled with the corresponding distance. The cloud extinction through the cloud centre is $A_V = 0.5$ mag. The *solid black* and *grey lines* are calculations using a cloud temperature of 100 and of 50 K, respectively. The *dashed line* gives the 1% flux of the attenuated SN light

3 Effect on the Distance Estimate

As we have seen for small cloud distances the relative contribution of the scattered light to the total flux increases as function of time. The peak of the luminosity curves appears therefore broader than it actually is. Because of the broader peak the supernova is wrongly identified as a more luminous one. This leads to an overestimate of the distance to produce the same observed flux.

The dust-scattered emission affects furthermore the reddening of the supernova spectrum. The relative contribution of the scattered light to the attenuated emission is stronger at shorter wavelengths so that the light appears to be less reddened by dust grains. This is visualized in Fig. 2 which shows the difference in reddening between the simple attenuated emission spectrum and our model. As the dust correction is underestimated the supernova appears again to be more distant.

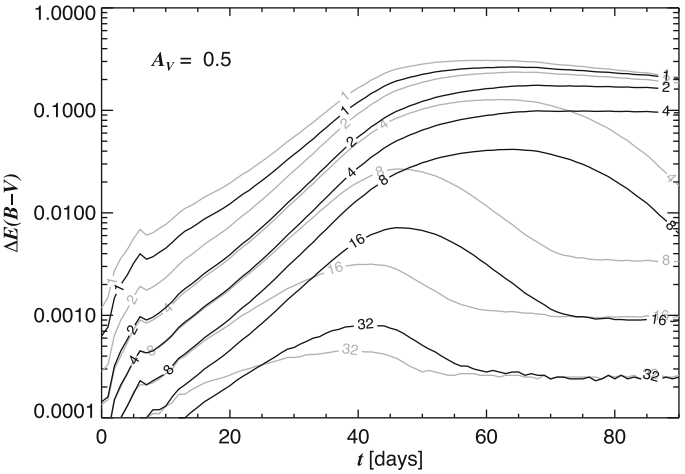


Fig. 2 Difference in reddening $\Delta E(B-V) = E(B-V) - E(B-V)^{\text{model}}$ of the dust-attenuated light $E(B-V) = A_B - A_V$ (without scattering) and the dust-attenuated light where we added the scattered emission of the model $E(B-V)^{\text{model}}$

The obtained offsets for a variety of model assumptions using a standard distance estimator tool (MLCS2k2, which uses multi-color light curve shapes [10] are shown in Fig. 3. The results are based on non-redshifted spectra using all Bessel filters of the $UBVRI$ system. The overestimate of the distance to the supernovae can be quite substantial. However, the effect decreases strongly with distance and disappears for small extinction values. For 50 K the effect increases furthermore as function of the cloud extinction $[A_V]_{\text{cl}}$. While the behaviour for 100 K is the same at small extinction values, it deviates above $A_V = 0.5$ mag where the effect starts to disappear. The explanation might lie in the weighting of the fluxes used in the distance estimator tool.

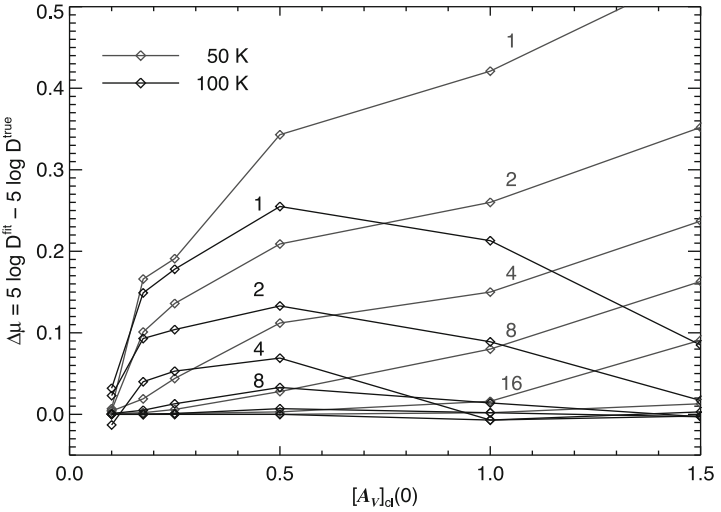


Fig. 3 Offsets of the distance modulus obtained by using a standard MLCS distance estimator tool. The offsets are shown as function of the cloud extinction $[A_V]_{\rm cl}$ through the cloud centre. Values corresponding to the same cloud distance $D_{\rm cl}$ (given in units of the corresponding cloud radius) are connected with *straight lines*

4 Probability Consideration

To derive an estimate of the mean effect on the distance modulus for a large sample of SN distances we assume that one kind of progenitors of the type Ia supernovae is well mixed within the thin dusty disc assumed to be infinite in diameter. The in-homogeneous density structure of the interstellar medium in this infinite slab is described by an idealized two-phase model where the cold clouds of relatively high density are randomly distributed in a medium of low-density gas. For simplicity the clouds are assumed to be all of the same size. A main parameter in such a two-phase model is the actual volume occupied by the clouds. In case of randomly distributed mono-sized clouds this filling factor is given by

$$ff = 1 - e^{-n_{\rm cl} V_{\rm cl}} \tag{6}$$

where $n_{\rm cl}$ is the number density of the clouds and where $V_{\rm cl} = (4/3)\pi R_{\rm cl}^3$ is the cloud volume. Observations of the clouds in the interstellar medium point to a filling factor of $\sim$ 10 %. The filling factor in our model might deviate considerably from this value because of the simplifications made. To take those into consideration we explored a larger range of filling factors down to values as low as 1%.

As a first approach the extinction caused by the in-homogeneous medium is described by the mean extinction. A more elaborate and more accurate description based on Poisson statistic provides essentially the same solution. If we neglect

the density profile of the clouds the mean extinction caused by such a two-phase medium at a distance L is given by

$$\langle A_V \rangle = \left\{ (1 - ff)[n_{\rm H}]_{\rm icl} R_{\rm cl} R_V \frac{E(B-V)}{N_{\rm H}} - \ln(1 - ff) 0.5 [A_V]_{\rm cl} \right\} \frac{L}{R_{\rm cl}} \quad (7)$$

where $R_V \approx 3.1$ and where $E(B-V)/N_{\rm H}$ is the "dust-to-gas ratio" which is found to be $(1/5.8)\, 10^{-21}$ cm^2 [1]. If we identify the inter-cloud material with the WNM with typical temperatures in the order of 10^4 K and assume pressure equilibrium between the different phases, the density in the inter-cloud medium is $[n_{\rm H}]_{\rm icl} \approx 1$ cm^{-3}. We find that for small filling factors the mean extinction is dominated by the dense gas. For the mean vertical extinction through the infinite slab we assumed $\langle A_V \rangle_{\rm slab} = 1$ mag roughly consistent with the observed vertical column density of neutral hydrogen in galactic discs.

The probability of a certain extinction value of randomly distributed sources inside the infinite slab seen from all possible angles is

$$p(x) = \begin{cases} 0.5 & \text{for } x < 1 \\ 0.5\, x^{-2} & \text{for } x > 1 \end{cases} \quad (8)$$

where $x = \langle A_V \rangle / \langle A_V \rangle_{\rm slab}$. As the distribution overall reflects the observed probabilities of extinction values of SN events, we believe that the model is sufficiently accurate. Important for cosmology are events with extinction values less than $A_V = 1$ mag. Events showing a higher attenuation are generally neglected to avoid inaccuracies caused by the uncertain reddening curve. In our model we consider therefore only half of all SN events.

The number of clouds in the line of sight varies strongly with location inside the slab and the viewing angle to the disc. However, the offsets obtained for the distant modulus decrease strongly with cloud distance. For simplicity we base therefore our estimate of the mean offset on the clouds closest to the supernova event. The probability to place at a certain distance $D_{\rm cl}$ the first cloud in the line of sight can be determined by using the covering factor or accumulative probability $p(< D_{\rm cl})$ to have at distance $D_{\rm cl}$ at least one single cloud in the line of sight. The derived normalized accumulative probability distributions $p(< D_{\rm cl})$ for certain filling factors ff and cloud sizes (extinction values) are shown in Fig. 4. The finite probability at large distances is caused by edge effects of the infinite slab. They are stronger for larger clouds and higher filling factors. We see that the probability to have at least one cloud at a distance $D_{\rm cl} = 10$ is relatively high with $\sim 50\%$ for a filling factor of 10%.

The mean offset is given by

$$\langle \Delta\mu \rangle = \int_0^\infty {\rm d}D_{\rm cl}\, \Delta\mu(D_{\rm cl}) \frac{{\rm d}p(< D_{\rm cl})}{{\rm d}D_{\rm cl}} \quad (9)$$

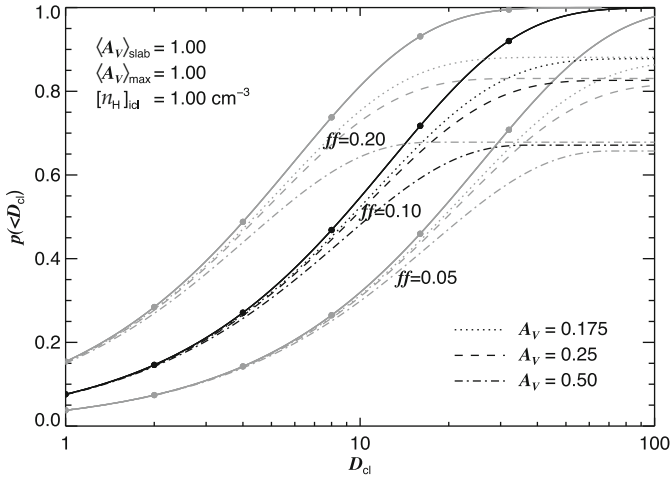


Fig. 4 Covering factor or accumulative probability to have in an infinite slab at distance $D_{\rm cl}$ at least one cloud in the line of sight. The filling factors of the clouds are assumed to be 0.05, 0.10, and 0.20. The density in the inter-cloud medium is assumed to be $[n_{\rm H}]_{\rm icl} = 1$ cm^{-3}. The vertical mean optical depth and the maximum optical depth taken into account in the probability calculation are assumed to be $\langle A_V \rangle_{\rm slab} = 1$ mag and $\langle A_V \rangle_{\rm max} = 1$ mag. The extinction A_V through the cloud centres is either 0.175, 0.25, or 0.50 mag. The corresponding probabilities shown as *dotted*, *dashed*, and *dashed–dotted lines* are compared with a medium without boundaries (*solid lines*)

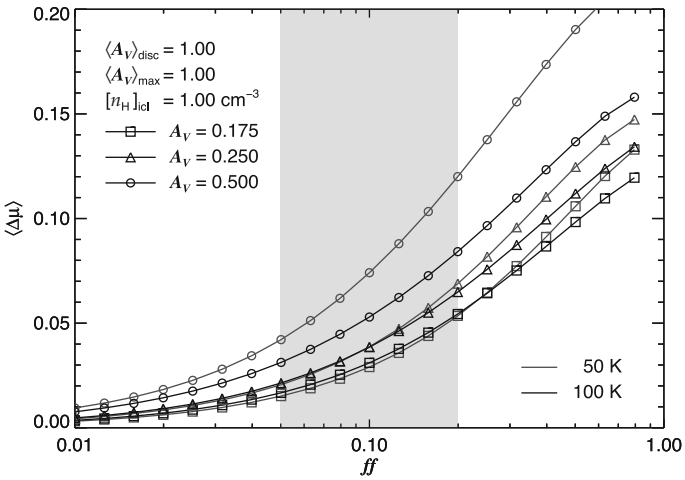


Fig. 5 Mean offset of the distance modulus for several assumptions of cloud extinction as function of filling factor of the clouds. The cloud temperature is assumed to be either 50 or 100 K. The parameters of the infinite slab are the same as in Fig. 4

The result for several assumptions of the cloud sizes (extinction values) for a large range of filling factors is shown in Fig. 5. The offsets for $ff = 10\%$ lie in the regime from 0.03 to 0.07. They are generally larger for more optically thick clouds and decrease considerably for smaller filling factors.

5 Conclusion

The dust-scattered emission caused by interstellar clouds in the line of sight between supernova and observer which are located in the host galaxy can have although a small but a non-negligible effect on the observed light curves. In those cases scattered light produces an apparently broader peak in the light curves and a less reddened SN spectrum. Both effects lead in turn to a positive offset in the distance estimate.

In general the dust needs to be quite close to the supernova event to have an effect. In the local Universe the distribution of the supernova seems to be comparable with the old stellar population with a scale height much larger than the one of the thin disc of normal galaxies. In this case the mean effect will be negligible. However, observations indicate that a certain fraction of type Ia SN is associated with the young stellar population which is therefore well mixed with the dust. Our understanding of the star formation rate as function of redshift furthermore suggests that this type is also more frequent at higher redshifts [19]. This implies that the effects discussed in this work become systematically more important with cosmological distance. It does not appear unreasonable that at redshift $z = 1$ roughly half of all events occurs in the thin dusty disc. Based on our simple in-homogeneous model of the interstellar medium those supernovae events may have a mean offset as high as $\langle \mu \rangle \approx 0.03$.

References

1. Bohlin, R.C., Savage, B.D., Drake, J.F. 1978, ApJ, 224, 132
2. Boulares, A., Cox, D.P., 1990, ApJ, 365, 544
3. Conley, A., Carlberg, R.G., Guy, J., Howell, D.A., Jha, S., Riess, A.G., Sullivan, M. 2007, ApJ, 664, L13
4. Curry, C.L., McKee, C.F. 2000, ApJ, 528, 734
5. Dopita, M.A., Sutherland, R.S. 2003, Astrophysics of the Diffuse Universe, Springer, Berlin/New York
6. Fischera, J., Dopita, M.A. 2008, ApJS, 176, 164
7. Fitzpatrick, E.L. 1999, PASP, 111, 63
8. Guy, J. et al. 2007, A&A, 466, 11
9. Henyey, L.G., Greenstein, J.L. 1941, ApJ, 93, 70
10. Jha, S., Riess, A.G., Kirshner, R.P. 2007, ApJ, 659, 122
11. Kessler, R., et al. 2009, ApJS, 185, 32
12. Li, A., Draine, B.T. 2001, ApJ, 554, 778
13. Lillie, C.F., Witt, A.N. 1976, ApJ, 208, 64
14. Morgan, D.H., Nandy, K., Thompson, G.I. 1976, MNRAS, 177, 531
15. Nugent, P., Kim, A., Perlmutter, S. 2002, PASP, 114, 803

16. Perlmutter, S., Schmidt, B. 2003, In: Supernovae and Gamma Ray Bursts, ed. K. Weiler, Springer, Berlin
17. Phillips, M.M. 1993, ApJ, 413, L105
18. Riess, A.G., Press, W.H., Kirshner, R. 1995, ApJ, 438, L17
19. Sulivan, M., Le Borgne, D., Pritchet, C.J. et al. 2006, ApJ, 648, 868
20. Wang, L. 2005, ApJ, 635, L33
21. Weingartner, J.C., Draine, B.T. 2001, ApJ, 548, 296

Endpiece – On Location in Dead Vlei

Location:
Dead Vlei, a white clay pan in the Namib Desert.

GPS coordinates:
24°45′45.99″ S, 15°17′39.44″ E

Description:
Trees in Dead Vlei, devoid of groundwater, are believed to have died about 1400 AD (Figs.1 and 2). Two delegates awed by the morphology of skeleton trees in the Vlei (Fig. 3).

D.L. Block et al. (eds.), *Galaxies and their Masks*,
DOI 10.1007/978-1-4419-7317-7,

Fig. 1 "The dry river-bed finds no thanks for its past." From the writings of Nobel laureate Rabindranath Tagore. Photo: David Block

Fig. 2 "My evening came among the alien trees and spoke in a language which my morning stars did not know." Penned by Nobel laureate Rabindranath Tagore. Photo: David Block

Fig. 3 "The mighty desert is burning for the love of a blade of grass who shakes her head and laughs and flies away." Penned by Nobel laureate Rabindranath Tagore. Photo: Bruce Elmegreen

The co-chairs of the Scientific Organizing Committee for the conference “Galaxies and their Masks” were Kenneth C. Freeman and David L. Block, seen here together with Lord and Lady Rosse at Birr Castle in Ireland. Some of the most exquisite early drawings of dusty masks within the Milky Way - *Shrouds of the Night* - were produced at Birr Castle. Photograph: Robert Groess

List of Participants

Roberto G. Abraham Department of Astronomy and Astrophysics, University of Toronto, Toronto, ON L6M 2N5, Canada, abraham@astro.utoronto.ca

Borja Anguiano Astrophysikalisches Institut Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany, baj@aip.de

Hong Bae Ann Pusan National University, Busan, Korea, hbann@pusan.ac.kr

Magda Arnaboldi European Southern Observatory, Garching, Germany; and Observatory of Turin, INAF, Turin, Italy, marnabol@eso.org

Martin Asplund Max-Planck-Institut für Astrophysik, Garching, Germany, asplund@mpa-garching.mpg.de

Maarten Baes Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281 S9, B-9000 Gent, Belgium, maarten.baes@ugent.be

John E. Beckman Instituto de Astrofísica de Canarias, Spain; Consejo Superior de Investigaciones Científicas, Spain; Departamento de Astrofísica, Universidad de La Laguna, Spain, jeb@iac.es

David L. Block School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa, igalaxy@iafrica.com; david.block@wits.ac.za

Jan J. Blom Physics and Astronomy, Springer, New York, USA, harry.blom@springer.com

Jean Brodie UCO/Lick Observatory, University of California, Santa Cruz, CA, USA, brodie@ucolick.org

Harvey Butcher RSAA, Mount Stromlo Observatory, The Australian National University, Weston Creek, ACT, Australia, butcher@mso.anu.edu.au

Gayandhi De Silva Anglo Australian Observatory, Sydney, NSW 2122, Australia, gdesilva@aao.gov.au

Ruben J. Díaz Gemini Observatory, AURA, Colina El Pino, La Serena, Chile, rdiaz@gemini.edu

Bruce G. Elmegreen IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, USA, bge@us.ibm.com

Jayanne English University of Manitoba, Winnipeg, MB, Canada R3T 2N2, jayanne_english@umanitoba.ca

Daniel Espada Instituto de Astrofísica de Andalucía – CSIC, Glorieta de la Astronomía s/n, 18080 Granada, Spain, daniel@iaa.es

Giovanni G. Fazio Harvard Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA, gfazio@cfa.harvard.edu

Annette Ferguson Institute for Astronomy, University of Edinburgh, Edinburgh, UK, ferguson@roe.ac.uk

Holland Ford Johns Hopkins University, Baltimore, MD, USA, ford@pha.jhu.edu

Kenneth C. Freeman Research School of Astronomy and Astrophysics, Mount Stromlo Observatory, ANU, Canberra, Australia, kcf@mso.anu.edu.au

Ortwin Gerhard MPE, Giessenbachstrasse, D-85748 Garching, Germany, gerhard@mpe.mpg.de

Trevor Gould Astronomer on Site, South Africa, trevgould@gmail.com

Carl J. Grillmair Spitzer Science Center, California Institute of Technology, Pasadena, CA 91125, USA, carl@ipac.caltech.edu

Robert Groess School of Computational and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa, robert.groess@wits.ac.za

Amina Helmi Kapteyn Astronomical Institute, University of Groningen, 9700 AV Groningen, The Netherlands, ahelmi@astro.rug.nl

Johan H. Knapen Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain; Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain, jhk@iac.es

Chiaki Kobayashi Research School of Astronomy and Astrophysics, The Australian National University; Mt. Stromlo Observatory, Weston ACT 2611, Australia, chiaki@mso.anu.edu.au

Konrad Kuijken Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands, kuijken@strw.leidenuniv.nl

Donald Lynden-Bell Institute of Astronomy, The Observatories, Cambridge CB3 0HA, UK, dlb@ast.cam.ac.uk

Francesca Matteucci Astronomy Division, Department of Physics, Trieste University, Via G.B. Tiepolo 11, 34134 Trieste, Italy; INAF, Trieste, Via G.B. Tiepolo 11, 34134 Trieste, Italy, matteucc@oats.inaf.it

Isabel Pérez Departamento de Física Teórica y del Cosmos, Universidad Granada, Granada, Spain; Instituto Carlos I de Física Teórica y Computación, Granada, Spain, isa@ugr.es

Daniel Pfenniger Geneva Observatory, University of Geneva, Sauverny, Switzerland, daniel.pfenniger@unige.ch

Ivânio Puerari INAOE, Santa María Tonantzintla, Mexico, School of Computational and Applied Mathematics, University of Witwatersrand, Johannesburg, South Africa, puerari@inaoep.mx

Mary E. Putman Columbia University, New York, NY 10027, USA, mputman@astro.columbia.edu

George Rhee Department of Physics and Astronomy, University of Nevada Las Vegas, 89154 NV USA, grhee@physics.unlv.edu

Brian P. Schmidt RSAA, Mount Stromlo Observatory, The Australian National University, Weston Creek, ACT 2611, Australia, brian@mso.anu.edu.au

Joseph Silk Department of Physics, Beecroft Institute for Cosmology and Particle Astrophysics, University of Oxford, Oxford OX1 3RH UK, j.silk1@physics.ox.ac.uk

Pippa Skotnes Centre for Curating the Archive, University of Cape Town, Cape Town, 7701, South Africa, pippa.skotnes@uct.ac.za

Thomas Y. Steiman-Cameron Indiana University, Bloomington, IN, USA, tomsc@astro.indiana.edu

R. Brent Tully Institute for Astronomy, University of Hawaii, Honolulu, Hawaii, tully@ifa.hawaii.edu

Petri Väisänen South African Astronomical Observatory, Cape Town, South Africa, petri@saao.ac.za

Pieter van der Kruit Kapteyn Astronomical Institute, University of Groningen, 9700AV Groningen, The Netherlands, vdkruit@astro.rug.nl

Ewine van Dishoeck Leiden Observatory, Leiden University, Leiden, The Netherlands, ewine@strw.leidenuniv.nl

Lourdes Verdes-Montenegro Instituto de Astrofísica de Andalucía – CSIC, Glorieta de la Astronomía s/n, 18080 Granada, Spain, lourdes@iaa.es

Richard Wielebinski Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany, rwielebinski@mpifr-bonn.mpg.de

Mary Williams Astrophysikalisches Institut Potsdam, Potsdam, Germany, mary@aip.de

Tim de Zeeuw European Southern Observatory, Garching, Germany, tdezeeuw@eso.org

CODA

David L. Block

As a CODA to this book “Galaxies and their Masks” - two final masks...

Alexander Solzhenitsyn penned these words:

“Not everything has a name. Some things lead us into a realm beyond words. Art thaws even the frozen, darkened soul, opening it to lofty spiritual experience. Through Art we are sometimes sent – indistinctly, briefly – revelations not to be achieved by rational thought.”

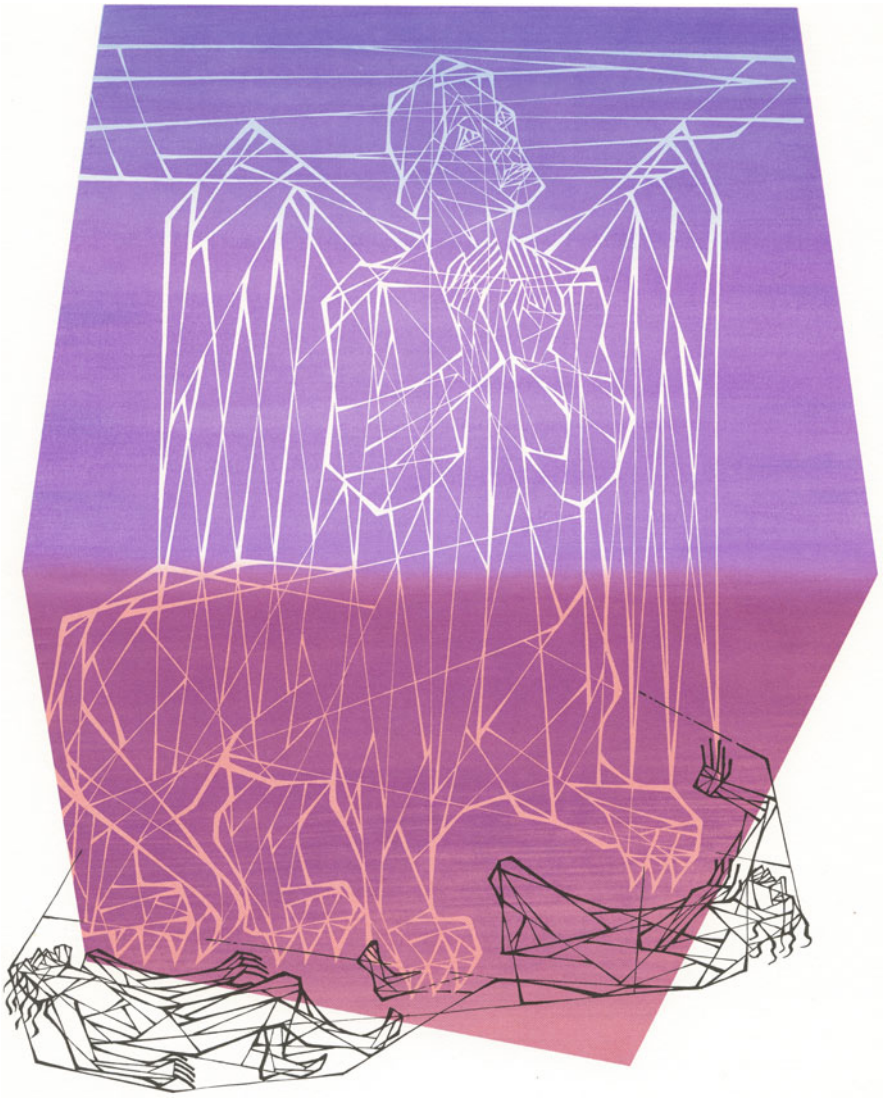


Fig. 1 Penetrating the mask of Man, as strikingly depicted by the artist Ismar David. “Man is neither angel nor beast; and the mischief is, that he who would be thought an angel, acts the beast” wrote Blaise Pascal. Courtesy Ismar David Collection, RIT Cary Graphic Arts Collection and the Limited Editions Club, New York

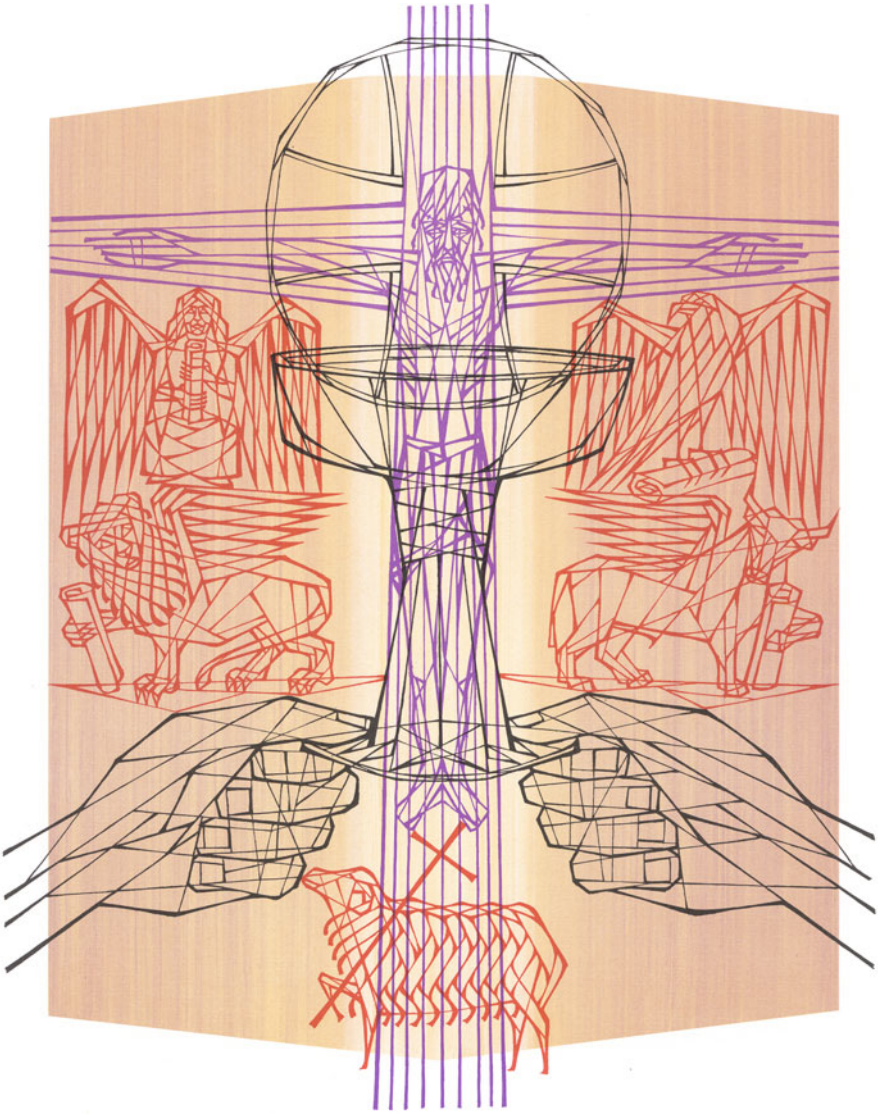


Fig. 2 Some may affirm that the mask of time was forever penetrated in the luminous figure of the Nazarene. "Before Abraham was, I Am" (John Chapter 8). From the hand of the great Jewish artist, Ismar David. Courtesy Ismar David Collection, RIT Cary Graphic Arts Collection and the Limited Editions Club, New York