ISSI Scientific Report 14

Vincent Minier · Roger-Maurice Bonnet
Vincent Bontems · Thijs de Graauw
Matt Griffin · Frank Helmich
Göran Pilbratt · Sergio Volonte

Inventing a Space Mission

The Story of the *Herschel* Space Observatory

INTERNATIONAL SPACE SCIENCE INSTITUTE

Springer

ISSI Scientific Report Series

Volume 14

More information about this series at http://www.springer.com/series/10151

Vincent Minier • Roger-Maurice Bonnet •
Vincent Bontems • Thijs de Graauw •
Matt Griffin • Frank Helmich •
Göran Pilbratt • Sergio Volonte

Inventing a Space Mission

The Story of the *Herschel* Space Observatory

Springer

Vincent Minier
Département d'astrophysique/Irfu
CEA Paris-Saclay
Gif-sur-Yvette, France

Roger-Maurice Bonnet
International Space Science Institute
Bern, Switzerland

Vincent Bontems
Larsim/Irfu
CEA Paris-Saclay
Gif-sur-Yvette, France

Thijs de Graauw
ESO-Chile
Santiago de Chile, Chile

Matt Griffin
School of Physics and Astronomy
Cardiff University
Cardiff, United Kingdom

Frank Helmich
SRON Netherlands Institute for
Space Research & University of Groningen
Groningen, The Netherlands

Göran Pilbratt
European Space Agency
ESTEC, Scientific Support Office
Noordwijk, The Netherlands

Sergio Volonte
International Space Science Institute
Bern, Switzerland

ISSI Scientific Report Series
ISBN 978-3-319-60023-9 ISBN 978-3-319-60024-6 (eBook)
DOI 10.1007/978-3-319-60024-6

Library of Congress Control Number: 2017942794

Cover figure: ESA's *Herschel* space observatory set against a background image of the Vela C star-forming region. The image was mapped using Herschel instruments PACS and SPIRE at wavelengths of 70, 160 and 250 microns.
Cover credit: © ESA/PACS & SPIRE Consortia, T. Hill, F. Motte, Laboratoire AIM Paris-Saclay, CEA/IRFU – CNRS/INSU – Uni. Paris Diderot, HOBYS Key Programme Consortium.

Frontispiece figure: The Herschel Space Observatory. Cut-away diagram.
Frontispiece credit: © ESA, ATG medialab

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The idea of this book originated from the ExplorNova research project at CEA—Université Paris-Saclay, France, led by Vincent Minier, also engaged at that time as an astrophysicist in observing programmes with the *Herschel* space observatory. The ExplorNova project aims at studying innovation in space technologies from the history, the design and the realisation of space science missions, selecting the European Space Agency (ESA) *Herschel* Space Observatory for far-infrared and submillimetre space astronomy, because of its blends of scientific, technical and management challenges. *Herschel* was an ESA space observatory with science instruments provided by European-led Principal Investigators' consortia and with important participation from NASA. The mission was launched in 2009 and in-flight operations ended in 2013 with success, nearly 30 years after its initial proposal in 1982.

Following discussions started in 2011 between Vincent Minier and Roger-Maurice Bonnet (ISSI Executive Director and former Director of ESA Scientific Program[1]), it was considered of great appeal to conduct this study through ISSI, which could offer an ideal framework suited to both the objectives of the project and ISSI. ISSI is a centre of interdisciplinarity in space science, promoting, through the participation of international experts, a deeper understanding of past space science projects to assist in reaching out towards new horizons, benefiting from the interaction of different cultures and experiences in the development and management of complex instrumentation. An ISSI Working Group, one of the four tools of the Institute set up by the ISSI Directorate for specific tasks, often of a technical nature, was selected in October 2012 by the ISSI Science Committee, then chaired by Professor Johan Bleeker, who considered this Working Group formula as best adapted to the study. The outcome of ISSI Working Groups' activities is generally published in volumes of the ISSI Scientific Report Series (SR).

[1]Over the period 1983–2001, from the year of the Assessment Report of the original proposal to the year when the industrial implementation commenced.

The overall goal of the ExplorNova Working Group was to analyse the creation and evolution of a space mission from its original concept to its operation in orbit through innovative design, risk and crisis assessment and management and their effects on the final achievements of the mission and how it contributed to furthering knowledge, discovering new phenomena and offering new perspectives in astronomy. Its originality and challenge lie in the reflexive nature of the work, including direct actors of the *Herschel* story as the writers of their own collective history.

The project started its four years working period in October 2012. The main contributors to the writing and edition of the book include, besides Vincent Minier and his colleague Vincent Bontems, Philosopher of Technology at CEA, who both designed the conceptual methodology for this collective study, the *Herschel*'s Principal Investigators—T. de Graauw and F. Helmich his successor, M. Griffin who also reviewed the content and English language of all chapters and A. Poglitsch who contributed by means of private interviews; the former ESA Coordinator for Astronomy and Fundamental Physics between 1986 and 2008, Sergio Volonte; the *Herschel* Project Scientist—Göran Pilbratt; and the Director of the ESA Science Program–Roger Maurice Bonnet.

All contributed either directly or through interviews in the publication of the book throughout its different chapters. Four years have been used for investigating how technical inventions and adaptations covering the main critical systems of *Herschel*, including both the satellite and its scientific payload, emerged through a series of earlier developments and from other infrared astronomy missions, which preceded *Herschel*.

This work of an authentic interdisciplinarity nature opens new perspectives in transferring the methodology applied to *Herschel* to other domains of space activities either in other scientific missions or for operational or commercial satellites in Earth observation, communication and navigation systems.

This book is addressed to a large public with broad interests in science, engineering, technology research and development as well as in the philosophy of innovation, including scientists, and in particular space scientists, engineers, managers and decision-makers. It will offer a new vision and new perspectives to students motivated in embarking on a career covering any of these domains.

Bern, Switzerland
December 2016

V. Minier
R.-M. Bonnet
V. Bontems
T. de Graauw
M. Griffin
F. Helmich
G. Pilbratt
S. Volonte

Acknowledgements

All participants to the Working Group acknowledge the financial support of ISSI and its help in facilitating their work, their accommodations while in Bern and their travel for some of the co-authors. Vincent Minier and Vincent Bontems acknowledge support from the ExplorNova research project funded by Région Pays de la Loire, Université Paris Saclay and CNES, France.

This book would not have been published without the contributions of many persons at:

- ISSI: Prof Rafael Rodrigo, Prof. Rudolf von Steiger, Prof. Tilman Spohn who chaired the ISSI Science Committee when the project was approved, Andrea Fischer, Jennifer Fankhauser, Saliba Saliba, Silvia Wenger and the whole ISSI staff.
- ESA: John Credland, Thomas Passvogel who provided precious information on the management and the financial fights, Bert Bastijns, Anton Linssen, Eric Morel who helped in the back-dating of the cost-at-completion of both *Herschel* and *Planck* and Mélanie Legru and Nathalie Tinjod who provided access to numerous archived ESA documents from the Science Programme Committee (SPC), the Industrial Policy Committee (IPC) and the ESA Council (C), which are not available in the public domain and are listed in the list of references of Chap. 3. They can nevertheless be consulted on request to the ESA archive at ESA HQ, Paris.
- CEA: Louis Rodriguez, Marc Sauvage, Laurent Vigroux for their interviews about *Herschel* bolometers, Lionel Duband for his interview on cryocooler technology, Michel Spiro for his help on coopetition and Pierre-Olivier Lagage for his initial support to elaborate such a project at CEA Saclay.
- CNES: Michel Rouzé, project manager who supported initial work in investigating instrument innovation within *Herschel*.
- Observatoire de Paris: Pierre Encrenaz for his initial and constant support to the project and help in collecting archives and interviews and Jean-Michel Lamarre for his interview.

- Université de Nantes: Stéphane Tirard, historian of science at Centre François Viète, who supported the ExplorNova project.
- Astrium Toulouse (Airbus): Michel Bouffard and Yves Toulemont for their interviews.
- IRAP (formerly CESR): Emmanuel Caux for his interview.
- Cardiff University: Peter Ade and Ken King for their interviews.
- UQAM (Université du Québec à Montréal): Yves Gingras and Pauline Huet for their work on bibliometrics.
- Université Lyon Jean Moulin: Daniel Parrochia for his outstanding help in our philosophical approach.
- MPE, Garching: Reinhard Genzel and Albrecht Poglistsch for their interviews.
- Institut d'Astrophysique de Paris: Francis Bernardeau for the help of the IAP for hosting our Paris working meetings.
- Institut Mines Telecom: Armand Hatchuel for his advice on the C-K theory and Jean-Louis Ermine for his advice on Knowledge management.
- SRON: P. Huizinga, W.J. Vreeling, G. de Lange, P. Huizinga, W.J. Vreeling, G. de Lange and W. Laauwen for their help in getting pictures and illustrations related to the HIFI instrument.
- CALTECH/JPL: Thomas Phillips for his contribution through an interview.

Special thanks must be addressed to:

- Professor Johan Bleeker at SRON who contributed to the formulation of the ESA Horizon 2000 long-term plan in 1983–1984 as Chairman of the ESA Survey Committee and provided detailed records of the history of FIRST.

Thanks also to the Principal Investigators and their Co-investigators, the instrument hardware teams and their support, as well as all the scientists and observers who, through their work in the analysis, interpretation and publication of the data, have made it possible to exploit the very rich harvest of this pioneering Far-Infrared mission for the benefit of the whole World Astronomical community. Their number, totalling several thousands of scientists, makes it impossible to list them all here. The story of *Herschel* would also not have been possible without the outstanding contributions of the European Space Industry, in particular Thales Alenia Space and Airbus.

Contents

Chapter 1
Inventing a Space Machine: Breaking the Borders of Knowledge, Technology and Management

If the machine is one of the aids man has created towards achieving further intellectual growth and attaining maturity, if he treats this powerful automaton of his as a challenge to his own development, if the exact arts fostered by the machine have their own contribution to make to the mind, and are aids in the orderly crystallisation of experience, then these contributions are vital ones indeed
(Lewis Mumford, *Technics and Civilization*, 1934).

Abstract This introduction presents some of the characteristics of space missions and considers why they are stimulating for a study on innovation and why the *Herschel* Space Observatory offers a particularly rich example for engaging an epistemology of innovation. Our socio-technological study assesses the historical, structural and sociological aspects, which are discussed in the book. The first three chapters describe the conditions which framed the *Herschel* story from invention through development, in-orbit operation and scientific discoveries. The second part of the book analyses the key innovations on which the main elements of *Herschel* relied and the complex management structure adopted among all partners involved. Finally, we discuss the role of innovation and of risk management in view of the need to foster creativity and ambition for the next generations of instruments.

A space mission is a machine carrying mechanical, optical or electronic instruments placed on board an artificial satellite or an interplanetary probe and aimed at fulfilling any kind of measurement. Such measurements can be for astronomy, studies of matter and life in reduced gravitational fields, exploring the planets, observing and surveying the Earth and its environment as well as offering services in telecommunications, weather and climate forecasts, navigation, military and civilian security and so on. The key elements which constitute a space mission

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_1

are the launch vehicle (usually not reusable[1]), the satellite and its payload made up of the specific instruments that allow it to carry out its measurements, plus all the ground systems for remote control and command and receiving and processing the data transmitted to the ground.

A key parameter of a space satellite is its mass, which ranges from a few kg for cubesats—so-called because they are no larger than a 10 cm cube—to several tonnes for the big science missions. The primary objectives of these small cubesats were originally to train university students, but they are now also considered for several scientific research applications. The mass determines the size and power of the launcher and therefore the cost of the mission. A rough estimate of the cost of small- to medium-sized missions gives 1/3 to the launcher, 1/3 to the satellite and 1/3 to its payload. For large missions using the largest launchers available, the satellite and the payload costs are much larger. For the *Herschel* Space Observatory, the launch cost was roughly 10% of the total. This book focuses on missions of larger size that for the sake of simplicity are hereafter called space science missions.

1.1 Why Choose a Space Science Mission for a Study About Innovation

The term space science encompasses all branches of astronomy, Solar System exploration (including the Sun, the planets, comets and asteroids), the interplanetary medium and the plasma physics phenomena which are present in the solar wind and the surroundings of magnetic bodies and their so-called magnetospheres. The main objectives of space science missions are to conduct observations above the Earth's atmosphere or in deep space. The rationale for developing such expensive and ambitious endeavours is mainly scientific, justified by curiosity and the quest for new discoveries. Space missions are needed to explore planets with orbiters or probes that land on their surfaces and to observe astronomical objects and phenomena at wavelengths that are absorbed by the atmosphere: gamma rays, X-rays and ultraviolet and far-infrared (FIR) radiation. Examples are the *Hubble* Space Telescope (HST), the Mars rover Curiosity, Rosetta and its Philae lander and of course the *Herschel* Space Observatory, the mission chosen for this book and hereinafter simply named *Herschel*. Because space missions explore the unknown, be it by studying planets, comets, the Sun or stars and galaxies, or probing yet unobserved parts of the electromagnetic spectrum, and operate far away from us, they are often perceived as mythic machines venturing into the invisible cosmos, delivering beautiful images and other scientific data on the extraordinary objects which populate this immensity. They are visionary, enabling their inventors to explore the Universe while exciting the curiosity of many people.

[1]With the exception of the NASA Space Shuttle and probably in the near future, some parts of the rocket such as the first stage if proven economical.

Since their goals are to produce new knowledge, space science missions are very seldom repeated with the same capabilities: each must provide a substantial improvement with respect to previous missions working in the same scientific area. For this reason, they are technically ambitious, presenting big challenges, with performance requirements sometimes several orders of magnitude better than their predecessor missions, and are often identified as 'mission impossible'. Since the launch of Sputnik 1 in 1957, they have rapidly evolved from simple concepts to complex sets of highly sophisticated instruments. Some examples are the long series of international Solar System and heliospheric missions; the European Space Agency (ESA) XMM-*Newton* High-Throughput X-ray mission and its high-angular resolution NASA companion, *Chandra*; the NASA/ESA *Hubble* Space Telescope (HST) and its successor the *James Web* Space Telescope (JWST); and of course *Herschel*. These and similar missions are confronted to the necessity of operating in the harsh environment of space characterised by extreme conditions of vacuum, pressure, temperature and the intense bombardment of cosmic rays and high-energy particles from the solar wind and coronal mass ejections that pervade the interplanetary medium. Satellites experience intense vibration and acceleration during launch and before becoming weightless in orbit. Unlike laboratory experiments, ground-based observatories, or major scientific facilities like CERN, space mission hardware cannot be repaired or upgraded once in orbit.[2] To ensure that they can survive the rigours of launch and the space environment and perform reliably over their complete operating lifetimes (typically several years or more), they must be designed with very high standards of robustness and reliability and exhaustively tested before launch. To achieve major scientific breakthroughs, they must use the most advanced technologies, constantly introducing new concepts, breaking new barriers and making them the most modern instruments of their kind. At the same time, they rely on high-level engineering and demonstrably reliable technologies. These two aspects—novelty and proven reliability—are often in tension with each other, requiring very careful design choices and testing. The lengthy and ambitious developments and tests often confront the participating individuals, institutes and funding agencies with very difficult dilemmas resulting from the rapid evolution of the technologies. They could either decide to rest on rather old but proven technological approaches or to jump on a newly discovered solution offering better performance, but at higher risk. To avoid single-point failures, which could abruptly and unexpectedly put an end to their function, space missions also normally incorporate redundancy in the design of their vital subsystems (with the exception of the launcher, for obvious reasons), allowing them to secure their objectives and increase their longevity.

As they become more complex and sophisticated, and more massive, space science missions have required more powerful launchers. Launch mass and payload

[2]With the unique exception of the HST, which was repaired and maintained in orbit by astronauts carried by the space Shuttle offering a regularly upgraded facility to the astronomers since its launch.

complexity are major cost drivers, and the most ambitious missions need large budgets often exceeding hundreds of millions or even billions of Euros, cost levels which are typically an order higher than needed for a similar machine designed to operate from the ground. These high costs mean that space projects must respect strict budget and schedule controls. Maintaining cost and schedule control can be a very difficult task, given the large number of scientists and engineers, and of institutes and countries, involved in the development. The relatively high cost of space science missions also dictates that they must be designed to carry out only scientific investigations that are completely impossible from the Earth.

Roles of Space Scientists

A principal investigator (PI), supported by several co-investigators (co-Is), usually leads the development and the management of an instrument on board a space science mission. These people must combine scientific expertise with talents in engineering and management. In addition, for the exploitation of the scientific results, the PI calls on a number of scientific experts in different domains of science related to the main objectives of the mission.

In the ESA system, the agency is responsible for all the elements of the mission with the exception of the scientific payload, which falls under the responsibility of the PIs and co-Is, whose financial resources are provided by their respective member states and not by ESA. The responsibilities of the PI are enormous, including leading the scientific and technical work of the instrument consortium and reporting both to national agencies for financial management and to ESA for the development of the payload and its interfaces with the spacecraft and for schedule management.

Over the various phases of their development, from the first ideas through the first concepts, and early design, their technical and industrial development, their operations in orbit and their scientific exploitation, space missions connect the scientific, engineering, political and public arenas. They bring together international and national agencies, research organisations, universities, industries and society at large, and last but certainly not least, they involve a complex network of person-to-person contacts and interactions. Involving teams composed of hundreds of scientists and engineers from scientific institutes and industry, in a large number of different countries, they require innovative management approaches. Although sometimes difficult and daunting, these managerial challenges are well worth meeting because these large-scale international partnerships enable the participation of the best technical and scientific experts in the world, as well as the assembly of the necessary financial resources for the development of the best quality scientific payload and ultimately the highest scientific impact.

For these reasons space science missions offer very broad and instructive opportunities to address the general process of the dynamics and typology of innovation and the conditions that made it possible, one essential aspect of what this book is all about. Their peculiar characteristics, in particular their long-term design and implementation, extending over decades for the largest, represent a very rich source of data for reconstructing their technical heritage and the succession of steps which characterise their development. It can start with an invention (a radical innovation), improve the functioning of an instrument (incremental innovation) or simply adapt a well-known object to a new environment (customisation), possibly leading to saturation of the performance and the substitution by a more competitive technology.

1.2 Why *Herschel*?

The *Herschel* Space Observatory, fully dedicated to far-infrared (FIR) and submillimetre-wavelength astronomy, is the first of its kind. Even though infrared radiation (IR) was discovered at the beginning of the nineteenth century, and X-ray radiation a century later, the first X-ray satellite was launched in 1974 while the first IR mission, IRAS, launched only in 1983. The number of X-ray missions launched so far totals more than 25, but the corresponding number of missions observing in the IR and FIR part of the spectrum is less than half that number. One of the reasons for this is that the technology needed to detect (IR) radiation is extremely challenging. As will be described in the following chapters, *Herschel* is indeed one of the most complex and ambitious astronomy missions yet launched. It was developed and operated by ESA, with several important international participations from Europe and elsewhere. Such ambitious and difficult endeavours will never proceed without problems. Our book analyses the occurrence and the resolution of the many crises, be they technical, financial or political, which inevitably appeared during the design and the development of an undisputable space science success story. How science, technology and management interacted and triggered innovation within *Herschel* and whether there are any general properties in the innovation processes that, once identified, could help in designing future space instruments are the main themes of this book.

Herschel indeed offers a unique opportunity to investigate such ideas. It consists of many technical systems, each of them having its own heritage. It is used here as a window through which one can follow technical lineages of interest, backward in time, and how they paved the way to new technological developments, identifying the starting point of innovation and the contextual reasons for its diffusion, saturation or substitution. For instance, the superfluid liquid helium cryostat (whose first of its kind was flown on IRAS in 1983) to cool the temperature of the instruments and detectors down to 1.7 K is based on the one developed for ESA's Infrared Space Observatory (ISO). Therefore, ISO (launched in 1995) and *Herschel* (launched 14 years later) provide two data points on the technical lineage curve of the cryostat. The bolometer cameras use different designs whose origins date back to the 1970s and 1980s, respectively, allowing the study of two competitive lineages in bolometer technologies. Our study also includes some failed and abandoned technologies, identified during interviews with experts and why they were dismissed. Deployable large space mirrors offer another good example, and a special chapter is devoted to the tortuous but fascinating story of the *Herschel* telescope primary mirror, the largest yet developed and launched. It also allows investigation of the synergies between industrial and military research and science laboratories in different areas, in particular the design of photodetecting systems. Finally, *Herschel* carried the most ambitious heterodyne astronomy instrument yet operated in space, and a special chapter describes the rocky road followed by the heterodyne techniques, which started in the wake of the radar developments during World War II, taking advantage of that heritage that eventually led, through a series of laboratory and ground-based experiments, to the development of the *Herschel* Heterodyne Instrument for the Far Infrared (HIFI).

Herschel also offers interesting lessons in the area of complex project management. As one out of the four cornerstones of the ESA Horizon 2000 programme extending over 20 years with a fixed budget, *Herschel* and its sister mission, *Planck*, were subjected to a very strict financial management rule of design to cost. That stringent constraint not only imposed rigorous financial control on all partners but also triggered technical innovation to identify the most high-performing, and at the same time the most economic, solutions to the development of scientific instruments and spacecraft systems and subsystems. In addition, scientific and technological novelty, risk handling, crisis management, international cooperation and leadership culture all offer extra dimensions to this study and an opportunity for improving the development of future and more ambitious science missions facing long developments at a time when international cooperation is an absolute must. Last but not least, the development and ultimate success of *Herschel* casts an interesting light on the paradigm of the Technology Readiness Level (TRL) risk management approach, which has since become systematically implemented by several of the large space agencies in the world. Divergence from this approach when the mission started probably explains why major innovations were produced during the development of *Herschel*'s science payload.

1.3 Our Approach

Based on the *Herschel* mission, the development and scientific output of which is an indisputable success story, this book leads us to propose a methodology for analysing the innovation process at various milestones of a space project and to deduce a typology of innovation for all large technical systems. This methodology introduces a novel way of capturing and storing information for the purpose of knowledge management. Because knowledge gets often lost in the history of mission development, finding a way to encapsulate the expertise of technicians, engineers, scientists and managers and to share it with the broad science and space community using a knowledge database and extracting prospective information for preparing future R&D activities were among our major goals in this study. It is regrettable that lessons learned from past endeavours about how to do things well, and how to avoid doing them badly, are often ignored by teams responsible for the management and development of new satellites for different reasons, such as the reluctance to revisit out-of-fashion approaches or just because of lack of time. With future major observatories now being planned and implemented, it will only prove costly in financial and scientific terms if the valuable experience of the *Herschel* mission is neglected.

Our socio-technological study rested on three main types of analyses that form the basis of the epistemology of innovation:

- Historical—covering political, social, management and scientific analysis and placing innovation in the historical context of *Herschel*, spanning three decades, through both a series of interviews of key participants in the mission development from the scientific community, industry and agencies and the consulting of archives associated to the project
- Structural—covering technological system analyses, identifying the technical lineages which benefited the development of *Herschel*, assessing their level of maturity and improvement with respect to previous technical devices and describing their saturation and how they were substituted by a new technical lineage in the case of invention
- Sociological—through interviews, study of transversality and recourse to bibliometrics, enabling identification of these researchers shifting across traditional scientific or intellectual boundaries, as well as the methodologies they follow, fostering human motivation to organise and manage new conditions leading to innovation

The book is divided into ten chapters following this introduction, organised in two main parts. The first following three chapters describe the history of the mission development and the scientific, sociological and political context which allowed *Herschel* to be developed, launched and operated. Chapter 2 sets the scene: the birth of IR and submillimetre astronomy in Europe and the programmatic framework of ESA Horizon 2000 long-term plan and its design-to-cost philosophy. Chapter 3 describes the hectic timeline, following the development of *Herschel* from the early design phase through launch and in-orbit operations. Chapter 4 resumes the scientific motivation behind *Herschel* and its evolution over the development of the spacecraft in response to scientific and technical developments during that long period and offers a broad view of the main scientific results, the discoveries and new knowledge that *Herschel* has produced. These three chapters synthesise the main history, conditions and limits that characterise the *Herschel* mission.

The second part of the book is mostly concerned with the key innovations on which *Herschel* relied. Chapter 5 outlines key philosophical and sociological concepts to understand innovation, such as the positive effect of design to cost, creativity, coopetition and the role of societal and human relations in different parts of the world, and it introduces our epistemology of innovation. Chapters 6–9 focus on technological innovations within the four most critical elements of *Herschel*: its telescope, the bolometer cameras, the heterodyne instrument and the cryostat.

Chapter 10 presents the management of *Herschel* at both the ESA level and inside the three scientific teams and instrument consortia.

Chapter 11 concludes on the characteristics of a successful mission and discusses knowledge management (KM) of space innovations. Guidelines for ensuring that knowledge is transferred, capitalised and disseminated are also proposed. A discussion about the future evolution of space science addresses key questions about the role of human innovation and of experimentalists, in securing a bright prospect for creativity and ambition in the development of the next generations of space and ground-based astronomy instruments.

Chapter 2
Creating the Historical and Strategic Framework for *Herschel*

Abstract Here we analyse the historical context which permitted the successful development of *Herschel*, from the discovery of IR radiation through the development of detection techniques, in particular bolometric detectors. We demonstrate how the European space astronomy community was able to reach maturity, how influential was the role of the IRAS mission and how infrared astronomy was able to slowly find its way to the European Space Research Organisation (ESRO). We emphasise the crucial role of the ESA Horizon 2000 long-term plan in allowing the development of FIRST (subsequently renamed *Herschel*) as early as 1985. The long-term plan and vision constituted a framework driven by a commitment to achieving scientific excellence, and by schedule and financial respect, in adherence to the original figures. Despite serious difficulties met by the ESA scientific programme during its development, due to budgetary constraints and launchers accidents affecting several missions of the plan, *Herschel* emerged as a major scientific success.

2.1 The Birth of Infrared and Submillimetre Astronomy

When in 1800 William Herschel (Fig. 2.1) placed a thermometer in sunlight after it passed through a prism, and discovered what he then called 'calorific rays', he certainly had no clue that more than two centuries later a telescope, named in his honour, would be orbiting the L2 point in the Earth/Sun system, a point in space named in honour of another famous European mathematician and astronomer, Joseph Louis Lagrange, who some 28 years before *Herschel's* discovery studied the equilibrium of small objects under the gravitational attraction of larger ones, in that case the Moon and the Sun. These two scientists illustrate the incredible wealth of European scientific and rationalistic culture at the turn of the nineteenth century and its continuing relevance to scientific research today.

William Herschel was able to show that infrared radiation, just like visible light, could be reflected, transmitted and absorbed. Figure 2.2, which illustrates the various parts of the electromagnetic spectrum expressed in wavelengths, shows the position of infrared light in the spectrum. The topic of this book concerns the *Herschel* Space Observatory, essentially covering the far-infrared (FIR) and submillimetre (submm) waves (Table 2.1), which are absorbed by the atmosphere

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_2

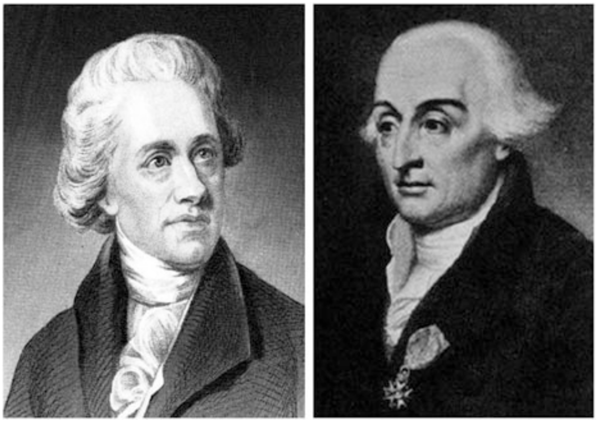


Fig. 2.1 Two great European scientists: William Herschel (*left*), astronomer and music composer, was born in Germany in 1738 and emigrated to Great Britain in 1757 where he died in 1822 after being elected the first President of the Royal Astronomical Society. © Skycaramba.com; Joseph Louis Lagrange (*right*) was born in Italy in 1736 where he founded the Science Academy of Torino and emigrated to Germany in 1766 and then to France in 1787 where he died in 1813. © Engraving by Robert Hart. British Museum

of the Earth. This part of the spectrum remained largely unexplored until the scientific curiosity of both radio and infrared astronomers, combined with new developments in technology, began to reveal a wealth of new science from observing at these wavelengths. This chapter has no ambition to offer the reader a comprehensive and detailed history of the development of this branch of astronomy. We refer readers to the excellent reviews of Lequeux (2009), Harwit (2001) and Rieke (2009), which offer many interesting details of importance to the present book.

Sensitive detection in the infrared requires detectors, and preferably instruments and telescopes, to be operated at very cold temperatures. This enhances the sensitivity of the detection system and also minimises the undesirable background of black body[1] emission from the observing system itself. The development of convenient cryogenic systems as well as novel types of detectors has proven indispensable for all scientific utilizations of IR, FIR and submm radiation. Therefore, technological and instrumental research has been a major driver of progress in all these applications. Pioneering observations were made, starting in the 1830s continuing through the nineteenth century, using thermocouples (which convert temperature differences into electric signals) to detect infrared radiation from the Sun, the Moon and the bright stars. These first steps were slowly followed by

[1]A black body is an idealised physical object that absorbs all incident electromagnetic radiation, regardless of frequency or angle of incidence. In thermal equilibrium, a black body radiates according to *Planck's* law, with a particular dependence on wavelength and a total power output that only depends on the temperature of the object.

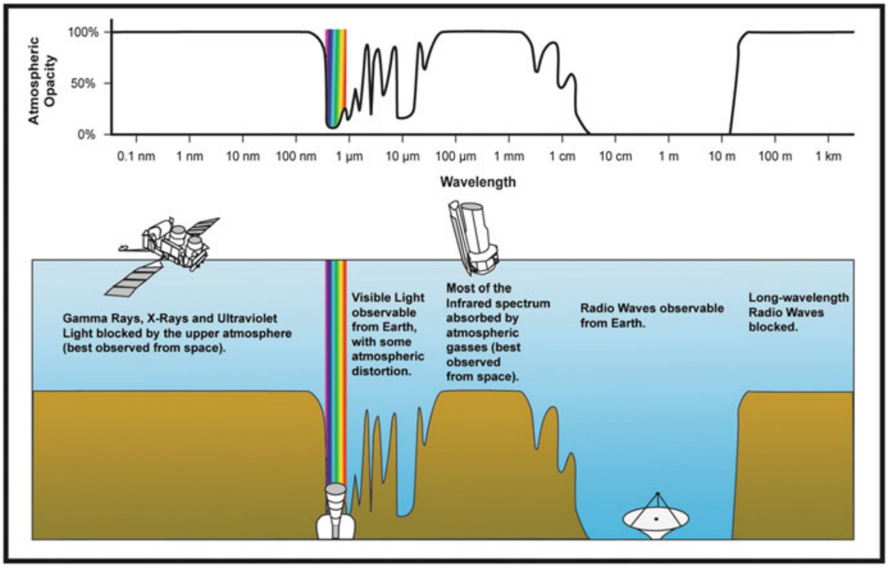


Fig. 2.2 The electromagnetic spectrum from very high-energy gamma rays on the left to radio waves on the right. Strong atmospheric absorption occurs in the high-energy range (left of the visible spectrum) and in the infrared, far-infrared, submillimetre and millimetre wavelength ranges, which can be overcome through the use of instruments carried onboard high-altitude balloons, rockets and artificial satellites. © IPAC Caltech

Table 2.1 Identification of the subdivisions in the infrared and radio ranges of the electromagnetic spectrum

Designation	Abbreviation	Wavelength	Frequency
Near-infrared	NIR	0.7–5 μm	430–60 THz
Mid-infrared	MIR	5–30 μm	60–10 THz
Far-infrared	FIR	30–300 μm	10–1 THz
Submillimetre	submm	300 μm–1 mm	1000–300 GHz
Millimetre	Mm or mm	1–10 mm	300–30 GHz
Microwave	Microwave	10 mm–1m	30 GHz–300 MHz
Radio	Radio	1 m–100,000 km	300 MHz–3 Hz

incremental advances in the early twentieth century until breakthroughs occurred in the development of new and far more sensitive infrared detectors and image converters mostly carried out by the military (e.g. Rogalski 2012). Absorption by atmospheric molecules, in particular water and carbon dioxide, requires the use of high-altitude observing sites or extra-atmospheric observations carried from aircraft, balloons, rockets or better still artificial satellites, which had to wait the advent of the space age 157 years after William Herschel's discovery.

2.1.1 The Military Connection

The military, particularly in the USA through the Department of Defense, made major advances in infrared technology. The image converter[2] developed for infrared radiation on the eve of World War II enabled imaging in total darkness, detecting shell batteries, rockets, missiles and their exhaust plumes, as well as the body of enemies at night. The development of lead sulphide (PbS) and indium antimonide (InSb) semiconductors (responding to radiation of about 3 μm) started in the 1930s in Germany, and such detectors were already used in airborne IR systems as of 1943, while infrared photoelectric detectors were developed during World War II in the USA. However, these technologies remained largely classified and so inaccessible to civilian users. A major breakthrough occurred in 1961 with the development of the germanium bolometer,[3] which achieved a huge increase in sensitivity (a factor of several hundred) over most of the FIR-submm part of the spectrum. The first cooled telescopes were flown on rockets and balloons, allowing several minutes to several hours of observations above most of the atmosphere with a factor 100 more sensitivity. The first all-sky surveys of celestial sources at 4, 10 and 20 μm were made by the US Air Force Cambridge Research Laboratory Hi-STAR project with the main objective of cataloguing astronomical infrared sources to avoid confusing rocket engines with stars, while at the same time creating a strong stimulus for future astronomy missions. Unfortunately, the submm and FIR spectral ranges were of little or no interest to the military because they did not correspond to light emissions from targets of importance to them. Consequently, funding for the development of FIR-submm detection techniques had to come from the more modest science budgets available to astronomers (Jamieson 1995). Even though the US community did benefit from military research after World War II, infrared astronomy started rather late in Europe. Furthermore, the need to employ cryogenic techniques, difficult and inconvenient to implement outside specialised laboratories, hampered the development of working astronomical instruments.

The development of radio astronomy also took advantage of the military heritage of radar developments during World War II, in particular in the USA (MIT), and from captured German radar reflectors placed in the UK, France, the Netherlands, Sweden and Czechoslovakia, offering future radio astronomers hands-on experience before other major dedicated facilities became in operation, in Cambridge, Dwingeloo, Greenbank, Nançay, Onsala and elsewhere.

[2]Also known as the infrared image tube, the image converter converts an invisible infrared image into a visible one. It consists of an infrared-sensitive, semitransparent photocathode on one end of an evacuated envelope and a phosphor screen on the other, with an electrostatic lens system between the two.

[3]The bolometer, invented in 1878 by Samuel P. Langley, was initially essentially a very sensitive thermometer made of a thin, blackened strip of platinum metal. Heat radiation falling on the strip changed its electrical resistance, which was recorded by a sensitive electric metre. Subsequently, this basic idea was developed into extremely sensitive bolometric detectors (see Chap. 7).

2.1.2 The Emerging Infrared Astronomy Community

Since they were discovered, infrared and radio celestial emissions have challenged scientists' curiosity. Atoms and molecules strongly emit or absorb radiation in the infrared to microwave range due to numerous vibrational and rotational quantum transitions with energy changes requiring emission or absorption of photons. Cold, dark clouds of molecular gas and interstellar dust glow with radiated heat as embedded stars irradiate them. Infrared radiation can also be used to detect embryonic star (also called protostars) before they begin to emit visible light.

Early attempts in the first quarter of the twentieth century at finding infrared stars were hindered by the fact that the work relied on narrow-field optical telescopes and photographic detection; the relatively short wavelengths to which photographic emulsions were sensitive (below 1 μm) further limited the work. While both young and very evolved stars can be prominent in the infrared, most fully formed stars are brightest in the visible region and are not strong infrared emitters, allowing surrounding cooler objects such as dust disks and planets to be more easily detected. Infrared light is also useful for observing the cores of active galaxies, which are often cloaked in gas and dust. Galaxies emit mainly in the visible due to the large number of stars, and for very distant galaxies, which are receding from us at high speed, the visible light is shifted by the Doppler effect towards longer wavelengths, making them more readily observed in the infrared. The primordial light emitted by the Universe after the Big Bang can only be detected in the mm to radio wavelength range.

Because of the technical challenges inherent to this spectral range, in particular in detector and cryogenic technologies and the poor transmission of the Earth's atmosphere, most observational astronomers tended to neglect the field, which was initially explored by engineers and physicists. As noted above, a breakthrough came with the development of PbS and InSb photon detectors and image converters in the first half of the twentieth century, at Caltech in particular, where the Mount Wilson 2-μm Sky Survey covered 75% of the sky. The success of military developments and the advent of radio astronomy prompted more astronomers in the 1950s–1960s to take notice (Lequeux 2009), and infrared astronomy gradually became established as an extension of visible and radio astronomy (see text box on the radio astronomy connection). These developments also stimulated exploration of the longer wavelength parts of the spectrum in the FIR, submm and mm range (Table 2.1). Furthermore, the growing number of discoveries, and the evaluation of new high-altitude observing sites partly overcoming the limits of atmospheric absorption, prompted astronomers to probe the cold Universe through observing at longer wavelengths. As well as new instruments for detection of black body emission of interstellar dust, heterodyne instruments, traditionally used at radio wavelengths, were extended downwards in wavelength to the submillimetre to enable astronomers to probe molecular spectral lines (Chap. 8). These developments opened a new field in experimental astronomy, blending pure astronomical interests with fast-paced instrumental and experimental research.

The mm-Radio Astronomy Connection

The first detections in the late 1960s of interstellar ammonia, water and hydroxyl in the radio range, followed by the discovery in 1970 of interstellar carbon monoxide at mm wavelengths, inspired the rapid development of ground-based facilities. The NRAO 12 m telescope, which started operating in the 1960s, became the facility where most interstellar molecules were first detected, in addition to the 5 m telescope located at the McDonald Observatory and the first antenna of the Owens Valley Radio Observatory mm interferometer (OVRO), operating at 350 GHz, in 1978. More radio telescopes were built on high-altitude sites and dedicated to submillimetre astronomy. The Institut de Radio Astronomie Millimétrique (IRAM), created in 1979 in Grenoble, started the construction of two large facilities: the NOEMA interferometer (currently an array of seven 15 m telescopes) on Le Plateau de Bure in the French Alps and the 30-m dish in Pico Veleta in Spain. Two submm telescopes were built at 4000 m altitude, on the summit of Mauna Kea in Hawaii: the James Clerk Maxwell Telescope (JCMT), which started operating in 1987, and the Caltech Submillimeter Observatory (CSO), which saw first light in 1986. In 1987, the European Southern Observatories (ESO) joined with Sweden to build the Swedish ESO Submillimetre Telescope (SEST), a 15 m IRAM-type antenna, on la Silla in Chile at 2600 m altitude. SEST was decommissioned in 2003 to transfer its staff to APEX, the 12 m Atacama Large Millimeter Array (ALMA) Pathfinder Experiment which, like ALMA, is located on the Chajnantor plateau at 5100 m altitude. There were numerous smaller projects like the Bordeaux interferometer and the Gornergrat 4 m antenna (KOSMA), which offered excellent learning and training opportunities for the growing generation of mm and submm range astronomers and instrument developers.

The first mm-wavelength radio observations were carried out by Soviets in 1959 using a 22-m dish in Pushchino operating at 7 mm. The Americans made their first mm-wavelength observations in the 1960s with the McDonald Observatory of the University of Texas, and the Kitt Peak telescope of the US National Radio Astronomical Observatory (NRAO), also known as the '36 foot telescope', where IR and mm astronomy came of age with the pioneering work of *Frank Low* (see text box below). In Europe, mm observations started a decade later. In Bordeaux around 1970, a small two-element interferometer (2.5-m antenna diameter with 80-m baseline) was constructed for solar observations and for mm interferometry experiments (Delannoy et al. 1973). In 1975 the first large mm antenna of 20 m diameter came into operation at the Swedish radio observatory in Onsala. At the end of the 1970s, IRAM, a French, German and Spanish observatory, was conceived, and its

6-antenna interferometer started operation by 1988. Its 30-m single dish began operations in 1989.

Frank Low

Formerly employed by Texas Instruments Laboratories where he had developed the gallium (Ga)-doped germanium (Ge) bolometer, Frank Low joined the NRAO in 1962 at the invitation of the Director, Frank Drake. His bolometer, which was 1000 times more sensitive than what had been available up to then at 10 μm wavelength, with a response extending out to 100 μm, became the workhorse for infrared and millimetre continuum astronomy for a decade. Low also developed a novel liquid helium cryostat which he started to market through his own company, Infrared Labs, in Tucson (Az) in 1967, together with Ge bolometers and InSb detectors, systems that were used by many astronomers worldwide. While at Kitt Peak, Low designed a 30 cm telescope for the Learjet aircraft and used it to make the first FIR observations of the galactic centre. To make these measurements possible, he invented an observing mode in which the telescope secondary mirror was rapidly moved back and forth (chopped) so as to shift the detector field of view between the source and an adjacent patch of sky. This allowed the large background from the atmosphere and telescope to be subtracted off, providing an accurate measurement of the astronomical source alone. This powerful technique subsequently became a crucial feature of IR and submm observatories.

Low's work was key to the development of submillimetre astronomy in the USA and Europe. He can be considered as the founding father of the IRAS satellite (Fig. 2.4), the first infrared satellite ever developed, which mapped the whole sky producing an archive of astronomical data that is still in use today and which provided an essential input to *Herschel*'s observations.

2.1.3 The Space Connection

As shown on Fig. 2.2, the Earth's atmosphere is not transparent to cosmic electromagnetic radiation throughout a large portion of the electromagnetic spectrum. Water vapour, for instance, strongly absorbs IR to mm radiation, except in some low-transparency windows (Fig. 2.3). Moving observatory facilities to high-altitude and dry sites, such as Mauna Kea in Hawaii and the Chajnantor plateau in Chile, had been a key requirement for developing submillimetre astronomy. Besides favourable high-altitude ground-based locations, aircraft, rockets, balloons and satellites offered powerful possibilities that enhanced the rapid development and success of modern infrared and submillimetre astronomy in the second half of the twentieth century (Harwit 2001). The tight military connection, together with the technological challenges inherent to this new branch of astronomy, explains why European researchers benefited from cooperative ventures with the USA, as well as

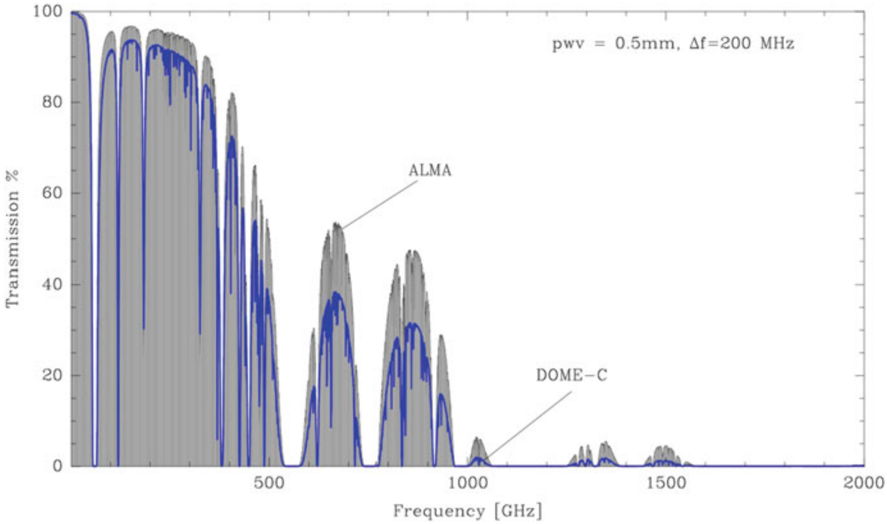


Fig. 2.3 Modelled atmospheric transmission for Dome C in Antarctica and Chajnantor/ALMA site in Chile for a precipitable water vapour (PWV) content of 0.5 mm. The spectral frequency resolution is 200 MHz. Grey plot represents atmospheric transmission at ALMA site while blue plot is atmospheric transmission at Dome C site. Many windows in the range 300 μm (1000 GHz) to 1 mm (300 GHz) appear with good transmission. The 200 μm (1500 GHz) window slightly opens at ALMA site and Dome C, but not sufficiently to ensure regular observations as in space. From Minier et al. (2008)

from individual European nations' own initiatives. Today, space is considered the ideal location for exploring these spectral domains, and it is therefore logical that infrared astronomy has gradually shifted from a purely ground-based activity to a branch of science accessible mainly to space-faring nations among which the USA, Europe, Japan and to a lesser extent Russia are leaders.

Cooling a telescope on the ground is impractical because ice would condense on its surface. The first cooled telescopes were small ones placed on rockets, allowing them to observe the sky for several minutes before re-entry, and the first infrared all-sky map resulted from a long series of flights conducted by the Air Force Cambridge Research Laboratory (Low et al. 1976). These projects surveyed the cosmos at wavelengths of 4.2, 11.0, 19.8 and 27.4 μm. Although the total accumulated observation time of these flights was only about 30 min, they successfully detected 2363 reliable infrared sources, which were published in the Air Force Geological Laboratory Four Colour Infrared Sky Surveys Catalogue (Price and Walker 1976) and (Price 1977). About 70% of these sources matched those of the Mount Wilson 2.2 micron survey.[4]

Airborne astronomy also began in the USA in 1966, when planetary scientist Gerard Kuiper flew a 30 cm telescope pointing out from the window of a Convair-990. Once it had been proven that airborne astronomy was feasible (see Frank Low text box), scientists were keen to use that technique routinely. From 1974 to 1995,

[4]http://coolcosmos.ipac.caltech.edu

NASA operated the Kuiper Airborne Observatory (KAO), named after Gerald Kuiper, which carried a 91 cm reflecting Cassegrain telescope in a converted C-141 military cargo plane. The KAO covered a range of 11,000 km, reaching an altitude of 14 km, and was designed for observations in the 1–500 μm spectral range. In Europe, particularly in the Netherlands, the UK and Germany, but also in France, astronomers used these NASA facilities. In 1969, having returned to France from the USA where he had conducted solar infrared observations with the Learjet, Pierre Lena at Meudon Observatory initiated a series of observations using the Caravelle-116 aircraft carrying the 32 cm OSIRIS telescope, which operated in the range 30–200 μm (Rouhan 2013; Vanhabost et al. 1977). OSIRIS also flew on the NASA/Ames Research Center Convair-990 in the course of two ASSESS missions[5] aimed at simulating the operations of Spacelab (Sect. 2.2.3). Since 1996, NASA and the German Aerospace Centre (DLR) have been developing and operating the Stratospheric Observatory for Infrared Astronomy (SOFIA), onboard a modified Boeing 747-SP aircraft that can allow its 2.5 m diameter telescope to conduct observations at altitudes of about 12 km. Featuring three times better image quality and over an order of magnitude higher sensitivity than the KAO, SOFIA is today the world's largest mobile astronomical observatory. It saw first light in 2010, a year after the launch of *Herschel*.

In parallel, balloons equipped with instruments for the FIR and submm observations have offered affordable opportunities to new generation of astronomers. Helium-filled, mylar balloons have carried infrared telescopes up to altitudes as high as 40 km, which the NASA Goddard Spaceflight Centre used in 1996 to survey the sky at 100 μm. Their programme led to the discovery of about 120 bright infrared sources near the plane of our galaxy. In Europe, in the mid-1970s, the AGLAE twin 14 cm telescopes (Serra et al. 1978) and the Dutch BIRAP platform with a 60 cm telescope, under the leadership of R. van Duinen (Harwit 2001), initiated a highly ambitious and successful balloon programme. Later, the German 1 m 'Golden Dragon' telescope (Drapatz et al. 1981), the UCL balloon platform (Poulter and Jennings 1983), the PIROG (Stegner et al. 1997) and the PILOT project (Bernard et al. 2016; Enger et al. 2009) followed the path and all helped to uncover the FIR universe. These projects were later followed up in France with the 2 m PRONAOS telescope (Lamarre et al. 1990) whose 540 μm to 1.1 mm photometric band allowed measurement of the Sunyaev-Zeldovich effect (Lamarre et al. 1998), while three other bands were used for characterising the transmission of cold dust in the galaxy. Even though its planned heterodyne receiver was not functional at the time, PRONAOS played an important role in the development of heterodyne technology for space facilities like band-1 in HIFI (Chap. 8). BLAST

[5]ASSESS (Wegmann et al. 1978), a joint exercise of NASA and ESA, was set up to enable scientists to become accustomed to an atmosphere-free, gravity-free working environment. This simulation of Spacelab conditions provided new data and a valuable insight into the possibilities and limitations of such a mission.

and BOOMERanG experiments also provided valuable scientific and technical information for FIR and submm space observatories.

A historical breakthrough in infrared astronomy occurred in 1977 when an international collaboration was formed by the Netherlands, the USA and Great Britain to develop the Infrared Astronomical Satellite, IRAS (Fig. 2.4), which pioneered satellite-based infrared astronomy not only in Europe but also in the world, opening a scientifically rich and productive period for a new branch of space

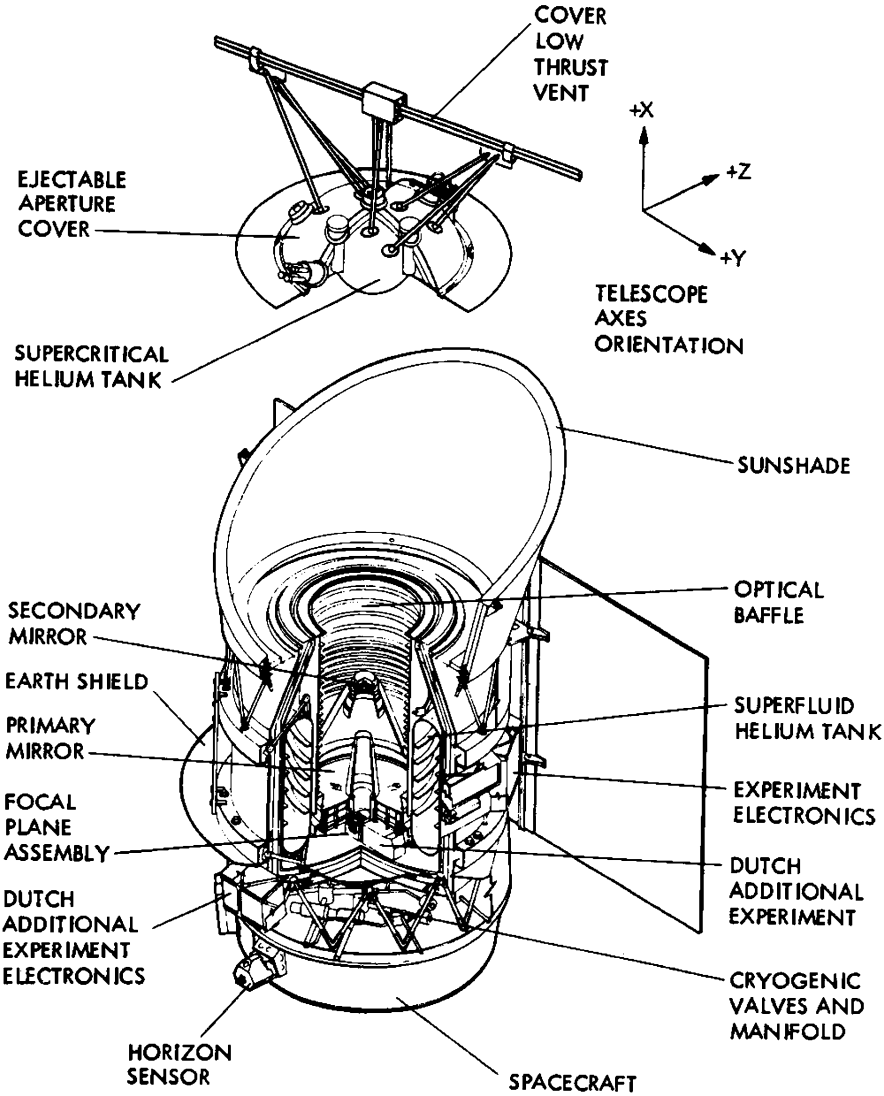


Fig. 2.4 Schematic of the NL-USA-UK Infrared Astronomy Satellite (IRAS). © NASA RP-1190

astronomy. The Americans built the telescope, the detectors and the liquid helium cryostat needed to cool the detectors and telescope; the British provided the satellite ground station and control centre; and the Dutch built the spacecraft, which included the onboard computers and the pointing system. IRAS was launched in 1983 (26 years after Sputnik 1) and operated for 9 months before its helium was exhausted. As the first infrared space mission, using the first ever space-borne helium cryostat, conducting the first all-sky infrared survey finding over 350,000 sources, IRAS paved the way for all subsequent infrared space missions.

Bearing in mind its ambitious design, its international connections and its cost, the IRAS project could have been carried out under the responsibility of the European Space Research Organisation (ESRO). But (ESRO), created in 1964, 7 years after the launch of Sputnik 1, was surprisingly ill prepared for developing such a complex project. It was only much later that infrared astronomy was introduced in the programmes of the European Space Agency (ESA), ESRO's successor, with the approval in 1983—the year of the IRAS launch—of the Infrared Space Observatory (ISO). That late entry of ESA into IR and submm infrared astronomy is analysed in the following section. It culminated with the development of *Herschel*, the largest space telescope yet launched at the time of writing.

Table 2.2 lists the 13 known infrared, FIR, submm and cosmology satellites launched so far by all space-faring nations and organisations since the beginning of the space age in 1957. As impressive as it appears, this number is nevertheless only about one third of the total number of high-energy and UV satellites launched to date. That difference can be attributed to the technological difficulties, and consequent costs, inherent to that branch of space astronomy and to the availability of ground-based observatories and balloon-borne experiments, which attracted the scientists' preference, precisely because of these development difficulties and the inherent high costs and risks of space systems.

2.2 The Birth of Space Infrared and Submillimetre Astronomy in Europe

2.2.1 Creation of a European Space Research Organisation

Eduardo Amaldi and *Pierre Auger*, who founded CERN, the European Centre for Nuclear Research in Geneva in 1954, were also the originators of the ESRO concept 4 years later. Later, in 1958, Amaldi established the fundamental scientific and technical principles, which would govern the future ESRO:

- An organisation dedicated to scientific research, essentially controlled by the scientific community in the definition of the programme following the fundamental principle of a bottom-up approach

Table 2.2 List of FIR and submm satellites/space observatories

Mission name (responsible organisations)	Effective aperture (cm)	Wavelength coverage	Year of launch/ mission length	Ref
IRAS (NASA/NIVR/SERC)	57 cm	5–100 μm	1983/10 months	[1]
COBE (NASA-GSFC)	19 cm	1 μm–1 cm	1989/4 years, 1 month	[2]
IRTS (Japan/ISAS)	15cm	1–1000 μm	1995/28 days	[3]
ISO (ESA)	60 cm	2.5–240 μm	1996/28 months	[4]
SWAS (NASA/JPL/Caltech)	44–70 cm	540–610 μm	1998/6 years	[5]
WIRE (NASA/JPL)	30 cm	12–25 μm	1999/failed	[6]
WMAP (NASA-GSFC)	1.4–1.6 m	3.2 mm–1.3 cm	2001/9 years, 1 month	[7]
Odin (Swedish Space Board)	1.1 m	0.5–3 mm	2001/6 years nominal	[8]
Spitzer (NASA/JPL/Caltech)	85 cm	3–180 μm	2003/6 years	[9]
Akari (JAXA-ISAS)	68.5 cm	2–200 μm	2006/5 years, 9 month	[10]
Herschel (ESA)	3.5 m	55–672 μm	2009/4 years, 1month	[11]
Planck (ESA)	1.5–1.9 m	3 μm–1.1 cm	2009/4 years 5 month	[12]
WISE (NASA/JPL)	40 cm	3–25 μm	2009/6 years, 10 day	[13]

[1] The *Nederlands Instituut voor Vliegtuigontwikkeling en Ruimtevaart (NIVR)* was the official space exploration agency of the Dutch government until 2009. The Science Engineering and Research Council was the UK agency in charge of astronomy and space research between 1965 and 1994
[2] Cosmic microwave background experiment
[3] Japan/ISAS sky survey from 1 to 1000 μm
[4] Including contributions from Japan (ISAS)
[5] Small Explorer Mission (SMEX)
[6] Wide-field Infrared Explorer was planned for 4 months but got lost due to the premature opening of its cryogenic lid
[7] Wilkinson Microwave Anisotropy Probe (Cosmic microwave background experiment)
[8] Dual FIR astronomy and aeronomy mission, developed in cooperation between Finland, Canada and France is still operating for aeronomy observations only
[9] Formerly called the Space Infrared Telescope Facility (SIRTF), the fourth and final of the NASA Great Observatories programme
[10] Akari also called ASTRO-F
[11] Including contributions from the USA and Canada
[12] Cosmic microwave background experiment, with large contributions from Caltech
[13] MIDEX

- A stress on the importance of national programmes in the parallel development of each nation's technical and scientific expertise, including a number of cooperative projects
- The development of common laboratories and facilities, with the proviso however that all the scientific work—planning of experiments, design and construction of the scientific instruments, interpretation of the results—should remain

under the responsibility of research groups in the Member States and not within the remit of ESRO

Amaldi was careful to avoid being too precise about the contents of the organisation's future programmes, suggesting that it developed in two phases of increased ambition. The first would give Europeans the time and the opportunity to enhance their know-how and to train their technical and engineering personnel. The second would put them at a level comparable to that attained by the USA and USSR, then leading protagonists in the field. Amaldi also insisted that the future space organisation should keep 'a real European character' and not just be 'a suggestion coming from the US'. These principles still characterise the European Space Agency today.

During the preparatory phase preceding the signature of the ESRO Convention in March 1964, the formulation of the programme was soon confronted with the competition between the domestic programmes of the Member States and those of the future common European venture. Member States were determined to keep the development of ESRO under tight control, in particular its scientific programme and of its budget. Surprisingly, a portion of the scientific community especially from the larger Member States was prompt in supporting these constraints, some of them even arguing against an increase of ESRO's budget during the first 8-year period (1963–1971) covered by the ESRO Convention.

2.2.2 Infrared Astronomy and the ESRO Programme

The skeleton of the ESRO scientific programme, worked out as early as 1961, before the signature of the Convention, foresaw the firing of some 435 sounding rockets and the launching of 17 satellites over the first 8 years. These ambitions were never realised, mainly due to financial troubles as well as technical and managerial immaturity. Scientific priorities and selection of missions were discussed within ESRO's Launching Programme Advisory Committee (LPAC), composed of four prominent scientists and chaired between 1964 and 1970 by *Reimar Lüst* and then by *Johannes Geiss*[6] until 1972.

Fundamental physics, plasma physics, magnetospheric physics and coordinated studies in the interplanetary medium and in the polar ionosphere, as well as cosmic ray observations and high-energy astrophysics in the X-ray and low-energy gamma ray regime, were identified as priorities, while UV astronomy and planetary science were ruled out as they were considered to be too expensive. Cooperation with

[6]Reimar Lüst was already a well-respected and famous plasma physicist and space scientist who later became the President of the Max Planck Society and Director General of ESA from 1984 to 1990. Johannes Geiss, also a space scientist, became well known in the course of the Apollo Programme when he proposed to measure the chemical composition of the Solar Wind with an instrument that Neil Armstrong and Buzz Aldrin deployed on the Moon just after their historic Apollo 11 landing.

NASA (or USSR) was then the only option for Europe to participate in large space telescopes and in the exploration of the Solar System. Infrared astronomy was not even mentioned despite the growing interest of the scientific community and the support of the Space Science Department at ESTEC, which dedicated the 1971 ESLAB/ESRIN Symposium on *Infrared Detection Techniques for Space Research*. The infrared was an area of astronomy, which according to Ernst Trendelenburg, its Chief, 'so far had been carried out from ground-based observatories or from aircraft, balloons or rockets but never satellites' (Trendelenburg 1971). Trendelenburg's interest was an important stimulus to the development of space infrared astronomy in Europe as it initiated some technological research at ESTEC in light-weight and low-cost cryogenic systems.

In 1972, just 168 ESRO sounding rockets had been launched (with 75% success rate), as well as five small satellites and the medium-sized, 3-axis stabilised TD-1 satellite,[7] globally covering the field of plasma and magnetic fields, auroral and polar ionosphere phenomena, cosmic rays and some solar astronomy. The Large Astronomical Satellite, a high-resolution ultraviolet spectrometer, was cancelled but returned several years later in a more modest version renamed IUE, the International Ultraviolet Explorer.

2.2.3 From ESRO to ESA: A 14-Year Stagnation Period in Space Science

Implementation of the LPAC recommendations was soon upset by a major reorganisation in 1970, led by several Member States including France, Belgium and Germany, transforming ESRO from a pure research organisation into an application-oriented agency responsible for the development of the European Ariane launcher (Krige and Russo 2000). They also got the support of the UK, which was keen on developing the MAROTS maritime satellite navigation project. ESRO was also asked to enter into cooperation with NASA on the Shuttle-based Space Transportation System through Spacelab. This so-called package deal led to a doubling of ESRO's overall budget but unfortunately drastically reduced scientific funding,[8] eventually reaching a minimum yearly level of 27 MAU (equivalent to 27 millions euros[9]) in the years 1975–1977, to be spent in the framework of a mandatory

[7]The TD satellites were so-called because they were launched with the US Thor-Delta rocket, which had capabilities that did not yet exist in Europe.

[8]France went as far as proposing to eliminate science from the new organisation, arguing that scientists could also fly their experiments on national and US satellites.

[9]The Accounting Unit (AU) was an internal financial measure established by ESRO, within which all financial transactions were recorded AU. The value of an AU was equivalent to that of the average value of the European Currency Unit (ECU) for the month of June of the preceding year, as determined by the Commission of the European Communities. In October 1995, the ESA Council determined that as of 1 January 1997, ESA should implement an 'All ECU System', and in

programme to which all Member States would contribute according to their GNP share. The mandatory programme would represent about 10% of the total budget of the new agency, now dominated by application programmes and cooperative ventures with NASA. The science budget, though financially extremely modest (equivalent to less than 10% that of NASA), surprisingly did not prevent ESA from developing rather ambitious projects, of a size larger than those foreseen in any of the national programmes.

In 1975, the overall European space science programme including ESA and its member states was characterised by three main components:

1. A modest ESRO/ESA science programme made of medium-class missions (excluding Earth sciences and microgravity), mostly centred on studies of the magnetosphere and the solar wind (HEOS-1, GEOS), and gamma- and X-ray astronomy, with two satellites, COS-B and EXOSAT, about to be launched or in development
2. An increasingly substantial set of national programmes (some of them in cooperation with NASA and the USSR), especially in the fields of solar physics (OSOs), heliospheric physics (HELIOS), plasma and magnetospheric physics (AMPTE) and high-energy astronomy[10]
3. A rapidly growing dependence upon NASA contributions both in the launches and the development of its satellites, suddenly amplified by the so-called ESRO package deal to satisfy nationalistic political interests which increased the pressure to take account of the new opportunities presumably offered by the Shuttle/Spacelab combination.

Infrared astronomy was still absent from the programme, notwithstanding the growing number of experiments and missions that were initiated in various Member States, in particular in the UK, the Netherlands, France, Germany and Sweden.

Following a recommendation by the ESRO Astrophysics Working Group and the still existing LPAC, studies were initiated at the beginning of 1974 on a Large Infrared Telescope for Spacelab (LIRTS), which would have a 2.8 m diameter telescope (larger than any other facility yet envisaged in any space science programme, including NASA) operating at ambient temperature and mounted on a stabilised platform onboard Spacelab.[11] The proposed payload of that huge telescope evidenced the great interest in photometry, polarimetry and Michelson interferometry and also included a heterodyne spectral line receiver—it is interesting that many of these features were subsequently incorporated in ISO and *Herschel*. In parallel, the X-ray Grazing Incidence Solar Telescope (GRIST) on

---

1997, the European Council fixed the conversion rate at one Euro for one ECU. As of 2000, ESA adopted the Euro for all its activities (credit: Eric Morel, ESA).

[10]By 1978, the totality of all European space projects, including the Member States and ESA, amounted to some 70 satellites with only 15 under ESA responsibility.

[11]British and Dutch scientists dominated the membership of the LIRTS Science Team, evidencing the prominent role of these two countries in infrared astronomy.

Spacelab, one of a set of solar telescopes to be developed along with NASA, was also studied. Although neither GRIST nor LIRTS was ever implemented, these two studies were important and influential as they enabled European astronomers to confirm their expertise in solar physics and to enter the new field of space infrared astronomy. In 1978, the Science Advisory Committee (SAC), the successor of the LPAC, in its *Recommendations on the Development of Space Science in the 1980's* (ESA SP-1015) recommended that LIRTS be reoriented and be studied in a 'free-flying' version, i.e. detached and independent from the Shuttle and Spacelab.[12]

At about the same time, in 1977, an international collaboration independent of ESRO was formed by the Netherlands, the USA and Great Britain to develop IRAS (Fig. 2.4). In 1978, under the responsibility of *Reinder van Duinen*, the IRAS Dutch principal investigator, a proposal for a Submm Telescope for Astronomy and Atmospheric Research (STAAR) was submitted to ESA in the framework of the ASSESS programme , with no success. Also in 1978, German astronomers started developing hardware for the GIRL 40 cm telescope to be flown on the Shuttle in the second half of the 1980s, while the Spacelab 2 payload (eventually launched in 1985) already identified a small liquid helium-cooled telescope (IRT). In 1979, anticipation of the results expected from the forthcoming IRAS all-sky survey led to the first proposal to ESA for a non-Shuttle-dependent mission: ISO, the Infrared Space Observatory. Thanks to rapid improvements in infrared detector technology, ISO would offer the scientific community detailed observations and analysis of some 30,000 infrared sources with 1000 times better sensitivity and 100 times better angular resolution at 12 μm than IRAS. ISO was eventually selected by ESA in 1983. At nearly the same time, a year before the launch of IRAS, in response to its call for ideas issued in July 1982, ESA received a proposal for a Far-Infrared and Submillimetre Space Telescope, FIRST, the ancestor of *Herschel*. Infrared astronomy, finally, had climbed all the steps leading to the start of a very promising and successful future. It was time for the ESA science programme to recognise the growing importance and scientific power of the field and the strong community involved.

2.2.4 Fighting to Recover ESRO's Lost Science Budget

By the late 1970s, the scientific programme of ESA had reached a high level of ambition and expectation, including large astronomical telescopes. It envisaged the participation in NASA's large space telescope (the future HST) and in the out-of-ecliptic International Solar Polar Mission (ISPM) comprising two satellites orbiting

[12]Created in 1975 by the ESA first Director General, Roy Gibson, the SAC absorbed some competencies of the former LPAC of ESRO. In 1982, the terms of reference radically changed. It became SSAC, the Space Science Advisory Committee, and it was tasked with advising the Directorate of Science under the authority of Director General.

above the poles of the Sun, one European and one American. It also shared responsibility for the IUE (in cooperation also with the UK) and for ISEE-B, one of three satellites in the International Sun-Earth Explorer programme led by NASA. Most of the ESA scientific programme was then clearly dependent upon NASA plans, with ambitions far above the financial resources of the new ESA.

In June 1979, the ESA Council, inspired by France and Britain, opposed a 50% increase of the, still capped, 27 MAU (1970 price level) annual budget of the mandatory programme,[13] to reach 120 MAU/year at 1978 price level for the classical space science disciplines as recommended by the SAC. The Council nevertheless accepted that Earth sciences and microgravity science on Spacelab be funded outside that envelope as optional programmes. In that context, two major events occurred which had a strong influence on the future development of the ESA space science programme:

1. The successful launch of the first Ariane-1 on Christmas Eve in 1979
2. The unilateral decision by NASA in 1981 to quit the joint ISPM mission, initiating a major crisis of trust between ESA and NASA

The success of Ariane meant that Europe was no longer reliant on the US for launching its satellites, and the ISPM crisis re-enforced the rapidly growing ESA desire for independence from NASA. In the wake of these events, ESA's Science Programme Committee (SPC)[14] decided to develop the Giotto and Hipparcos missions and to initiate studies on ISO, Cluster and SOHO, all missions possible without NASA participation. Had ESA come of age? Not quite so, because the capping of the budget since 1971 was no longer compatible with the level of maturity reached by the European scientific community, and a change of science policy had to be adopted. Until then, the selection of missions was made coup-by-coup, budget permitting. It was the result of increasingly tough competition between a growing number of good missions proposed by a rapidly developing community of more than 2000 scientists in Europe, covering the whole range of space science disciplines: astrophysics, solar and heliospheric physics and planetary exploration. Proposals were submitted to the SAC advising the ESA Director General for final choice through the Astronomy Working Group (AWG) and the Solar System Working Group (SSWG). But that ad hoc procedure did not reflect a well-thought-out strategic science policy that would deploy resources most effectively and ensure the best exploitation of the growing talents of the community. It was also lacking in the coordinated technological preparation necessary to undertake the more and more complex missions that were envisaged. In other words, ESA did not have a long-term programme based on a clear science policy. The

[13] Totaling 76 MAU in 1978 price levels.

[14] The Science Programme Committee, SPC, is a decision committee overlooking the mandatory scientific programme. One delegate, usually assisted by a science advisor, represents each participating member state. The SPC votes on all financial matters such as the approval of the budget of the science programme and decides the approval of new projects, as well as their continuation.

frustration of the science community was high and resulted in some bitter criticisms of the ESA management. The election in early 1983 of a new ESA Science Programme Director[15] at the head of the science programme just after the selection of ISO in March 1983 offered an opportunity to bring about radical change.

2.2.5 Formulation of an ESA Strategic Long-Term View: Horizon 2000

Immediately after he took his duties, the new Director developed a broad-based process that would ultimately result in the setting up of the Horizon 2000 long-term programme founded on the science-driven bottom-up approach dear to Amaldi, established through the support of all partners: scientists, industry and member states (Bleeker and Bonnet 1984; Krige and Russo 2000). The formulation of that programme was carried out in earnest between October 1983 and July 1984. Thanks to the support and the drive of Vittorio Manno, Head Coordinator of ESA's Scientific Programme, in November 1983 ESA issued a call for mission concepts addressed to the broad European and international science community encompassing more than 2000 scientists. By the end of 1983, ESA had received 77 replies. A Survey Committee, chaired by Dutch astronomer Johan Bleeker, was formed, supported by Vittorio Manno and five panels and topical teams involving in total some 50 scientists, tasked with analysing priorities and requirements as well as identifying the technologies necessary for implementing the missions. The Survey Committee finally built up the architecture of a coherent overall programme in the course of a 3-day meeting in Venice in June 1984.

The Horizon 2000 long-term programme, which can be compared with the decadal surveys regularly established by NASA and the US National Research Council, aimed at defining a comprehensive and coherent programme spread over a 20-year (not 10-year) period. It encompassed all classical space science disciplines (with the exception of Earth and microgravity sciences), and its balance had to satisfy the whole community involved. No discipline was excluded, contrary to the ESRO LPAC policy of 1970. It was based on four main cornerstones, all representing flagship areas of science, complemented by ten opportunities for medium and small missions (Fig. 2.5). It involved and unified the more than 2000 scientists who participated in its preparation. It also proposed a mission sequence arranged within an appropriate time schedule (Fig. 2.6) and sketched the technological research and development and industrial preparation indispensable for its implementation.

In Fig. 2.5, the four Cornerstones represent the priorities identified by the Survey Committee as being of a highly competitive scientific content and reflecting the areas of expertise and excellence of the community for which continuity of work,

[15]Incidentally, the author of this paragraph!

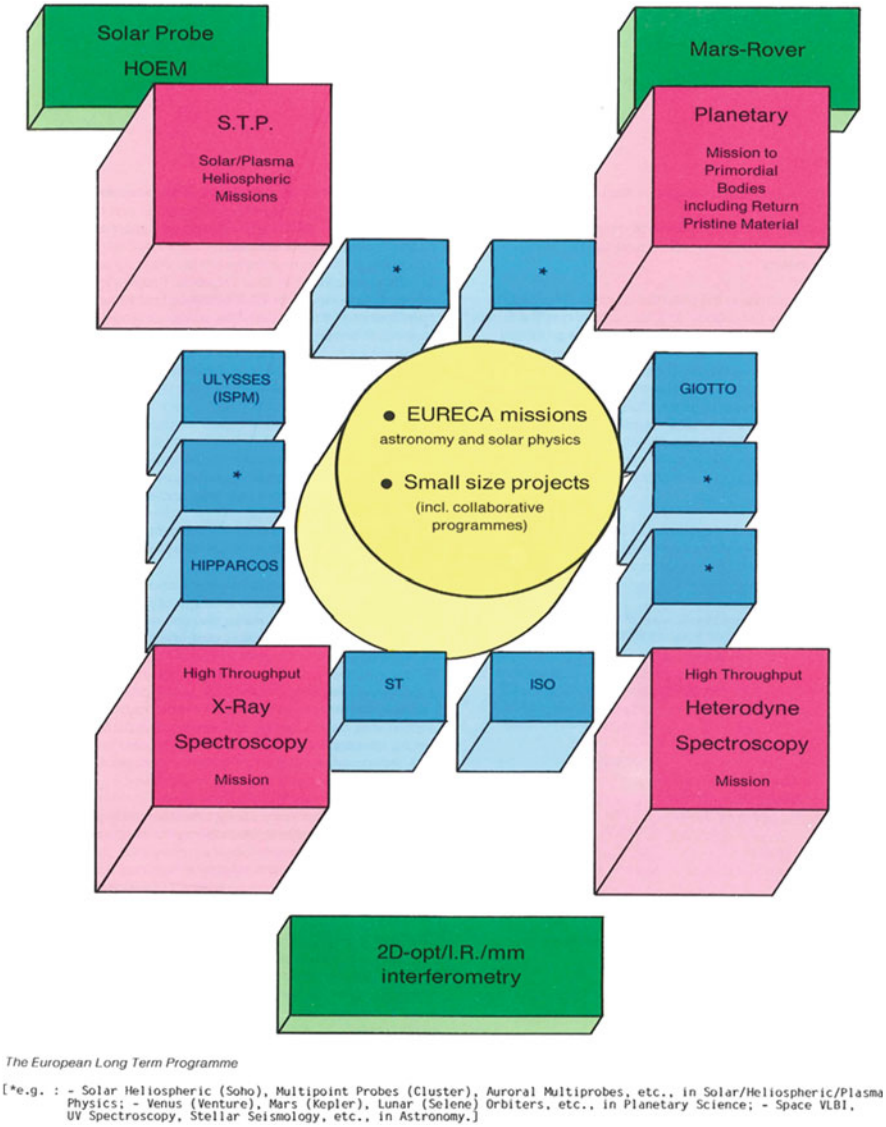


Fig. 2.5 Original architecture of Horizon 2000 as designed in Venice, June 1984, Bleeker and Bonnet 1984. © ESA

both in research institutes and in industry, was an essential science policy element in their selection in the program:

- The Solar Terrestrial Programme would include SOHO and Cluster, two medium-sized missions already under study at ESA, which represented an important contribution to solar physics and plasma physics by the respective communities in Europe and unified these through common physical problems.

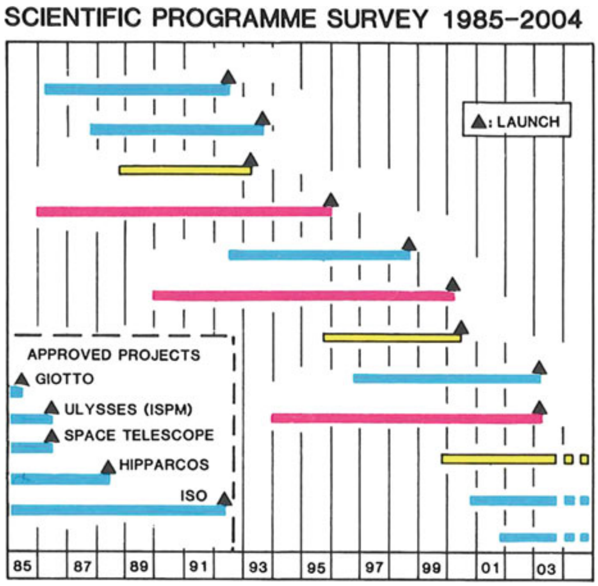


Fig. 2.6 Sequence of missions in Horizon 2000 as envisaged in 1983, showing the continuity with the pre-Horizon 2000 programme. Note that the launch dates for Ulysses and the *Hubble* Space Telescope were later shifted by 4 years as a consequence of the Challenger accident and Hipparcos 1 year because of the 1988 Ariane 4 accident. The first Cornerstone, constituted by SOHO and Cluster, is here represented as two medium-sized missions: they were consolidated in a single Cornerstone after the June 1984 Survey Committee Venice meeting. This sequence is based on an assumed yearly budget increase above inflation of 7% over 10 years. Bleeker and Bonnet 1984 © ESA

- The Mission to Primordial Bodies represented the large interest of the community in that field, illustrated by the important effort developed in the preparation of the Giotto mission to Halley's comet. It identified the possibility of returning pristine material from either an asteroid or a comet. That Cornerstone would later be redefined as Rosetta, also a comet mission.
- The High Throughput 0.1–20 keV X-ray Spectroscopy Cornerstone represented both the very high level of interest and the strong capability of the European high-energy astrophysics community, which had submitted some ten mission concepts, among which was the XMM multi-mirror mission, under assessment study at ESA. Its scientific objectives were considered by the Survey Committee of the highest priority. However, in view of the design-to-cost approach it intended to follow, ESA decided not to identify specifically the Cornerstone as XMM because it wanted to preserve the possibility of reassessing that mission concept, while maintaining as much of its original scientific objectives. Ultimately, the Cornerstone was named XMM-*Newton*, a new version of the original XMM.
- The High Throughput Heterodyne Spectroscopy mission followed a similar evolution as XMM for exactly the same reason: ESA did not want to be bound by the FIRST mission concept which had to be reassessed in view of keeping its

cost within a fixed envelope. Eventually the Cornerstone was later renamed *Herschel*.

2.2.6 The Horizon 2000 Philosophy: Scientific Excellence, Schedule Respect and Design-to-Cost

The new approach required a major change in the management of the science programme and in the way its missions were to be selected. Before 1982, the procedure was the result of a strict competition at the end of phase-A, after prioritisation by the Astronomy and Solar System Working Groups in their respective fields of competence and final recommendation made by the top-level advisory body, the Space Science Advisory Committee (SSAC). In contrast, the Cornerstones would now be considered as fixed elements of the future ESA Science programme, representing both its long-term set of high-priority goals and its central core: none should be submitted to re-competition or later be abandoned. They were conceived to be under the control of ESA and not dependent upon other agencies policies, schedules or budget availabilities. International cooperation was obviously welcome and encouraged for enhancing the missions whenever possible but in full respect of their ESA established schedule and financial constraints. They could be redesigned and modified, provided that their original scientific objectives were not modified or abandoned, in the event that cost over-runs might occur or that new advances in science or in relevant technologies might arise in the course of their development. In contrast, the selection of medium and small missions, which represented the flexible part of the programme, would follow the fully competitive process that prevailed in the pre-Horizon 2000 era for all missions.

How Did FIRST Become One of the Four Cornerstones of Horizon 2000?

Paradoxically, FIRST was selected in spite of the fact that only two IR astronomy proposals were submitted in response to the 1983 call for mission concepts: a small infrared telescope called SMIT submitted by a French group and, FIRST, submitted by a team of 14 European scientists (cf. Sect. 3.2). This surprisingly small number can be interpreted as resulting from a self-screening approach by the infrared astronomy community, which wished to preserve the innovative and ambitious character of FIRST as a mission combining several important and unique capabilities. The submm domain was the last remaining gap in the electromagnetic spectrum left unexplored, allowing study of the continuum radiation from dust and a large number of very important atomic and molecular transitions essential for studying the physics and chemistry of the cool Universe in the temperature range

(continued)

3–1000 K. It also offered the appealing capacity of an order of magnitude increase in angular resolution (at the same wavelength) with respect to ISO. It reflected the excellence and expertise of the European science community acquired through ground-based, aircraft-, balloon- and satellite-borne experiments in the technology of submm instrumentation. It would allow for the development of a timely mission in submm imaging and spectroscopy and would offer, according to the Survey Committee words, 'a unique opportunity to take the lead in studies of star and planetary system formation, origin and evolution of galaxies and small scale granularity of the cosmic background radiation' (Bleeker and Bonnet 1984).

The 20-year time span of the programme was considered to be realistic, given the capabilities of the scientific community and of industry, the state of maturity of the technologies necessary for the development of satellites, subsystems and their respective scientific payloads as well as the financial resources of the research institutes and of the Member States. The programme was to be considered as a whole set, directly implying that all missions needed to be implemented within very strict cost limits for the ESA share of their development, which in the ESA system does not include the cost of scientific payloads and of data analysis. That approach automatically meant that the pre-Horizon 2000 missions already in development[16] had also to obey as far as possible the rules imposed on all Horizon 2000 missions and therefore become an integral part of it (Fig. 2.5).

Based on a parameterized cost analysis of missions already developed, the Cornerstones were allocated a cost envelope (to ESA) of 400 MAU (1984 economic conditions),[17] equivalent to 2 years of the envisaged annual budget. Medium mission costs were arbitrarily capped at half that value, i.e. 200 MAU, equivalent to 1 year of the annual budget, and small missions at 100 MAU. The 200 MAU annual budget level was compatible with an average frequency of launches, all missions considered, of one per year over 20 years. It represented an overall budget increase, above inflation, of little more than 50%, with respect to the 1983–1984 financial situation, to be reached progressively at an optimum annual rate of 7% over no more than 10 years.

[16] These included Giotto, Hipparcos, ISO and ESA's participation to HST, and Ulysses.

[17] With the exception of the STSP Cornerstone regrouping SOHO and Cluster (totaling five satellites and representing the scientific interests of about 30% of the whole European Space Science Community), which was allocated a higher limit of 470 MAU.

2.2.7 Paving the Way to Success

The months following the Venice meeting were dedicated to 'selling' the plan to its stakeholders: the scientific community, the Member States and the industry, eventually paving the road to a successful decision by the ESA council, the only body to vote (with a unanimous decision required) on the level of resources of the scientific programme. Interestingly, a substantial number of scientists, even though they had fully participated in the formulation of the plan, were sceptical that it might eventually achieve the ambitious goal of putting financial stagnation to a rest. That attitude was particularly evident early in the process among the ESA project scientists of the ESTEC-based Space Science Department (SSD) who were used to the previous 'all-competitive' selection approach and somewhat jaded after having witnessed so many unsuccessful attempts to increase the scientific budget.

Even the newly elected Director General, Reimar Lüst, first Science Director of ESRO and Chairman of the former ESRO LPAC, was doubtful that a plan extending over 20 years would convince the Member States to unblock the required budget so long before its full completion. He was reluctant to commit ESA's scientific budget and resources over a 20-year period, which would rely on sustained support of all the SPC and Council Member States, through a series of some seven consecutive budget-planning cycles every 3 years, to adopting a continuously increasing level of resources of the budget. The Science Programme Director decided to demonstrate to him the support of the scientific community and invited him together with the (fully supportive) Chairman of the Survey Committee, Johan Bleeker, to a dinner in a Paris restaurant close to Place de l'Étoile, where he was briefed. The good atmosphere of that meeting, together with the frankness of the exchanges, which placed strong emphasis on the necessity to engage as early as possible in the technology preparative activities prior to starting the ambitious missions of the programme, helped to convince him that he should support the plan, to which he subsequently lent his full support. Indeed, his support was essential, and that first step was already a very positive sign for the future of Horizon 2000. The Member States—which had been left a little bit aside during the formulation of the plan—were also briefed through a series of meetings organised in each of the capitals, in which representatives of the scientific community participated. These meetings eventually attracted their interest and convinced some of them that the advantages of the new ESA strategy meant that it was worthy of support. Industrialists also participated in these meetings, not because they had to be convinced, but because they were strong supporters of the new strategy, realising its potential in targeting early the technological developments that would ultimately give their future bids a higher chance of winning forthcoming competitions. Industry thus played an important role in the delicate road leading to the success of Horizon 2000.

Reimar Lüst presented the programme in January 1985 in Rome to the ESA Council at ministerial level. The main agenda items were the development of the Ariane 5 launcher, the participation in the future NASA space station and the adoption of the new level of resources for the scientific programme. The atmosphere at the meeting was positive: ministers agreed to start preparatory work on Ariane 5 and agreed to start the development of Columbus (ESA's participation to the Space Station), while taking note of the French decision to undertake the Hermes crewed space plane programme (van den Abeelen 2017). They also granted the science programme a 5% increase over the 1985–1989 period (instead of the 7% requested), with the level of funding reaching 162 MAU (1984 prices) in 1989,[18] to be subsequently extended at the next ESA Ministerial Meeting in The Hague in 1987. As reported by Krige and Russo (2000): *With Member states committing themselves to almost double the ESA expenditure in seven years, reaching about 2600 MAU by 1993 (1986 prices), it was not an outrageous idea to spend less than 8% of this money for pure science*. The clever diplomacy exerted by Reimar Lüst and the scientists' intense lobbying eventually achieved success: it put to an end the stagnation of 15 years of the ESRO/ESA space science budget, allowing the Horizon 2000 programme to start.

The successful outcome of the Rome Council meeting would probably not have been possible without the existence of the Horizon 2000 programme. In Rome, the Ministers were faced not only with a general plea for 'more money for science', as had been the case with the SAC Report of 1979, but with a coherent and comprehensive development plan for European space science activities over 20 years, including a mix of large- and medium-sized missions, an appropriate balance among the various disciplines, a suitable development schedule and a well-defined financial framework. The need for advanced long-lead technological developments made it possible to introduce into the programme those ambitious missions, which could not be approved in the pre-Horizon 2000 era because their feasibility could not be established. Furthermore, the plan would provide a reference around which the various sectors of the space science community including national programmes might eventually organise themselves in order to take the best possible share of the missions. Finally, the plan would help planning future collaborative efforts with NASA and other national space agencies.

[18]France was in favour of a more modest increase of 3–4%, possibly giving up one of the Cornerstones, a suggestion strongly opposed by the ESA Director and the Survey Committee Chairman, who pointed out that 'the Cornerstones were already the outcome of a careful selection procedure ensuring a balance between the various scientific areas that would be jeopardized' unless this balance was respected. The chairman of the SSAC added: *If this minimum threshold was not reached*, *Europe would lapse back into* '*a posteriori*' *planning*, *since the effect of removing one Cornerstone would be to break up the coherence of the plan and destroy all the others* (Krige and Russo 2000).

2.2.8 Implementation of Horizon 2000

Respecting both the costs ceilings imposed on the Cornerstones and smaller-size projects as well as their schedules, no longer mission by mission, but for the totality and the integrity of an entire programme, engendered a true revolution in mission management, leading ESA engineers and scientists, the scientific community and industry to obey the Horizon 2000 rigour. For the first time, the SSAC and the SPC were committed to an implementation plan involving an identified series of large missions not open for re-competition. The SPC delegates had also to accept a new approach for the decision-making process, with a 20-year horizon instead of just the few years spanning individual phases B and C/D. That, for some, was much longer than their own mandate's duration. That was also the case of the ESA management where the continuity was essential, and opportunely facilitated by the unusually long duration of the eventual tenure of the Director, totalling 18 years in the post. The rigour was essential to safeguarding the scientific integrity of missions' objectives and the coherence of the programme. Project managers and project scientists also played an important role in identifying possible synergies between projects in order to take advantage whenever possible of using recurrent subsystems, eventually resulting in cost savings for subsequent missions (e.g. Rosetta serving Mars Express, Mars Express serving Venus Express, XMM-*Newton* serving Integral).

The first decision step was to assume that the 5% annual budget increases would continue beyond 1989. That was confirmed at the ministerial meeting in The Hague, in 1987, which agreed that the budget should reach 216.7 MAU in 1992.[19] Another delicate issue was the sequence of the Cornerstones after fine-tuning their technical definition to fit their cost limit and the confirmation of their technological readiness. That sequencing exercise started in the course of 1985–1986 through a series of four dedicated workshops (Table 2.3) to which the scientific community and representatives of Member States and industry were invited. ESA also organised a series of rendezvous with the main prime and subcontractors, during which the progress in design and implementation in the design-to-cost framework were systematically reviewed. That sustained and friendly dialogue was important for the overall success of the Horizon 2000 strategy.

Table 2.3 Sequence of the four Cornerstones dedicated workshops

Cornerstone	Workshop date	Place
STSP (SOHO-Cluster)	30 April–May 1985	Garmisch-Partenkirchen(D)
XMM (XMM-*Newton*)	24–26 June 1985	Lyngby (Denmark)
FIRST (*Herschel*)	4–7 June 1986	Segovia (Spain)
CNSR (Rosetta)	15–17 July 1986	Canterbury (UK)

[19]This is equivalent to 198.3 MAU at 1984 prices.

The candidate for first Cornerstone position was naturally the STSP, formed of SOHO and Cluster, both projects already under phase-A study and with relevant technologies within reach. Second in line was the High Throughput X-ray Spectroscopy mission, (XMM) whose technology readiness was at a lower level. Both had nevertheless to be significantly descoped to fit their costs within the assigned limits, as was also the case for the two other Cornerstones.

Concerning the Comet Nucleus Sample Return mission, occupying the position of 'planetary' mission in Fig. 2.5, the concept of returning to Earth a sample from a comet was abandoned after NASA, which expressed some interest in participating, eventually decided not to support their share. A purely European venture, renamed Rosetta (including the possibility to incorporate a small lander for surface science) was approved by the SPC in November 1993, positioning it as the third Cornerstone after it was confirmed that the submillimetre astronomy mission was still above the assigned cost limit (Chap. 3). FIRST, later renamed *Herschel* at a meeting in Toledo in December 2000, became the last of the series. The concept of flying an early version of FIRST onboard the International Space Station (ISS), whose facilities for in-orbit assembly and cryogenic servicing by astronauts were strongly advertised by NASA and also by the ESA ISS management as attractive, had to be abandoned,[20] but its study required more work by the teams involved and introduced substantial delays in the development of the Cornerstone.

The Horizon 2000 concept had several very positive side effects. It unified the scientific community behind the integrity of the programme. Every scientist involved in the Cornerstones had an interest in securing its overall balance: if one of the four faced difficulties, the remaining three would also suffer. Those who were not fully involved could use the opportunities offered by medium and small missions to satisfy their interests. The very challenging mission objectives and the design-to-cost approach stimulated the ingenuity of both industry and the experimenters, who could invest efforts in the development of new technologies early in the process, and in establishing new international cooperative ventures. Indeed, the solid framework of the plan changed the shape of international cooperation. The overall programme could survive independently of any US or other non-European defective connections. While Cornerstones had to remain under ESA leadership and be consistent with ESA's own technical and financial means, cooperation with non-European partners was not at all excluded, in fact encouraged, as it could bring added capabilities to these European-led missions. In fact, immediately after its approval in Rome, Horizon 2000 attracted the interest of the numerous other nations, with USA, Russia, Japan and later China among the most visible. It also reinforced ESA's position vis-à-vis NASA from one of subservient partner as in the ESRO and ESA pre-Horizon 2000 era, to one of leader. Cooperative agreements

[20] The option looked attractive at first sight, but the drawbacks of having to refuel 6000 l of helium every year, and, knowing from the IRT experience on Spacelab 2 that the observing conditions on a manned platform in Low Earth Orbit are not optimum, forced the ESA Scientific Programme Directorate to reject it.

were negotiated on all four Cornerstones, benefiting the overall programme. However, the programme was robust enough to survive even in case of unilateral withdrawals from the considered partners. That was in particular the case with Rosetta on which NASA proposed providing one of two comet landers and later withdrew, leaving the mission with just the European lander (Philae), the success of which during the cometary encounter in 2014 has provided an ultimate confirmation of the value of the Horizon 2000 concept.

2.2.9 Problems on the Horizon

At mid-course, 10 years into the development of Horizon 2000 (Fig. 2.7), it was time to review its implementation, assess its performance, possibly correcting for its defects and preparing its prolongation over the next 20 years. Such a review was also recommended by the ESA Council of 1992 in view of major decisions to be taken at the following one scheduled for the end of 1995 in Toulouse. Naturally, not everything went as expected. Difficulties of a managerial, technical, political and conceptual nature were encountered. They were overcome through management redirections eventually allowing the integrity of the overall programme to be protected and most of its objectives to be fulfilled. In conducting that review and the continuation of the programme, ESA followed the same approach as for the formulation of Horizon 2000 10 years before: it organised a call for ideas whose replies were analysed by a Survey Committee, chaired this time by *Lodevijk Woltjer* assisted by Dr. Giacomo Cavallo.

In the course of their discussions, the new Survey Committee observed that the cost envelopes assigned to Rosetta and FIRST as well as medium-sized missions were becoming difficult to maintain. There was however no discussion that these might be increased as that would unavoidably lead to longer intervals between all

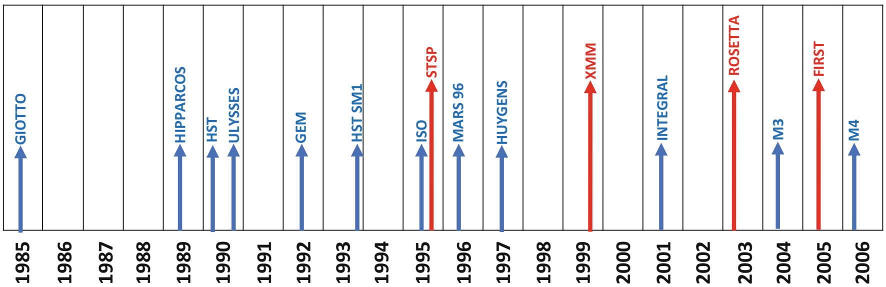


Fig. 2.7 Launch dates of the Horizon 2000 missions established as of 1995. Note the launch of FIRST (*Herschel*) already foreseen end of 2005 and of M3 (*Planck*) in the second half of 2004 (ESA SP-1180). Mars 96 refers to the (non-ESA) Russian mission that was lost on November 1996, to which several Member States provided a substantial share of the payload. In January 1997, Mars Express was proposed as a replacement and eventually occupied the medium-class mission M4 slot. © *ISSI*, *RM Bonnet*

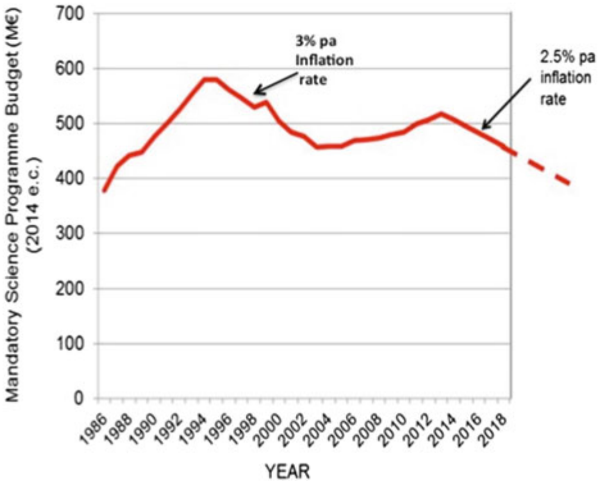


Fig. 2.8 Evolution of the ESA science budget from 1986, following the 1985 Rome ESA Council meeting, which raised the annual budget from 378 millions euros in 1986 to 578 millions euros in 1995, expressed in 2014 economic conditions. The two periods of negative gradient beginning in 1996 at a level of 578 millions euros after the Toulouse meeting and in 2013 at a level of 510 millions euros after the Naples meetings correspond to when flat cash was imposed by the Member States. From Southwood (2014)

missions. Some savings, though not yet very large, were already identified through the use of recurrent systems for some missions. The Committee therefore concluded that the budgets allocated to Cornerstones and medium-sized missions should remain unchanged, but that whenever economies might materialise, they should mainly be used to increase the frequency of missions rather than to simply make them bigger, even suggesting that 'mission proposals at lower cost should be regarded with particular favour'.

The new plan was presented to the ministers in Toulouse in November 1995. However, the good spirit of the January 1985 Rome meeting was absent there. Under the pressure of Germany and the UK, the ministers decided not to compensate the yearly budget for inflation anymore, meaning that the purchasing power of the science programme would decline 3% every year as of 1996 (when it reached 578 millions euros), a trend that lasted until 2001, when ministers granted the programme a total of 1869 millions euros for the period 2002–2006 to maintain operations for missions already launched (*Hubble* Space Telescope, Ulysses, Cluster II, SOHO, *Huygens*/*Cassini*, XMM-*Newton*), continue missions in the development phase (Integral, Rosetta, Mars Express, Smart-1, *Herschel* and *Planck*) as well as the work on missions under study (*Bepi Colombo*, GAIA LISA-Smart-2 and Solar Orbiter (Fig. 2.8).

Political decisions were not the only obstacles that Horizon 2000 had to face. Delays and cost over-runs were encountered as a consequence of several launcher problems. These included the Challenger accident of 28 January 1986, which killed its full crew of seven astronauts at the eve of the planned launches of *Ulysses* and of the *Hubble* Space Telescope, both scheduled for October 1986.[21] HST was eventually launched on 24 April 1990[22] and *Ulysses* on 6 October 1990. That 4-year delay affecting both projects resulted in a total cost increases for both of about 150 MAU, something that threw a shadow on the implementation of Horizon 2000, especially on the first Cornerstone, STSP. SOHO was originally planned for a launch on the Space Shuttle free of charge for ESA, and the consequences of the accident on NASA's budget were not fully clear, with consequent uncertainty about what solution would eventually be selected for its launch. As a consequence, Cluster might become a victim of the extra cost of SOHO if ESA had to pay for its launch, with the risk of damaging the coherence of the STSP Cornerstone. NASA and ESA finally agreed to a free SOHO launch on an Atlas-II-AS, which occurred on 2 December 1995. The SPC delegations in a remarkable spirit of support to Horizon 2000 eventually confirmed the integrity of the STSP double-mission architecture, insisting that every effort should be made by the Science Directorate to keep its costs within the 470 MAU limit, de facto agreeing that Cluster should use the second flight of the new Ariane 5 launcher. This was an option that unfortunately never materialised as Cluster was shifted to the first test flight (designated 501), which resulted in the tragic loss of its four spacecraft (Cavallo 1996)—see text box below.

Immediately after the 501 accident, the French government strongly objected to a replacement of Cluster, accusing ESA of irresponsibility for having agreed *to fly such an important mission* on the first flight of a new rocket. ESA equally strongly opposed that position as one that would have destroyed the architecture of the first Cornerstone and having knock-on effects damaging the entire programme. In less than 2 years, a solution was found through the opportunity offered by the Russian Soyuz launcher. Cluster, renamed Cluster II, was placed in orbit using two successive Soyuz launches from Baikonour[23] on 16 July and 9 August 2000. That solution had its price though, affecting severely the development of *Herschel*, as ESA had to introduce in its financial projections the costs of developing four new Cluster spacecraft plus their adaptation and launch on two Soyuz. Furthermore, the SPC endorsed it under the condition that ESA also supports the costs of redeveloping four new sets of the Cluster payload (normally not under ESA financial responsibility), which it eventually did. Unfortunately, the costs of accommodating these

[21]The ESA Ulysses project team was already preparing for the launch of the Discovery Shuttle at NASA's Kennedy Space Center from where they witnessed live the Challenger tragedy.

[22]In fact, the HST ground software was not ready in 1986, and indeed was barely ready by the April 1990 launch.

[23]The capacity of Soyuz was not enough to launch the four satellites at once. Therefore two separate consecutive launches were necessary to launch all satellites of the Cluster mission.

unforeseen additions in the Horizon 2000 programme forced a substantial restructuring of the development of both *Herschel* and *Planck* (selected in 1997 and occupying the M3 position in Fig. 2.7) to be launched on the same Ariane 5 and to reach their respective halo orbits around L2. This is further described in more detail in Chap. 3. In fact, both *Herschel* and *Planck*, resting on very advanced technologies, required a longer than expected preparatory work, and their launch could not occur earlier than 2007. That solution, although not at all popular in the infrared astronomy community, eventually proved to be a great success: on top of the substantial scientific achievements of the two missions, it increased the visibility of the acquired ESA leadership in far-infrared astronomy, with two major flagship missions launched at the same time.

Reasons for Selecting the Test Flight of Ariane 5 to Launch Cluster

The Ariane 5 launch option was retained by ESA for two reasons. First, the political pressure from the French space agency, CNES, which was in charge of the development of Ariane 5, was very high that the science programme should use the second flight of the new launcher at marginal cost. ESA accepted that solution because it allowed maintaining the cost of the Cornerstone within its allocated financial envelope of 470 millions euros. Furthermore, Ariane 5 was designed to launch astronauts onboard the Hermes manned-tended Shuttle and to be a very safe launcher, an argument that helped obtain the agreement of the scientific community for that option. The first launch of Ariane 5, designated N° 501, was offered to ARTEMIS, the Advanced Relay and TEchnology MISsion, an experimental data relay satellite. Eventually, in the course of its development, ARTEMIS suffered substantial delays and missed the 1995 launch opportunity (in fact it was launched in 2001 onboard Ariane 510). Consequently, Cluster was swapped from flight 502 to the first Ariane 5 flight, 501, a change of opportunity that was readily accepted by the Cluster PI's, since it secured an earlier flight of the mission. Cluster was launched on 4 June 1996 from Kourou. Unfortunately, the four Cluster satellites were lost 37 seconds after takeoff when the rocket exploded, the victim of a software error.

Less dramatic, but also costly, was the delay of 1 year in the launch of the Hipparcos astrometry mission, originally foreseen for 1988. Hipparcos was eventually launched towards a geostationary orbit on 8 August 1989. Unfortunately, the apogee boost motor that was supposed to propel the satellite into its final geostationary orbit did not ignite, and the mission was forced to operate on a non-optimal orbit. Despite the strong pressure from the astrometrist community who wanted an immediate replacement of the mission, ESA, concerned about the consequences of that option for the integrity of the programme, strongly opposed that pressure. It agreed however to make all possible efforts to improve the status and operability of Hipparcos, raising as much as possible its transfer orbit using the attitude control thrusters. Eventually, the mission objectives were more than 100% achieved, and the mission can a posteriori be considered as a major success, having pioneered the field of space astrometry. On 11 December 2002, Ariane 5 failed to place a 10 ton commercial payload on its geostationary transfer orbit. As a consequence, the following flight, which was supposed to launch Rosetta towards Comet *Wirtanen*, was postponed, and Rosetta was delayed by 1 year and retargeted to Comet 67/P *Churyumov-Gerasimenko*. The greater distance of that comet was to add 3 years to its already long trip across the Solar System.

Conceptually, Horizon 2000 should have been accomplished in 20 years, counting from the January 1985 Rome meeting, with its last project supposed to come to completion in 2005. The problems just described explain the main reasons why that goal was not achieved, with *Herschel*, as the last Cornerstone, finally launched in 2009, i.e. with a 4-year delay. In its determined policy of maintaining the integrity and balance of the programme, ESA's Science Programme Directorate was criticised for having introduced rigidity and sacrificing the frequency of small and medium missions. These criticisms were particularly strong from part of the scientific community and from some of the SPC delegations, which were deprived of their important role in approving mission selections. From those perspectives, these criticisms are certainly justified. However, referring to Figs. 2.6 and 2.9 and the opportunities for four small-to-medium missions identified therein, that loss of flexibility was compensated by a continuous recourse to adaptability.

Indeed, M4 cannot be identified as such in the history of Horizon 2000-2000+, extending the Horizon 2000 programme. That is partly because as early as January 1997, the SSAC suggested to compensate the Solar System community for the loss of Mars 96 by developing Mars Express and to make it 'cheaper and faster' than usual, taking larger risks, and using whenever feasible commonalities which existed with some of the Rosetta subsystems. Mars Express was launched on 2 June 2003, 4 and 1/2 years after its approval by the SPC. A similar approach was used using Mars Express this time to enable Venus Express, launched 2 years later on 9 November 2005, adding its contribution to the M4 slot. Both missions can be identified as major successes in Solar System exploration. Simultaneously, the astronomy community, looking at the possibility of contributing to the successor of HST, was keen to support an ESA involvement in the NASA-led New Generation Space Telescope, NGST (later to be renamed the *James Webb* Space Telescope, JWST). Together with Huygens, Integral, *Planck*, SMART-1, LISA Pathfinder and pre-Horizon 2000 missions, these opportunities testify that the non-Cornerstone share of Horizon 2000 has been extremely well balanced and of great scientific success.

An important element of the original Horizon 2000 plan (Fig. 2.5) that met difficulties in implementation was the intention to include flight opportunities using Space Station and Eureca (a small reusable platform designed to be launched from and subsequently retrieved by the Space Shuttle). The Horizon 2000+ Survey Committee was not very enthusiastic about small missions, especially for astronomy, as they considered them to be the prerogatives of the Member States. Instead they favoured a scheme whereby ESA would collaborate at a small ratio with its

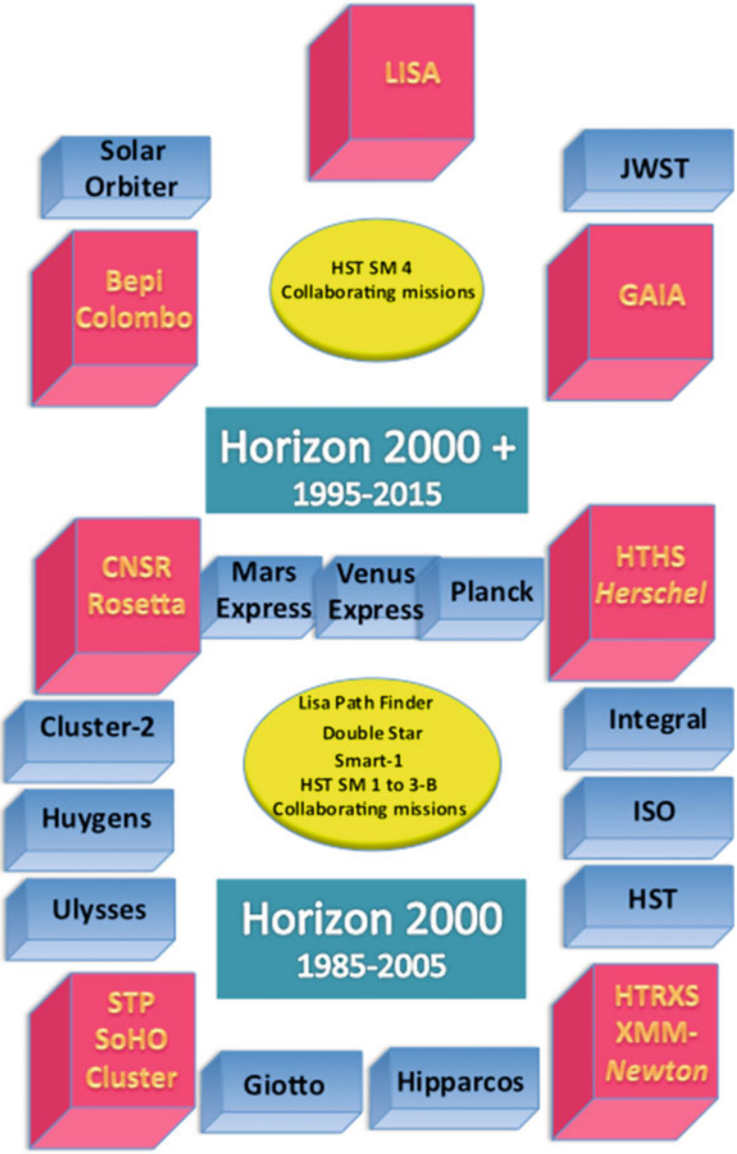


Fig. 2.9 ESA scientific missions implemented in the framework of the original Horizon 2000 concept (see Fig. 2.5) and its prolongation Horizon 2000+ spanning the period 1985–2015. © *ISSI, RM Bonnet*

own Member States' missions. That was the case with the stellar seismology and planet-finder mission, Corot, led by the French space agency, CNES. The only Horizon 2000 small mission that clearly materialised was SMART-1, a lunar mission testing solar electric propulsion and successfully launched in 2001. The second one, SMART-2, later renamed LISA Pathfinder, evolved from small mission status into an M-class mission, as a precursor to the future Horizon 2000+ LISA Cornerstone. Furthermore, in the course of its development, Horizon 2000 opened several opportunities for mission extensions (spacecraft status permitting) such as those for IUE, Giotto, Ulysses and nearly all Cornerstones, which were not foreseen originally in Horizon 2000 and can be considered as extended scientific opportunities of a 'small mission' category.

Continuity beyond 2015 was assured by ESA through the Cosmic Vision scheme established by Professor David Southwood in 2005 and by Professor Alvaro Gimenez Cañete after 2011. Missions like Bepi Colombo, LISA, JWST and Solar Orbiter though included in the Horizon 2000+ plan would be launched beyond 2015 during the period corresponding to the Cosmic Vision scheme (ESA BR-247). The Horizon 2000 Cornerstones in carmine boxes are placed in chronological order from the bottom with astronomy missions on the right and Solar System missions on the left, while the medium missions in blue boxes are also arranged in chronologic order and placed clockwise from Giotto to Venus Express for Solar System missions and counter clockwise from Hipparcos to *Planck* and similarly for the Horizon 2000+ component.

2.2.10 Conclusion

The summary above leads us to conclude that overall, Horizon 2000 has been a great success. What it promised was delivered and more. It brought to an end the policy of permanent re-competition that characterised the previous ESA system and impeded ambitious long-term strategic planning. With the spectacular success of Rosetta, following that of *Herschel* and *Planck*, XMM-*Newton*, Cluster, SOHO, Mars and Venus Express, HST, Integral and *Cassini*, Horizon 2000 has fulfilled its goals. It has allowed Europe, through ESA, to launch the then largest ever space telescopes for IR and X-ray astronomy, to achieve the most distant landing ever accomplished on an extraterrestrial object and the first landing on a comet nucleus and to fly the first mission to the Moon using solar electric propulsion. With SOHO, it has created a true revolution in solar observations, and with Cluster it has performed the first multilocation studies of the Earth's magnetosphere.

Horizon 2000 has also given European scientists leading roles in ultra-high-resolution astrometry, X-ray and FIR astronomy, cosmology, solar physics and Solar System exploration including comets, Mars, Venus, Saturn/Titan and soon Jupiter in the framework of Cosmic Vision, its prolongation. Thanks to the programme, the European space science community is now considered as an essential and powerful partner on the worldwide scene. Looking back to the ESRO times, it has transformed ESA into a leading space agency, offering collaborative opportunities on its missions rather than seeking the hospitality of other agencies. Most essential in this success has been the unflinching support of all partners involved and the continuous work of ESA to maintain the continuity and integrity of the programme. An illustration is offered by the fate of two missions that, without Horizon 2000, would most probably not have been launched yet.

Before Horizon 2000 came into being, Cluster was in competition with SOHO. Combining both missions into one Cornerstone serving more than 30% of the space science community cemented the support of the complementary 70%. Without Horizon 2000, Cluster II would never have emerged after the loss of Cluster I in 1996. Furthermore, in the course of the formulation of Horizon 2000+, some of the members of the Survey Committee, while reviewing the achievement of the plan 10 years after its start, questioned the validity of keeping *Herschel* in the programme bearing in mind the forthcoming development of NGST (even though the scientific capabilities of the two missions were quite different and complementary). ESA was steadfast in not allowing this, and *Herschel* completed its in-flight mission nominally in 2013, 4 years after its launch, while JWST is still on the ground with (at the time of writing) a possible launch in 2018. Without the Horizon 2000 strategy, *Herschel* would probably still be on the drawing board in the best case. These considerations confirm the value of the approach followed by ESA in the formulation of its long-term plans, an approach now continued in the implementation of Cosmic Vision, the successor of Horizon 2000 and 2000+, through the selection of two large-mission concepts (an X-ray observatory and a gravitational wave interferometer) as ‘pillar’ in its long-term plan.

Twenty years after the birth of ESRO, following a long and difficult road on which technical and political difficulties abounded, ESA’s Horizon 2000 programme has contributed spectacularly to making Europe a world leader in IR and submillimetre astronomy thanks to three of its flagship missions, ISO, *Planck* and *Herschel*. The *Herschel* story as described in the following chapters offers a rich example of how to maintain that prominent and responsible position for the next generations of space astronomers.

References

Bernard J-P, Ade P, André Y, Aumont J (2016) PILOT: a balloon-borne experiment to measure the polarized FIR emission of dust grains in the interstellar medium. Exp Astron 42(2):199–227

Bleeker JAM, Bonnet R-M (1984) Space Science Horizon 2000. ESA SP-1070:1–137

Cavallo G (1996) The selection of the Ariane-5 test flight for cluster’s launch. ESA Bulletin 87

Delannoy J, Lacroix J, Blum EJ (1973) Proc IEEE 61:1282–1284

Drapatz S et al (1981) The 1m balloon-borne telescope “GOLDEN DRAGON” for far-infrared astronomy. Mitteilungen der Astronomischen Gesellschaft 54:91; 1981 MitAG.54...91D

Enger C et al (2009) Status of the PILOT Balloon–Borne experiment. In: 19th ESA symposium on European rocket and balloon programmes and related research, ESA SP-671

Harwit M (2001) The early days of infrared space astronomy. In: Bleeker JAM, Geiss J, Huber MCE (eds) The century of space science. Kluwer Academic, Dordrecht, pp 301–330

Jamieson JA (1995) Ballistic misssile defense and infrared technology upon the Russian revolution and the Soviet Empire. Proc IRIS, Infrared Instrumentation Society 40(1):13

Krige J, Russo A (2000) A history of the European Space Agency 1958–1987. ESA SP-1235

Lamarre JM et al (1990) French activity in infrared astronomy from stratospheric balloons. Proc SPIE 1341, Infrared Technology XVI:183
Lamarre JM et al (1998) First measurement of the submillimeter Sunyaev-Zeldovich effect. ApJ 507:L5–L8
Lequeux J (2009) Early infrared astronomy. J Astron Hist Heritage 12(2):125–140
Low F et al (1976) An observational study of the AFCRL infrared sky survey. ApJ 206:L153–L155
Minier V, Olmi L, Lagage P-O, Spinoglio L et al (2008) Submm/FIR astronomy in Antarctica: potential for a large telescope. 2nd ARENA conference, Potsdam. EAS Publications Series 33:21–40
Poulter G, Jennings R (1983) A far-infrared Fabry-Pérot interferometer and grating spectrometer for Balloon-borne astronomy. Infrared Phys 23(1):43–50. doi:10.1016/0020-0891(83)90065-9
Price SD, Walker RG (1976) The AFGL four color infrared sky survey: catalog of observations at 4.2, 19.8, and 27.4 μm. AFGL-TR-76-0208. ADA034448
Price SD (1977) The AFGL four color infrared sky survey: supplemental catalog. AFGL-TR-77-0160. ADA048048
Rieke GH (2009) History of infrared telescopes and astronomy. In: 400 Years of astronomical telescopes. Springer, Berlin
Rogalski A (2012) History of infrared detectors. Opto-Electr Rev 20(03):279–308
Rouhan D (2013) OSIRIS et Astroplane, Les premières Observations Infrarouges en France. L'Astronomie 127:1
Serra G, Puget JL, Ryter CE, Wijnbergen JJ (1978) The far-infrared emission of interstellar matter between galactic longitudes l=36 deg and l= 55 deg. ApJ 222:L21–L25
Southwood D (2014) Space Science and Policy. A&G 55(6):6.26–6.32. doi:10.1093/astrogeo/atu249
Stegner K et al (1997) The PIROG program. In: Proceedings of the 13th ESA symposium on rockets and balloon programmes and related research, ESA SP-397
Trendelenburg EA (1971) Infrared detection techniques for space research. In: Manno V, Ring J (eds) Proceedings of the fifth ESLAB/ESRIN symposium. D.Reidel, Dordrecht, pp 1–2
van den Abeelen, L (2017) Space plane HERMES, Europe's dream of independent manned space flight. Springer ed, Hilversum, The Netherlands. ISBN:978-3-319-444472-7
Vanhabost G, Gigan P, Hammal K et al (1977) Space Sci Instrum 3:343
Wegmann HM et al (1978) ASSESS-II: a simulated mission of Spacelab. Nature 275:15–19. doi:10.1038/275015a0

Chapter 3
Herschel Mission Overview

Abstract Here we describe the historical evolution of the FIRST/*Herschel* Cornerstone within the framework of ESA's Horizon 2000 plan, from the initial concept proposed by the scientific community in November 1982 through the various studies, the industrial development, the launch and operations and the post-operations phase which ended in mid-2017. It illustrates the effect of the design-to-cost approach on the technical and managerial developments of *Herschel* over the more than three decades of the project. The chapter details the long series of challenges faced by the scientists, ESA and its member states and industry in the face of several crises and of many technological, managerial and political difficulties. Eventually these were overcome through innovation in technology and technical and financial management, preserving the scientific excellence of the mission. The chapter also emphasises the consequence of the combination of the *Herschel* and *Planck* missions within a single industrial project, launched together, on the overall final cost, somewhat lower than the total allocation set in the original Horizon 2000 plan for these two missions.

3.1 Introduction

The preceding chapter has described the historical and political/strategic context which made it possible to introduce into the scientific programme of ESA one of the most ambitious astronomy missions ever undertaken. FIRST, later renamed *Herschel*, was one of the four Cornerstones of the Horizon 2000 programme (ESA 1984) extending over two decades, initiated in Rome in 1985 by the unanimous decision of the ESA Council meeting at ministerial level (cf. Sect. 2.2.5). In Rome, the ESA ministers also agreed on the overall programme budget. As the main philosophy of Horizon 2000 was based on three essential principles, 'scientific excellence, design-to-cost and schedule respect', every element of the programme had to be developed within fixed cost limits (cf. Sect. 2.2.6). This approach, though restrictive, was key to granting each of the Cornerstones solid programmatic security as well as sustained support by the ESA executive, offering the benefit of initiating early enough the necessary innovative technological developments. However, the chronology of the implementation of the missions was the responsibility of ESA's science director. The STSP and high-throughput X-ray Cornerstones (CS),

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_3

the technological readiness of which was more advanced, logically granted them earlier launch opportunities, and they became, respectively, the first (CS1-SoHO-Cluster) and second (CS2-XMM-*Newton*) Cornerstones, thereby placing FIRST and the Comet Nucleus Sample Return missions in competition for the CS3 position (cf. Sect. 2.2.8). Section 3.3.4 describes the process, which led to *Herschel* occupying the CS4 position.

This chapter intends, from the vantage point of the ESA executive, to show how the design-to-cost approach influenced the evolution of the design and management of *Herschel* over 26 years, from the initial mission concept submitted to ESA in 1982 to its launch in 2009. This approach led to overcome an impressive series of challenges, thanks to the innovation in technology and technical and financial management to reach the highest scientific level in its field. It vividly illustrates the remarkable attitude of all partners involved: scientists, member states, industry and ESA management, in finding solutions to keep the scientific objectives and capabilities of the mission at the level of a Horizon 2000 Cornerstone.

3.2 FIRST in the Pre- and Post-Horizon 2000 Period (1974–1986)

Many space missions have had a visionary scientific leader—or a small group of scientists—at the origin of their conception. FIRST is no exception to that rule. In 1974, the year when the European Space Agency came officially in existence, Thijs de Graauw (one of the co-authors of this book) participated in a study group held at the ESA European Space Research and Technology Centre (ESTEC, the Netherlands) where he made an impression on Dr. Brian Fitton, head of the Surface Physics Division (later renamed as the Astronomy Division). Fitton offered him a staff position at ESTEC, which he took up after finishing his PhD on heterodyne detection in the infrared. This work brought him in contact with the submillimetre astronomical spectroscopy community, mostly found at that time in the USA at the National Radio Astronomy Observatory (NRAO) in Green Bank and Kitt Peak, at NASA and the Jet Propulsion Laboratory (JPL). Critically, he was involved, as the only European, in the work of the 'The Submillimeter Space Telescope Working Group' formed at JPL in May 1977, which concluded in their report issued in early 1978, that '*The submillimeter spectral band between wavelengths of 100 μm and 1 mm is uniquely valuable for the investigation of a wide range of astronomical subjects*'.

In 1982, still at ESTEC, he was well aware that ESA would soon release a call for mission proposals. He then took the initiative of organizing under the auspices of ESA a meeting in Noordwijkerhout, near ESTEC, in May 1982 (Fig. 3.1) with the objectives of reviewing the scientific importance of submillimetre observations and the unique science to be addressed and the future steps to implement a space mission dedicated to this yet unexplored wavelength range. The conclusions of the

Fig. 3.1 Group picture in De Keukenhof from the ESA Workshop held in Noordwijkerhout, 10–12 May 1982. © *ESA*

workshop led the scientific community to work on the preparation of a submillimetre space mission concept. ESA issued a call for mission proposal in July 1982, and in November of that year, Gisbert Winnewisser, from Cologne, submitted, on behalf of a wide international community of radio, submillimetre and infrared astronomers, a proposal for a Far-Infrared and Submillimeter Space Telescope (FIRST).

The proposal was based on a deployable 8 m-class telescope radiation-cooled to 150 K, operating in the range 100 μm–1 mm, providing a 3.5 arc-sec beam, and with a sensitivity of 10 mJy[1] at the shortest wavelength. This spectral range represented the last major window of the electromagnetic spectrum to be explored after the Infrared Astronomical Satellite (IRAS) launched in 1983 and the Infrared Space Observatory (ISO), selected by ESA the same year and launched in 1995, both operating in the infrared range (IRAS ~8–120 μm and ISO ~2–240 μm). The proposed payload of FIRST (ESA 1983) consisted of an imaging multiband photometer for the range 100–800 μm, a high-resolution (10^4) spectrometer for the range 100–200 μm and a very high-resolution (10^6) super heterodyne instrument package operating between 300 and 600 μm. The first two instruments would be cryogenically cooled in a liquid helium/hydrogen cryostat ensuring a 2-year lifetime. Stirling cycle refrigerators providing a 5-year lifetime would cool the heterodyne mixers and preamplifiers. The combination of such a package together with a large aperture telescope would enable observation of the uncovered submillimetre wavelength range with unprecedented sensitivity and spatial and spectral resolution.

[1]The Jansky (or Jy) is a unit of spectral flux density or spectral irradiance, equivalent to 10^{-26} watts per square meter per Hertz.

Fig. 3.2 Cover of the FIRST assessment study report. © *ESA*

The proposal was eventually selected competitively by the ESA advisory structure for an internal assessment study, carried out in 1982–1983. The composition of the FIRST assessment study team is given in Appendix 1. The study resulted in a report (ESA 1983) (Fig. 3.2) highlighting FIRST as a space observatory based on a self-deployable 8 m Cassegrain telescope, with a surface error of 8 μm rms, coupled to the proposed three-instrument payload for spectroscopy and photometry, cooled with a combination of a cryostat and active Stirling cycle refrigerators. The satellite would be launched on an Ariane 4 rocket from Kourou into a highly eccentric orbit

(HEO) of the ISO type. It was considered at that time that the various new technologies required for FIRST, for instance, in the areas of heterodyne instrumentation, cryostat design (to be based on the ISO cryostat), lightweight structures and the Ariane 4 launcher capabilities, would be available in time for a launch in the 1994 time frame.

Notwithstanding the novel scientific objectives of FIRST and its ability to further expand and capitalise on the expected results of IRAS and ISO, to the great dismay of the proposers, the mission was not considered by ESA for further study. However, the discussions held in the ESA advisory structure highlighted the growing importance of this emerging field of space astronomy and opened the way to include a High Throughput Heterodyne Spectroscopy mission in the Horizon 2000 plan (ESA 1984) as described in Chap. 2 and in the following textbox quoting Johan Bleeker, Chairman of the Survey Committee, that established the architecture of Horizon 2000.

Quoting Johan Bleeker, Chairman of the Horizon 2000 Survey Committee

During the deliberations of the Survey Committee meeting in Venice in June 1984, I raised the point, shared by other members of the Committee, that major next-generation space astronomy missions would have to accommodate high-resolution spectroscopy as a mandatory feature, since they ought to deliver the detailed fingerprints of the astrophysical processes at work in the particular wave band concerned, thereby requiring high-throughput instruments to yield a significant spectral flux for unambiguous interpretation. This was true of both the X-ray and submillimetre spectral domains, leading to a certain coherence and symmetry in the final Horizon 2000 structure (Fig. 2.5) where the two astronomy Cornerstones eventually adopt nearly the same title. In its Sect. 12, the Horizon 2000 report outlines the importance of spectroscopy in 'Infrared, Sub-millimeter and Radio Astronomy ... a wavelength range encompassing spectroscopic studies of atoms and molecules of galactic and extragalactic objects, covering specific details of star formation in individual molecular clouds, the determination of the present rate of expansion of the Universe and the studies of the solar system (planets, comets and so on)'. The report also lists the required development of enabling instrument technologies up-front, including heterodyne mixer technology, yielding far superior spectral resolution as compared to noncoherent techniques. This unique capability and its unprecedented diagnostic power was, at least for me, the reason to explicitly label this in the cornerstone designation. Of course, over time, people have expressed different views (preferences) on the main thrusts of FIRST/*Herschel*. Nevertheless, in my perception high-resolution spectroscopy was pivotal for FIRST (and its reincarnation *Herschel*) as conceived right from the beginning in H 2000. The results from HIFI speak for themselves on this issue.

ESA organised dedicated workshops to review and discuss the scientific and technical requirements of each Cornerstone mission (Table 2.3). The workshop 'Space-Borne Sub-Millimetre Astronomy Mission' (ESA 1986) devoted to the High Throughput Heterodyne Spectroscopy mission took place on 4–7 June 1986 in Segovia (Spain). The outcome called for FIRST to represent the objectives of the Cornerstone based on an 8 m-class telescope (f/10, surface error 8–10 μm rms[2]) deployable in orbit with a payload complement consisting of a combination of cryogenically cooled direct detection and heterodyne instruments for high-throughput spectroscopy and photometry of the cold universe in the range 50 μm–1 mm.

During 1986, mainly for internal ESA political reasons to promote the use of the International Space Station (ISS), FIRST was also studied as a passenger of a co-orbiting platform to the ISS in the context of the Columbus Programme, ESA's participation in the ISS (see Chap. 2). In the course of the study, modifications occurred in the co-orbiting platform and a free-flyer configuration in low Earth orbit (LEO) was then envisaged. The study was not conclusive because the LEO appeared to be unsuitable as it provided a low observing efficiency and imposed very severe thermal constraints to the extent that in-orbit refuelling of as much as 6000 litres of liquid He (three times as much as the ISO cryostat capacity) would be needed to keep the mission operational over its 3-year nominal lifetime (Aerospatiale 1987). Consequently, the ISS-related scenario was abandoned and all further studies conducted from 1987 onwards concentrated on HEO scenarios and were carried out exclusively in the context of a free-flying mission independent of the ISS and under the full responsibility of the ESA science directorate.

NASA Submillimetre LDR Studies

Studies of a mission related to the science addressed by FIRST but much more ambitious, the Large Deployable Reflector (LDR), had their origin in 1977 with two parallel proposals from NASA Ames Research Center (ARC) and Jet Propulsion Laboratory (JPL). The proposals were for Space Shuttle-deployed telescopes with 10–30 m diameter apertures at ambient temperature for the submillimetre band. A joint NASA ARC/JPL study team was formed in 1979, and a first feasibility study was positively concluded by the Lockheed Missiles & Space Company as reported to NASA in 1980 (NASA 1980). From 1982, the LDR was studied extensively by NASA-JPL (NASA 1985). A study report was released in 1985 and various workshops were devoted to LDR, e.g. LDR workshop (NASA 1983) and Asilomar III LDR workshop (NASA 1988). Typically, it consisted of a 20 m aperture intended to observe in the 30 μm–1 mm range but required a significant number of technologies that did not exist. LDR was supposed to be launched

(continued)

[2]See Sect. 6.1 on Telescope principle.

by the Space Shuttle and deployed and released by astronauts. NASA eventually abandoned this concept at end of the 1980s following the 1986 Challenger accident and the uncertainties of the ISS programme.

3.3 The Race to the CS3 H 2000 Position: From Community Wishes to SPC Adoption

3.3.1 Mission Definition Study (1987–1990)

Following the 1986 Segovia meeting, a Science Advisory Group (SAG) was established in 1987, under the chairmanship of Reinhard Genzel from the Max Planck Institute for Extraterrestrial Physics (MPE) in Germany, with the task of defining the scope of the mission by trading off scientific requirements against technical complexity and the rigid 400 million euros (1984 EC) Cornerstone cost constraint. The SAG was also charged with outlining the necessary planning of activities compatible with the selection of the third Cornerstone (CS3), foreseen to take place in 1991 for a launch planned in 2003. The SAG was supported by a Payload Working Group (PWG) and a Telescope Working Group (TWG) and later by a Tiger Team composed of scientists with specific expertise in the issues at stake. The composition of these groups is given in Appendix 2.

The mission definition activities started along two major lines: a nominal mission based on the 8 m-class concept which had emerged from the 1986 Segovia workshop and less ambitious smaller concepts based on trading off telescope size and mission duration to provide a backup at lower cost. The analysis of these concepts led to the selection of a 24-hours HEO orbit. The concepts were extensively studied via an industrial System Definition Study (SDS) in the course of 1990–1991, with the aim of defining a mission that could be realised within the 400 million euros (1984 EC) limit.

It was soon realised, however, that the nominal mission would not meet these cost constraints, so that international cooperation would also have to be considered. To this end, initial discussions with NASA took place in November 1989 aimed at exchanging information between the two agencies on their respective far-infrared/submillimetre programmes (see textbox on the Large Deployable Reflector). Although the final scientific goals of both agencies were almost identical, the programmes leading to these goals did not appear to be compatible. Nevertheless, the idea of cooperating looked attractive to both agencies and actually, several other possible ways of cooperation were pursued in the course of the SDS study.

3.3.2 System Definition Study (1990–1992)

After the SAG had provided support to ESA in the evaluation of the industrial bids, the SDS contract was awarded to Dornier Systems (ESA 1989). The study was conducted in three phases during the period 1990–1991. An initial 2-month phase reviewed the previous studies (8-m deployable and 4-m class fixed telescopes and related technologies from the ESA).

The second phase, lasting 7 months, was aimed at trading-off various solutions:

- Telescope size and technology
- Cooling concept: liquid He cryostat or cryocoolers
- Focal plane layout: full payload or infrared only payload or heterodyne-only payload
- Service module

Two baseline system designs were studied. One was for an autonomous ESA mission with an overall cost within the Cornerstone budget allocation. The second was for an extended mission (e.g. larger telescope) relying on cooperation with other agencies. The third phase, lasting 8 months, was devoted to the detailed system design, including cost and schedule analyses of one of these baselines. The baseline to be selected at the end of Phase 2 for detailed study during Phase 3 would depend on whether or not a possible partner could be found for the mission.

Phase 2 was completed with a presentation of the results in December 1990. On the basis of technical risks, total programme cost and scientific priorities, the SAG recommended that the baseline mission for Phase 3 would include the following elements:

- A non-deployable 4.5-m telescope with a passive Carbon Fibre Reinforced Plastic (CFRP) reflector, diffraction limited at 150 μm, and for which two concepts were considered: a multi-panel dish studied by Matra and a single-panel concept as backup, led by Dornier
- A single He II cryostat
- A payload complement with both heterodyne and direct detection instruments
- A dual launch by Ariane 5 into a geostationary transfer orbit

A dedicated symposium ‘From Ground-Based to Space-Borne Submillimeter Astronomy’ was held in Liège (Belgium) on 3–5 July 1990 (ESA 1990a) to discuss and review the mission concept and the science objectives arising from Phase 2, which included a payload consisting of two high-resolution heterodyne spectrometers and two direct detection instruments for photometry and spectroscopy. It was also presented at the 24th ESLAB meeting (ESA 1990b) held on 17–19 September 1990 in Friedrichshafen (Germany) but was abandoned because of cost considerations.

Phase 3 started in January 1991, based on an ESA-only mission and the elements mentioned above, but still constrained by the ‘golden rule’ that the total programme cost should meet the Cornerstone budget limit. In that respect, the SAG agreed to trade off mission lifetime and the instrument complement if needed. In case the emerging mission would not satisfy the cost constraints or the expected scientific capabilities, further reconsideration of mission objectives would have to take place. The results of Phase 3 were presented to the SAG and the PWG/TWG on 5 November 1991, based on the Liège symposium recommendations: a passively cooled 4.5-m non-deployable telescope, diffraction limited at 150 μm, with a payload complement consisting of two heterodyne spectrometers supplemented by one direct detection spectrometer/photometer covering selected bands in the range 100 μm–1 mm. The payload was accommodated in a cryostat containing 3200 l of superfluid He to ensure a working temperature of 2 K and allowing a lifetime of approximately 3 years. The meeting concluded that the major technical and scientific objectives of the mission could be achieved, but for reasons of cost, the spacecraft configuration was constrained to be compatible with a dual launch with Ariane 5.

An essential element of the mission was the telescope primary reflector. Two passively cooled 4.5-m non-deployable reflector concepts, diffraction limited at 150 μm, were studied based on the Matra multiple panel concept as baseline and on the Dornier single-panel option as a backup, both showing that the predictable distortions (gravity, thermal, moisture) were not critical, even though the final surface accuracy was mainly driven by unpredictable error sources (e.g. non-isotropic behaviour of materials, manufacturing technology, vibration-induced displacements). Eventually, the preliminary results confirmed the feasibility of manufacturing a telescope meeting the required performance, and further studies continued throughout 1992.

Unfortunately, the resulting cost of the baseline mission was still too high: two independent estimates were derived which were respectively between 52 and 110 million euros above the cornerstone limit of 520 million euros (1990 EC). The first was based on the Dornier SDS study, while the second resulted from an ESA in-house evaluation based on a comparison with the ISO mission and included an extra 30 million euros for possible further preparatory studies before entering full industrial development called phase B (ESA 1992a) (see textbox on Mission Phase Cycle). The situation was discussed with the SAG in mid-March 1992. Cost drivers were identified, which the SAG used to define a reduced concept entailing a drastic effort to arrive at the allowable cost ceiling and whose scientific performance would be compared with other projects under consideration in the field (ground-based telescopes, Antarctica, balloons and other space missions). To decrease mass, the main reflector diameter was reduced to a 3-m dish with limited pointing requirements to avoid new developments of costly high-accuracy star trackers, and the He cryostat was deleted, implying that the direct detection instruments were removed from the payload, leaving the heterodyne-only part relying solely on mechanical coolers. However, the SAG also assessed the implication of a payload that could still include some direct detection instrumentation. This activity was completed in June 1992, and the descoped mission served as input to a rider study (ESA 1992b) to be completed by mid-1993 in time for the selection of the next Cornerstone at the end of the year.

Mission Phase Cycle
Usually, space missions' development goes through five main phases:

- Pre-Phase A, a conceptual/feasibility study
- Phase A, a preliminary design confirming proof of concept
- Phase B, the definition phase converting the preliminary plan into a baseline technical solution
- Phase C/D, the detailed design followed by the implementation phase, usually committed to the industry
- Phase E covers the launch and operations of the spacecraft once in orbit

3.3.3 Rider Study (1992–1993)

The rider study, carried out by Dornier Systems supported by the Tiger Team set up by the SAG, started in November 1992 together with trade off studies of various spacecraft configurations with the objective of saving mass. The period February–May 1993 was devoted to the detailed design of a proposed concept based on a heterodyne-only mission (HOM) with a 3-m telescope and a system of mechanical coolers carrying a heterodyne-only payload with three instruments limited to 1 THz as maximum frequency. The HOM was eventually rejected by the community because it was not compatible with the scientific capabilities of a Cornerstone, i.e. a large mission of highly competitive scientific content, providing unique innovative science and representing an area of expertise and excellence of the European space science community. The payload complement was consequently reconsidered so that it could accommodate any type of instrument (heterodyne or direct detection), provided it would not drive the system design more than would a mission with a 1 THz heterodyne instrument. With this proviso, the rider study was completed in mid-June 1993 featuring a mission concept based on:

1. A passively cooled 3-m Cassegrain monolithic telescope with a wave front error better than 11 μm rms and an effective area of at least 5 m^2
2. A cooling system of closed-cycle Stirling refrigerators providing a 4 K environment for the science payload and a lifetime of at least two years
3. A payload consisting of two instruments:

   a. An imaging direct detection instrument, the Far Infrared (FIR) spectrometer-photometer, using photoconductor and bolometer detector arrays for high-resolution spectroscopy (3×10^4) in the range 85–210 μm, medium-resolution spectroscopy (3×10^3) up to 300 μm and photometry in the range 85–900 μm

b. A multichannel coherent instrument, the multi-frequency heterodyne (MFH) receiver for very high-resolution spectroscopy with superconductor-insulator-superconductor (SIS) and Schottky mixers covering the range 490–1130 GHz in five selected bands

4. A launch in 2003, as upper passenger of a shared Ariane 5, into an ISO-type low-inclination 24-h HEO orbit

The revised concept was reduced significantly from the initial mission scenario, but the cost estimate was still in excess of the Horizon 2000 Cornerstone allocation (708 million euros compared to 625 million euros at 1993 EC) (ESA 1993a, 1994a). In fact, the same problem was also affecting the competing CNSR Cornerstone, Rosetta, placing both in more or less equal positions in the competition to become CS3. However, since the decision on the selection of the next Cornerstone was scheduled for the SPC meeting of November 1993, the SAG, supported by the Tiger Team and the Telescope and the Payload Working Groups, decided to proceed on the basis of the revised mission concept resulting from the rider study and fully described in the study report, the so-called Red Report, issued in September 1993 (ESA 1993b), some 10 years after the first assessment study of the original mission proposal.

3.3.4 Selection of FIRST as Cornerstone 4 of Horizon 2000 (1993)

The results of the preparatory studies of the two competing Cornerstone missions were presented to the ESA advisory structure in June 1993. It was recognised that both missions fulfilled the scientific requirements of a Cornerstone. Nevertheless, on the basis of programmatic considerations and technological readiness, Rosetta was recommended to be CS 3, forcing FIRST to become CS4. It was emphasised that under no circumstances should the scientific capabilities of FIRST be reduced further. On the positive side, it was noted that the extended development time for FIRST should be used opportunely for further qualification of critical technologies for the telescope, the cooling system and the payload, prior to the release of the Announcement of Opportunity to select the scientific instruments. In addition, it was recommended that, for both missions, all avenues be explored with regard to international cooperation in order to reduce the cost without jeopardising the science objectives. At its meeting of 4–5 November 1993, the SPC unanimously endorsed the recommendation to carry out Rosetta as CS3 before FIRST (ESA 1993c), implying that FIRST would launch at the end of 2006. Figure 3.5 (see Sect. 3.6.2) shows the evolution of the FIRST/*Herschel* launch date from inclusion of the mission as CS 4 to the launch in May 2009.

3.4 FIRST in the New Horizon 2000 Plan (1993–1997)

3.4.1 The Traumatic 1994–1997 Series of Crises

As explained in Chap. 2 (cf. Sect. 2.2.9), following the recommendation of the ESA Council of 1992 and in order to make room for the possible continuation of the Horizon 2000 programme beyond its completion assumed to happen in 2005, ESA initiated a complete review of its implementation. At its meeting of June 1994, the SPC approved the ESA proposal to reschedule the implementation of FIRST before the fourth medium mission (M4), thus bringing the launch of FIRST from late 2006 to end 2005/early 2006 (ESA 1994e) (Fig. 2.7). This decision reinstated the 3-year interval between M3 and M4 (Fig. 2.6) implying an early start of the FIRST pre-Phase B activities in line with Phase B starting in early 2000 instead of mid-2001 and with the following milestones:

- Mission Definition Freeze: late 1996
- Release of the Announcement of Opportunity (AO) for the payload instruments: late 1996
- Selection of payload instruments: late 1997
- Completion of the technology studies: mid-1998
- Issue of the Invitation to Tender (ITT) to industry for Phase B: early 1999

In addition to the severe programmatic and financial problems resulting from the 1995 Toulouse Ministerial meeting, which replaced a continuous 5% yearly increase of the Science budget over 10 years by a 3% decrease over an unknown period, other missions' specific issues emerged at the same time (cf. Sect. 2.2.9):

- In November 1996, the Russian Mars-96 mission, involving extensive payload participation by a number of ESA Member States, was the victim of a launch failure. Under pressure from the Mars community and the Member states involved in Mars-96, the decision was made to implement the ESA Mars Express mission, making reuse of the Rosetta platform and elements of the Mars-96 payload, a financially attractive option.
- In the course of 1996, NASA approached ESA to participate in their New Generation Space Telescope (NGST), later renamed *James Webb* Space Telescope (JWST), the successor of the *Hubble* Space Telescope (HST), a proposal which was immediately supported by the astronomy community, not hesitating to sacrifice, if necessary, some of the missions foreseen in Horizon 2000 and Horizon 2000+.
- In July 1996, the Cobras-Samba mission devoted to the investigation of the cosmic microwave background (CMB) radiation (later more elegantly renamed *Planck*) was selected by the SPC as the third medium mission (M3) of Horizon 2000 (ESA 1996a) and confirmed in November (ESA 1996c).
- In early 1997, the SPC made the decision to fly Cluster II in 2000 to compensate for the loss of the original mission after 30 seconds of flight, making it possible

to implement fully the complete STP Cornerstone of Horizon 2000 consisting of the SoHO and Cluster missions (ESA 1997a).

These successive crises and their associated financial and schedule problems, and also the occurrence of new and exciting scientific opportunities, necessarily implied a drastic revision of the implementation of the Horizon 2000 programme.

3.4.2 Managing the Crises (1994–1997)

As a consequence of the revision of the plan, FIRST and *Planck* underwent a significant cost reduction exercise. As both were cryogenic missions, sharing similar operational orbits around the Lagrange point L2, the possibility emerged to fly both of them together on a single Ariane 5 thereby reducing launch costs. A second cost saving might result from combining the two payloads on a common platform within one industrial procurement proposal from a single prime contractor for the two combined projects. Such a solution would also offer the advantage of using a single and smaller ESA project team intended to provide a further 10% reduction in the overall cost of the merged option. An ESA internal study, supported by a FIRST/*Planck* Tiger Team of external scientists involved in both missions, (team composition given in Appendix 3) did not identify any technical or scientific showstoppers to the merging. Encouraged by the substantial economies this option could provide, ESA went even further and fixed a target of saving the total cost of the *Planck* M3 mission, thus implementing the merged missions at the total cost of a Cornerstone minus 10%, corresponding to a financial envelope of 632 million euros (1996 EC) (cf. Sect. 3.8 which tracks the evolution of the cost at completion (CaC) of *Herschel* and *Planck*).

In mid-April 1997, '*The Far infrared and Submillimeter Universe*' symposium (ESA 1997b) devoted to the science of FIRST was held in Grenoble. A large audience including US scientists and NASA representatives attended. The FIRST/*Planck* merger received a rather unenthusiastic response. Immediately after the symposium, the ESA advisory bodies met to discuss the FIRST/*Planck* situation and eventually agreed to study both the merger option and the effects of the targeted savings, as long as the main scientific objectives of both missions could be preserved. They also recommended that the combined mission be launched not later than the end of 2005 or the beginning of 2006, in view of avoiding a possible loss of competitiveness of *Planck*, which was then confronted by strong competition from other space missions and ground-based instruments, and to favour international collaboration on FIRST, a prospect carrying some appeal after NASA (JPL) had offered to provide a CFRP telescope for FIRST (cf. Sect. 3.4.4.2).

In June 1997, the SPC approved the principle of implementing the combined FIRST/*Planck* mission in a single programme to launch in 2005, within the cost target of 632 million euros (1996 EC), assuming that the main scientific objectives of both missions would be preserved (ESA 1997c). However, the SPC also urged

ESA to study an alternative, lower cost, dedicated *Planck* mission (sometimes referred to as *Planck* Express in allusion to the fast and cheap Mars Express mission) in the event that the merger solution did not prove to be feasible. In fact, some SPC delegations indicated that a launch date at the end of 2005 for the combined FIRST/*Planck* mission would create financial difficulties for a timely provision of the payloads and requested a minimum 3-year interval after the launch of Rosetta, then planned for 2003. The launch date for the baseline scenario of the combined FIRST/*Planck* mission was then delayed to mid-2006 (see Sect. 3.6.2). At the same time, the same delegations claimed that there would be no problems with a launch of *Planck* in 2004 and of FIRST in 2007.

The revised plan, eventually approved by the SPC in November 1997 (ESA 1997d), included the concept of the merger of FIRST and *Planck* within a single Cornerstone financial envelope reduced by 10% with respect to the original Horizon 2000 targets and a launch by Ariane 5 in mid-2006. Such a drastic cost reduction required innovative approaches to technological and managerial concepts as well as the search for new opportunities for international cooperation. In this context and as a result of related discussions, NASA agreed to investigate providing hardware items to ESA. Similarly, ways of cooperation with Member States' space agencies were explored in parallel, albeit with little success.

*3.4.3 Selection of the Combined FIRST/*Planck* Option (1997–1998)*

In July 1997, ESA initiated technical studies of the combined FIRST/*Planck* mission, involving industry and a science team from members of both missions. The industrial effort consisted in two competitive contracts awarded to Matra Marconi Space (MMS) and Aerospatiale. An alternative option referred to as the 'carrier', was also proposed by industry. It consisted of two spacecraft mated together and launched by a single Ariane 5. In parallel, ESA conducted an internal study of a separate FIRST and of a cheaper *Planck*, referred to as the 'stand-alone' option. The aim was to obtain, in spring 1998, all the technical, scientific and programmatic information necessary to decide on the final concept meeting the cost target agreed by SPC in June 1997 while preserving the main scientific objectives of both missions.

In summary, three possible options were considered:

1. The 'merged' option consisting of one single spacecraft with two payloads launched by an Ariane 5 with direct injection into a transfer trajectory to a Lissajous orbit around L2,[3] with interleaved scientific observations for a total of 3 years for FIRST and 15 months (two sky surveys) for *Planck*.

[3]In orbital mechanics, a Lissajous orbit, named after Jules Antoine Lissajous, is a quasi-periodic orbital trajectory that an object can follow around a Lagrange point of a three-body system without requiring any propulsion.

2. The ‘carrier’ option consisting of two spacecraft mated together and launched by a single Ariane 5 with direct injection into a trajectory towards L2 and early separation for parallel scientific observations (3 years FIRST and 15 months *Planck*) from two different Lissajous orbits.
3. The ‘stand-alone’ option based on two spacecraft launched separately 1 year apart with a Soyuz-Fregat launcher for *Planck* and an Ariane 5 for FIRST with direct injection into a transfer orbit to L2. Scientific observations would span 15 months for *Planck* and 3 years for FIRST.

The three options were discussed with potential principal investigators (PIs) and co-PIs as well as by the advisory bodies prior to the May 1998 SPC meeting and were found to be compatible but with varying levels of satisfaction and confidence, with the essential scientific requirements of FIRST and *Planck*. In particular, the five instrument PI teams (three FIRST and two *Planck*) were extremely unenthusiastic about the merger, regarding it as a non-optimal and risky option designed purely to limit costs at the expense of scientific reliability and flexibility. It was clear that the ‘carrier’ was the most attractive option from an overall scientific and technical standpoint and the PIs urged ESA to find ways to implement that scenario. In addition, several factors emerged which enabled substantial cost savings wholly or partly applicable to the three options:

- A cheaper launch vehicle for the *Planck* ‘stand-alone’ option and a shared launch for the other options
- A single industrial prime contractor and industrial consortium for the procurement of both spacecraft
- A single ESA project team to manage both projects
- A common ground segment and a single operations team
- Common and reduced schedules and reuse of development equipment procured from the same companies for the two projects

The cost estimates of the various options were (all expressed in 1997 EC) 690 million euros for the ‘merged’ option, 757 million euros for the ‘carrier’ and 837 million euros for the ‘stand-alone’, to be compared with the original estimates of 700 million euros for FIRST and 381 million euros for *Planck*. However, the cheapest 690 million euros option was still in excess of the 654 million euros target set in the updated Horizons 2000 implementation plan.

In May 1998, the SPC did not know yet what would be its new level of resources and the purchasing power of the science programme to be approved by the Ministerial Council meeting to take place later in December. Nevertheless, ESA requested the Committee to approve the FIRST/*Planck* mission on the basis of the merged option within the financial target of 654 million euros (1997 EC) with a launch in mid-2006, realising that only the merged option might be compatible with this envelope. However, after a thorough and intense debate, the delegations concluded that the merged option would be too complex for both spacecraft and payload development and would entail much more risk than the carrier option. Accordingly, they recommended implementation of the carrier option, despite its higher cost, assuming that it should be decreased in the course of development. To compensate for the extra cost, ESA then proposed to delay the launch date from mid-2006 to the first quarter of 2007 while maintaining a target cost of 654 million euros (1997 EC). The SPC unanimously agreed the proposed solution.[4]

3.4.4 Optimization Studies of Mission Critical Elements (1994–2000)

The version of FIRST thus selected was based on the outcome of the industrial studies performed in 1992–1993, prior to the CS4 decision. The main critical elements of the spacecraft included a 3-m diameter Cassegrain telescope inside a sunshade and a payload module (PLM) requiring an improved cryogenic subsystem involving Stirling cycle coolers, not yet fully developed, to provide a 4 K environment to the scientific instruments and guarantee a lifetime of at least 2 years. As the advanced detectors were to operate at even lower temperatures in the 0.1–4 K range, further cooling would need to be provided by the instruments themselves. All these new cryogenic developments would have to be qualified against environmental conditions and scientific specifications by the time the payload would be selected. A launch in mid-2006 required a start of Phase B no later than mid-2001, implying that the considerable amount of study and development work required for all FIRST critical elements had to be completed in time for the release in September 2000 of the Invitation To Tender (ITT) for prime contractor selection.

ESA set up a technology development programme (ESA 1994b, c), including qualification and test of the large telescope and the Stirling cooler cryogenic subsystem, with qualification of the 20 K and 4 K coolers (Matra Marconi Space-MMS), a study to assess the implications of using ISO cryostat technologies (Daimler-Benz Aerospace-DASA) and a feasibility study of large arrays of detectors (ESA 1994d).

3.4.4.1 Spacecraft and Cryogenic System(s) Developments (1994–1996)

In the period 1994–1996, two different spacecraft concepts were studied based on the one hand on the design of the Stirling cryocoolers system, taking into account their increased technical maturity and performance (see details in Chaps. 8 and 9) and, on the other hand, a cryostat concept based on the then well-proven ISO cryostat technology. The presentation of the final results took place at ESTEC on 27 June 1995 and confirmed that an ISO-like cryostat was a viable alternative to the Stirling cooler system. This conclusion led to trade-off studies between the cryocooler and the cryostat options in the period mid-1995 to mid-1996, based on

[4]This paragraph includes detailed information from a private note provided by A. Linssen whom we acknowledge with thanks.

a 3-m telescope and on the science inputs provided by the SAG and the associated working groups. The final results were presented at ESTEC on 25–26 June 1996, and the study report was made available in October 1996, giving priority to the ISO-like cryostat option.

In the fight against cost increases, the reuse of the XMM platform for FIRST was also studied by ESA for both options (ESA 1996b). Both studies considered the use of 24- and 48-hour HEO and Lissajous orbits around L2. The final presentation of these studies took place in November 1996 at Friedrichshafen. As no clear advantages could be identified in favour of reusing the XMM platform, in particular because of the long time separating the respective developments schedules of XMM and FIRST, the idea was abandoned.

3.4.4.2 Telescope Studies (1995–2001)[5]

In mid-1995, a contract was placed with MMS for the development and qualification of the FIRST 3-m telescope, with Dornier as subcontractor for the primary reflector dish made of gold-coated CFRP (Carbon Fibre Reinforced Plastic) (ESA 1995). A number of difficulties were encountered, which delayed the development and qualification work by more than a year and implied an increase of the cost at completion by a factor of two. In this context, in November 1996, it was decided to abandon the CFRP technology considered as incompatible with the specifications and to look more closely to other alternatives and in particular to silicon carbide type materials, which had also been studied by MMS and Dornier within the same contract. In parallel, ESA had placed a small technology contract with Aerospatiale and Neyrpic Framatome Mécanique (NFM) to investigate the feasibility of using an all-aluminium telescope with active thermal control (ESA 1997e). The preliminary results of this study led to the release of an ITT for the manufacturing of a 3-m aluminium telescope, eventually assigned to Aerospatiale (ESA 1997f). Continuing work on the silicon carbide (SiC) technology by MMS was nevertheless kept as a backup. In the course of 1997, a revision of the payload scientific performance furthermore suggested to increase the diameter of the main reflector of the FIRST Telescope to 3.5 m (ESA 1997g).

Opportunely, in 1997, an understanding was reached between ESA and NASA for a potential cooperation whereby NASA would supply a CFRP telescope in exchange for a reasonable representation of US scientists in the future Time Allocation Committee. ESA initially required a 3.0–3.5-m telescope, but NASA studied a 3.5–3.8-m telescope for some time (1998–1999). The larger telescope was considered as an option for ESA but NASA eventually settled to 3.5 m by early 2000. The CFRP material for the US telescope was developed using a different process (high-temperature curing) than the early ESA development and had already been applied in a similar smaller scale for the NASA Microwave Limb Sounder

[5]cf. Chap. 6 for technical details.

mirror. The important cost saving resulting from the cooperation with NASA led ESA in May 1997 to baseline this solution for the FIRST telescope.

On the other hand, the mass of a 3.5-m aluminium telescope, approximately 200 kg heavier than the NASA/JPL baseline or the SiC alternative, proved to be incompatible with the Ariane 5 capabilities. In addition, since the machining demonstration of an aluminium test article failed and also because only one potential supplier could be identified for the procurement of a 3.5-m aluminium blank, this technology was eventually abandoned.

A NASA/JPL feasibility study[6] started with Composite Optics Inc. (COI), San Diego, in early 1998, for the design, manufacturing and testing of a subscale 2-m diameter demonstrator. By mid-2000, it became clear that the ultimate goals for the performance of the telescope could not be achieved with the proposed technology. Furthermore, at the end of 2000, due also to budgetary reasons, JPL was forced to stop all activities on the work, while waiting for approval of a budget extension. In mid-March 2001, NASA officially informed ESA that they were unable to deliver the telescope due to tight budget constraints, as well as development problems and cost overruns faced in their contributions to the FIRST and *Planck* scientific instruments.

Fortunately, the SiC technology studies continued by ESA in parallel eventually proved to be compatible with the requirements of a 3.5-m telescope, i.e. high image quality, very low distortion, stability over a wide temperature range (90–350 K) and low mass. A 1.35-m SiC demonstrator mirror was developed and successfully tested (ESA 1997h). The mechanical integrity of the mirror was confirmed following representative vibration and acoustic noise tests followed by optical verifications performed between April 1999 and April 2000. It was therefore decided to manufacture a full-size segment of the final 3.5-m diameter SiC primary mirror, to be completed in May 2001 (ESA 1999a). This successful development work led to the release of a contract awarded to MMS (renamed Astrium SAS, France) and its subcontractors Boostec (France) and Opteon (Finland), for the delivery of a qualified flight model and flight spare elements by mid-2005 (ESA 2001a). Figure 3.3 illustrates the evolution of the telescope primary mirror diameter over the entire development of *Herschel*.

[6]In the early stage of the study, the possibility of increasing the size of the primary to 3.8 m was investigated but soon abandoned as the diameter had to remain constrained to 3.5 m imposed by the fairing of the launcher.

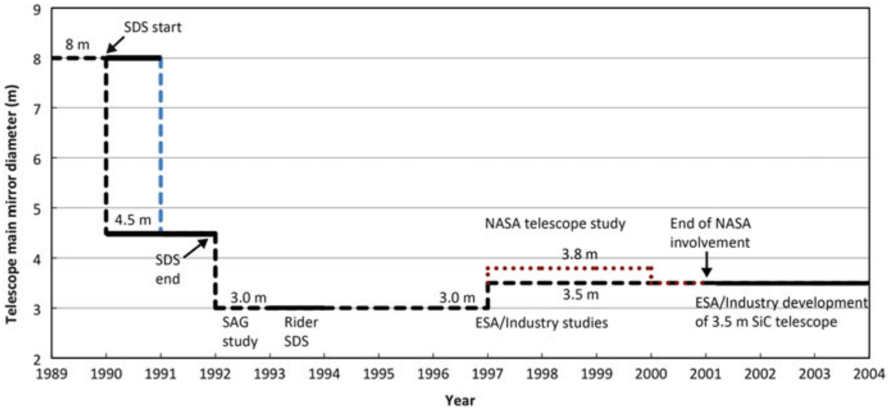


Fig. 3.3 Evolution of the FIRST/*Herschel* telescope primary mirror diameter from the 8 m-class original proposal to the 3.5-m flight model. © *ISSI (S. Volonte) & CEA (V. Minier)*

3.5 Selecting the FIRST Scientific Payload (1997–1999)

*3.5.1 The Announcement of Opportunity for the Combined FIRST/*Planck *Mission*

After the SPC had approved the principle of combining the FIRST/*Planck* mission in a single programme with a single launch by Ariane 5, the respective Science Management Plans (SMP) describing the various elements of the scientific exploitation of FIRST (ESA 1997i), *Planck* (ESA 1997j) and the combined FIRST/*Planck* (ESA 1997k) missions were approved by SPC in the period between June and September 1997. Since, except for the merged concept, the two missions were completely independent once in orbit, most of the elements described in their respective Science Management Plans were preserved. The integrated Science Management Plan clarified three issues relevant to the merged option: the composition and functions of a common FIRST/*Planck* Science Team, the coordinated project development and implementation plan and the coordinated schedule of operations, noting that in case of conflicts with either the FIRST or *Planck* Science Management Plans, the integrated SMP would take precedence.

Based on a fully merged mission concept (not yet approved by the SPC), the Announcements of Opportunity (AO) for the FIRST and *Planck* payloads were released in early October 1997 with proposals due by mid-February 1998. The worldwide scientific community was invited to submit proposals for either payload, and, in the case of FIRST, proposals were also invited for Mission Scientist (MS) positions. The role of the MS would be to independently monitor the development of the instruments with emphasis on their complementarity; provide independent advice on science operations, instrument calibration, operational modes and data reduction software; and review, advise and assist on optimising the scientific core

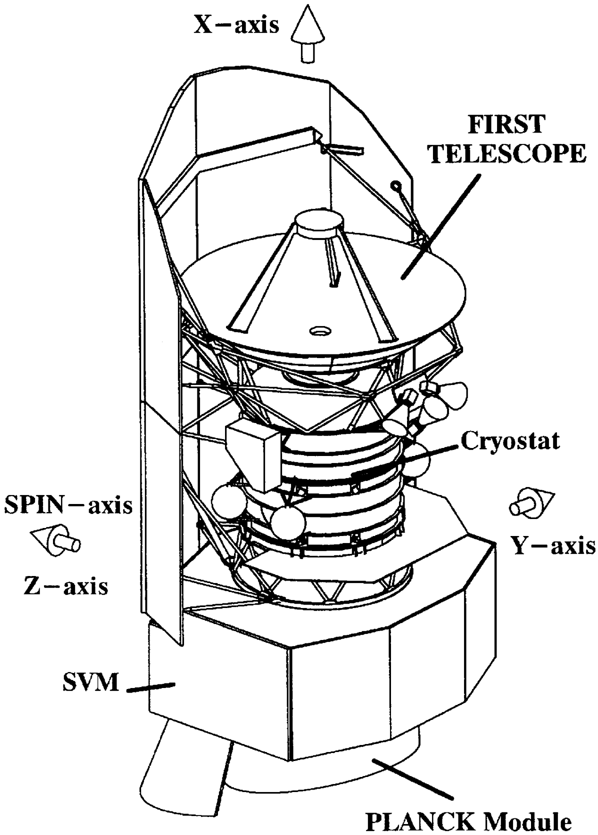


Fig. 3.4 The merged FIRST/*Planck* spacecraft mechanical configuration as shown in the FIRST/*Planck* Announcement of Opportunity (AO). The satellite has a total height of 9.3 m, a maximum width of 4.3 m and a launch mass of 4.5 tons. The *Herschel* 3.5-mm diameter CFRP telescope is protected by an open sunshade/sunshield structure, and the payload is housed inside the superfluid helium cryostat required to provide a lifetime of 4.5 years in the L2 orbit. The lifetime includes commissioning and 15 months of *Planck* operations. © *ESA*

observing programme based on the scientific objectives of the mission. The selection by the SPC would take place in two steps:

1. Preselection of instruments and PIs in May 1998, pending approval by SPC of the final mission option
2. Confirmation in February 1999, on the basis of the approved final mission option, after optimization of the preselected proposals and clarification of funding agreements with the relevant national agencies

At the end of an extensive scientific and technical evaluation of the trade-off study results, including risk, choice of orbit and cost, carried out by ESA, and taking account of the SAG recommendation, the decision was made to implement FIRST with a cryostat spacecraft in orbit around L2. The AO was based on a 3.5-m telescope assembly inside a fixed sunshade, a payload module with the cryogenically cooled focal plane science instruments and a service module accommodating the payload electronics at ambient temperature (cf. Fig. 3.4).

The model payload consisted of three scientific instruments, each also having a dedicated Instrument Control Centre (ICC), to be funded by the national space agencies and scientific institutes of the proposing instrument consortia:

1. A heterodyne instrument for high to very high-resolution spectroscopy in the 250–600 μm range
2. A photoconductor instrument for imaging, spectroscopy and photometry in the 85–200 μm range
3. A bolometer instrument for imaging photometry in the 200–600 μm range and for spectroscopy in the 200–350 μm range

The role of the ICCs would be to insure the successful operation of the respective instruments by developing and maintaining instrument observing modes and providing specialised software and procedures for the processing of the data generated. They would be part of the science ground segment, operating in partnership with the (FIRST) *Herschel* Science Centre (HSC) funded by ESA, the sole interface between the mission and the investigators in the science community.

The ground segment is fully described in Chap. 4, Fig. 4.5.

*3.5.2 Selection of the FIRST/*Planck *Scientific Payload*

After scientific evaluation of the proposals by two separate peer-review committees, in May 1998, the SPC endorsed the selection of the FIRST Mission Scientists (ESA 1998a) and preselected the payloads of FIRST and *Planck* (ESA 1998b) compatible with the ‘carrier’ option. These were subsequently confirmed by the SPC in February 1999 (ESA 1999b) for a joint launch in the first quarter of 2007. In the meantime, the PIs were intensively involved in securing adequate funding for their instruments by their respective national space agencies and negotiating whenever necessary rearrangements of distribution of tasks within the consortia.

The three instruments with their respective ICCs selected for FIRST were:

- The heterodyne instrument for FIRST (HIFI) for very high-resolution spectra of astronomical objects in seven frequency bands (PI: Thijs de Graauw, succeeded in 2008 by F. Helmich)
- The Photoconductor Array Camera and Spectrometer (PACS) for infrared imaging and photometry and low-resolution spectroscopy up to about 200 μm, based on two arrays of stressed Ge-Ga photoconductors with 16 × 25 detectors (PI: Albrecht Poglitsch, succeeded in 2014 by E. Sturm)
- The Spectral and Photometric Imaging Receiver (SPIRE), for photometry and Fourier transform spectroscopy at wavelengths longer than 200 μm using 0.3 K bolometer arrays (PI: Matt Griffin)

Table 3.1 gives a more detailed description of these three instruments in launch configuration.

Table 3.1 *Herschel* main characteristics at launch (Pilbratt et al. 2010)

Spacecraft	
Size height/width	7.4/4.0 m
Mass wet (incl. helium)/dry	3400 (335)/2800 kg
Incl. telescope/science instruments	315/426 kg
Power total/science instruments	1200/506 W
Science data rate (max. average)	130 kbps
Solar aspect angle (wrt tel. boresight)	60°–110° (120°)
Absolute pointing (68%)	2″
Telescope	
Primary physical/effective diameter	3.5/3.28 m
Secondary diameter	30.8 cm
System/primary f-number	8.70/0.5
Wave front error best-focus (centre/edge)	4.8/5.5 μm
Angular resolution	$7'' \times (\lambda_{obs}/100\ \mu m)$
Operational temperature	85 K
Science instrument main characteristics	
HIFI	
Heterodyne spectrometer	
Wavelength coverage	157–212 and 240–625 μm
Field-of-view (FOV)	Single pixel on sky
Detectors	5 × 2 SIS and 2 × 2 HEB mixers
Spectrometers	Auto-correlator and acousto-optical
Spectral resolving power	Typically 10^6
PACS	
Two-band imaging photometer	
Wavelength coverage	60–85 or 85–130 and 130–210 μm
Field-of-view (FOV)	$0.5F\lambda$ sampled $1.75' \times 3.5'$
Detectors	64 × 32 and 32 × 16 pixel bolometer arrays
Integral field spectrometer	
Wavelength coverage	55–210 μm
Field-of-view (FOV)	(5 × 5 pixel) $47'' \times 47''$
Detectors	Two 25 × 16 pixel Ge:Ga arrays
Spectral resolving power	1000–4000
SPIRE	
Three-band imaging photometer	
Wavelength bands ($\lambda/\Delta\lambda \sim 3$)	250, 350, 500 μm
Field-of-view (FOV)	$2F\lambda$ sampled $4' \times 8'$
Detectors	139, 88 and 43 pixel NTD bolometer arrays
Imaging Fourier transform spectrometer	
Wavelength coverage	194–324 and 316–671 μm
Field-of-view (FOV)	$2F\lambda$ sampled circular $2.6'$
Detectors	37 and 19 pixel NTD bolometer arrays
Spectral resolving power	370–1300 (high)/20–60 (low)

Note: Acronyms relating to the detectors: superconductor-insulator-superconductor (SIS), hot electron bolometer (HEB), gallium-doped germanium (Ge:Ga), and neutron transmutation doped (NTD)

The selected HIFI and SPIRE instruments included US scientific groups contributing enabling instrument technology sponsored by NASA. In addition, NASA was also responsible for the NASA HSC, providing additional support and funding of science exploitation for investigators based in the USA (Pilbratt et al. 2010).

As indicated above, a number of mission scientists were also selected to become members of the *Herschel* Science Team (HST). In addition, a *Herschel* Optical System Scientist position was added in 2001 (ESA 2001a, cf. Sect. 3.6.1).

The HST composition is given in Appendix 4.

3.6 Project Development (2000–2009)

3.6.1 Preparation of the FIRST/Planck Industrial Project Development (2000–2001)

A system-optimisation study of the carrier option, incorporating the selected payloads and the NASA-provided 3.5-m FIRST telescope (still the baseline at the time), was carried out in early 2000 by Alcatel at system-level supported by Astrium for the payload modules (PLM). The final presentations took place at ESTEC in June 2000. This activity led to the release of the Invitation To Tender (ITT) to industry in September 2000 requesting industry to set up a single development programme for the delivery of FIRST and *Planck* for a launch in early 2007 (ESA 2000). Two proposals were received by ESA at the end of 2000, one from Alcatel and the other from Astrium. In both, the development of the PLM was assigned to Astrium.

In December 2000, 'The Promise of FIRST' symposium (ESA 2001b) devoted to infrared space missions, with emphasis on FIRST, was held in Toledo (Spain). The carrier concept with the selected payload was presented to the scientific community and the scientific objectives of the mission discussed in the context of past and future infrared missions. On this occasion, following a proposal made by G. Pilbratt, the project scientist, FIRST was renamed the '*Herschel* Space Observatory' to mark the bicentennial anniversary of the discovery of the infrared radiation by the British astronomer William Herschel.

In March 2001, ESA's Industrial Policy Committee approved the *Herschel*/*Planck* implementation phase contract proposal eventually awarded to Thales Alenia Space France (TAS-F, formerly Alcatel Space Industries) as prime contractor for the overall programme development and for the PLM integration and test of *Planck* with a core team including TAS-Italy, formerly Alenia Spazio, for the *Herschel* and *Planck* service modules, and Astrium-Germany for the payload module, integration and test of the newly named *Herschel* satellite (ESA 2001c). At the same time, NASA informed ESA officially that they were unable to deliver the *Herschel* telescope (cf. Sect. 3.4.4.2), and in order to keep more direct control of the tight schedule, ESA took over the responsibility for the procurement of the 3.5-m diameter SiC telescope, awarded to Astrium-SAS (ESA 2001d).

In May 2001, the SPC approved the appointment of a *Herschel* Optical System Scientist in charge of providing independent support to ESA on all aspects of the optical systems of *Herschel* with emphasis on the design and performance of the end-to-end optical systems, keeping in sight the scientific goals of the mission (ESA 2001a).

3.6.2 Herschel/Planck Last Race Towards Launch (2001–2009)

The following paragraphs illustrate the hectic race in which all *Herschel* and *Planck* partners were involved with the aim of fulfilling as much as possible the science objectives, the schedule and the CaC of both missions.

April 2001: The *Herschel/Planck* industrial development formally starts, involving the industrial core team. The Preliminary Design Review (PDR) is completed in October 2002. By the end of 2003, the industrial consortium is built up and about 140 subcontractors are selected (ESA 2004a).

May 2002: At its meeting in Andenes, Norway, the SPC approves the selection of the *Eddington* mission for the detection and study of exoplanets as an element of the new Cosmic Vision plan (ESA 2002a), using a copy of the *Herschel* service module to be implemented in the context of the *Herschel/Planck* project (ESA 2002b).

May 2003: Since the beginning of the *Herschel/Planck* development programme, ESA had to assume responsibility for a number of tasks and deliveries such as the *Herschel* telescope (cf. Sect. 3.6.1) not included within the financial envelope of the mission at the time of selection and confirmation of the *Herschel/Planck* payload in 1999. Furthermore, in order to secure the scientific objectives and the schedule of both missions ESA had to financially support elements of the payload for which, despite all efforts of the instrument PIs, adequate funding from the national agencies was not available. This included the establishment of a Central Parts Procurement Agency, providing the instruments with expensive (electrical) components. Eventually ESA took the decision of not asking for the reimbursement of the components which was a great relief to the instrument teams and helped maintain the schedule. Accordingly, in May 2003, the SPC approves the resulting additional costs amounting to 73.74 million euros (including the procurement of the *Herschel* telescope), thus bringing the cost at completion for the *Herschel/Planck/Eddington* group of missions to a new total of 1051.40 million euros (mixed 2003 EC) (ESA 2003a).

November 2003: The SPC discusses the proposal of the science directorate for rebuilding a financially affordable science programme within the budgetary constraints of the approved level of resources and decides to exclude *Eddington* from the *Herschel/Planck/Eddington* group, leading to the immediate cancellation of *Eddington* and thus reducing the *Herschel/Planck* CaC to 860.32 million euros (mixed 2004 EC) (ESA 2003b).

November 2004: The SPC is informed of the results of the *Herschel*/*Planck* Critical Design Review (CDR) conducted between July and September 2004, revealing that the planned launch date in early 2007 can no longer be achieved as a consequence of the late completion of the detailed system design and problems of interfaces with the scientific instruments, impacting on hardware and software deliveries. In addition, the development of the individual instruments for both missions encounters technical problems and schedule difficulties partly due to the lack of funding (fortunately rectified earlier by the SPC in May 2003). Consequently, the launch is postponed to the beginning of August 2007. All these problems imply a total increase of some 20% of the CaC approved by SPC in November 2003, corresponding to a new financial envelope of 1038.4 million euros at mixed 2004 EC (ESA 2004b). As the SPC is unable to approve the increase, the chairman Risto Pellinen requests ESA to set up an investigation group made of Member States' delegations, chaired by him, to analyse the reasons for the problems and report back in February 2005 (ESA 2005a).

February 2005: On the basis of information collected through interviews with industry, the PIs and members of the ESA science directorate involved in the approval of the mission, the investigation group recommends the increase of the CaC (ESA 2005b) to the revised envelope for *Herschel*/*Planck* of 1038.4 million euros (2004 EC) eventually approved by the SPC (ESA 2005c).

November 2005: The SPC is informed that the development of the instrument flight models is on the critical path with delivery dates 3–4 months late (ESA 2005d, e).

Throughout 2006, in spite of several appropriate recovery measures adopted by the PIs, the instrument flight model deliveries remain critical (ESA 2006a, b). Industry is tasked to optimise their development programme in order to minimise the impact of the late instrument availability. In November, the SPC is informed that a consistent satellite schedule has been established leading to a launch in May 2008 (ESA 2006c).

May 2007: The schedule of the project remains driven by the availability of the scientific instruments and problems in the spacecraft development programme. Although the *Planck* instruments have been delivered for integration to the spacecraft in November 2006, the *Herschel* PACS and SPIRE instruments are delivered in March 2007 and HIFI later in May. The launch slips to 31 July 2008 (ESA 2007). ESA states that any further delay beyond that date will induce a further increase in the predicted CaC previously adopted for the project.

February 2008: The SPC is informed that, despite intensive efforts throughout the year to keep the schedule and assuming that the programme will not encounter any major difficulties in the remaining tests, the earliest realistic launch date would be 31 October 2008 (ESA 2008a). The SPC is also informed that the associated financial impact is around 40 million euros, an amount that cannot be accommodated in the approved *Herschel*/*Planck* CaC. Consequently, because of the residual uncertainties in the launch date, there is no sense to ask for the approval of a new CaC at that time.

June 2008: ESA reports to the SPC that due to remaining uncertainties affecting the spacecraft assembly, integration and test activities, the launch can no longer be planned for 2008. Negotiations with Arianespace indicate that the earliest launch slot would open in January 2009 (ESA 2008b). At an expenditure rate of some 6 million euros per month and in consideration of the already stated 40 million euros additional cost, the CaC of *Herschel-Planck* will then increase by about 60 million euros altogether, and any further delay cannot be covered from the science programme contingency (ESA 2008c).

November 2008: The SPC is informed that *Herschel* and *Planck* are within final flight-acceptance testing. Unfortunately, due to liquid helium contamination problems affecting the *Herschel* cryostat during the summer, in addition to leaks on the HFI *Planck* instrument, the programme is again delayed. The schedule is updated accordingly and a new launch slot is agreed with Arianespace starting on 10 April 2009 (ESA 2008d). In the last quarter of 2008, after a successful launch-simulation of the filling and conditioning of the helium cryostat, *Herschel* undergoes a last environmental test including the Thermal Balance and Thermal Vacuum (TB/TV) test of the fully integrated spacecraft in the Large Space Simulation facility at ESTEC. The TB/TV test demonstrates that the launch transient behaviour of the complete cryogenic system follows the prediction very closely, clearing this major performance verification aspect. This test also demonstrates the functional performance and operation of the spacecraft and of the instruments under situations close to orbit thermal conditions. A science operations validation test is also carried out simulating 5 days of nominal operations in orbit as requested by the mission control centre in ESOC. The test activities are fully successful thus demonstrating readiness for launch. The spacecraft is returned to the clean room in the early days of January 2009 for a final functional test.

February 2009: *Herschel* and *Planck* are air transported to Kourou for a launch set for 16 April 2009 (ESA 2009a). Unfortunately, on site, the readiness of the ground segment accuses a 2-week delay, and a late iteration with the launcher on the *Herschel* flight readiness leads to an additional slip of 1 week. An additional problem with a component of the launcher upper stage further delays the launch by another week. All these delays result in an increase of 106.6 million euros of the previously approved financial envelope (ESA 2004b) with a CaC of the *Herschel*/*Planck* mission now totalling 1176.0 million euros, mixed 2009 EC (ESA 2009b).

14 May 2009: *Herschel* and *Planck* are launched successfully at 13:12 UTC on-board an Ariane 5-ECA from the Kourou Space Centre. Both spacecraft (first *Herschel* followed by *Planck*) are released within 30 minutes after lift-off, very precisely into their respective transfer orbit towards L2 (ESA 2009c). From then on, *Herschel* and *Planck* become two physically separated missions.

Figure 3.5 illustrates the time evolution of the FIRST/*Herschel* launch date from its inclusion in the Horizon 2000 plan in 1985.

The main characteristics of the *Herschel* mission at launch are summarised in Table 3.1 (from Pilbratt et al. 2010).

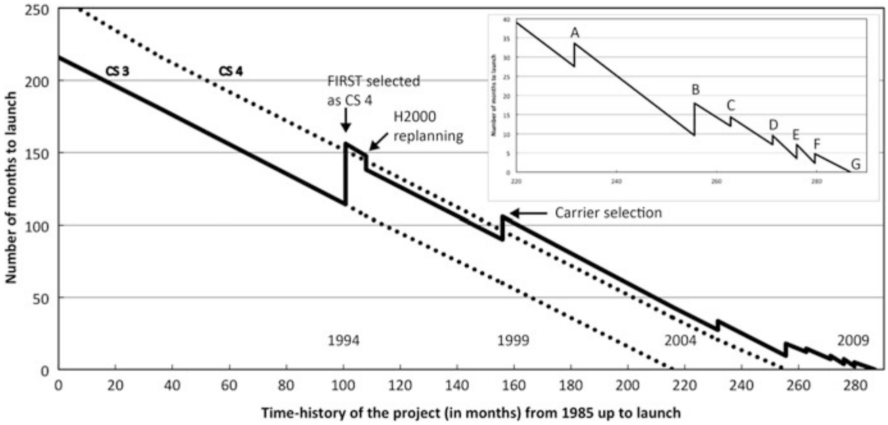


Fig. 3.5 Time-history of FIRST/*Herschel* from inclusion of the mission as a cornerstone of H 2000 in 1985 to the launch on 14 May 2009. The diagram illustrates the evolution of the planned launch date through the various stages of the studies and the project development. The figure also shows the launch dates originally planned in mid-2003 for CS3 and mid-2006 for CS4. In the blow-up extending from 2004 to 2009, A corresponds to the end of the Critical Design Review, B to a schedule slip due to foreseen late instrument delivery, C to a late delivery of the instrument flight models, D to uncertainties in the spacecraft tests, E to late assembly integration and tests and F to contamination problems in the *Herschel* cryostat. G corresponds to the launch on 14 May 2009. © *ISSI (Volonte) & CEA (V. Minier)*

3.7 *Herschel* In-orbit (2009–2013) and the Post-operations Phase (2013–2017)

After the initial 4 days of 'Launch and Early Operations Phase' (LEOP), the 'Commissioning Phase' (CoP) is initiated with the cooling down of the *Herschel* telescope to the predicted temperature below 120 K, permitting the cryocover protecting the cryostat from condensation of outgassed volatiles to be opened precisely 1 month after launch. The very first observations are then made (cf. Chap. 4, Sect. 4.3). The CoP phase is concluded with the successful In-Orbit Commissioning Review in July 2009. The overall responsibility for *Herschel* is handed over to the Science Operations Department at the European Space Astronomy Centre (ESAC), near Madrid. The Performance Verification Phase (PVP) is then started opening the way to the Routine Science Phase. By mid-October, 70% of the PVP is complete. The spacecraft is in very good health with the cryogenic chains working nominally and maintaining the bolometers of PACS and SPIRE at their operating temperatures of 0.3 K. The predicted cryostat lifetime is expected to slightly exceed the 3.5-year requirement. The ground segment is also working extremely well and providing high-quality data to the instrument teams. *Herschel* has already started performing some Science Demonstration and Routine Science Phase observations, with the data products being delivered to the users within 16 days (ESA 2009d).

In August 2009 the HIFI instrument suffers a significant failure subsequently traced to its prime electronic signal chain (cf. Chaps. 8 and 9). A recovery plan for operations is then formulated and a new observing programme of high-priority scientific objectives is redefined for HIFI in agreement with the *Herschel* Science Team (cf. Sect. 4.3).

In the course of January 2010, *Herschel* completes its Performance Verification and most of its Science Demonstration Phase observations. While SPIRE and PACS are performing mainly Routine Science Phase observations with most of their standard observing modes already released for public use, the HIFI instrument failure is eventually repaired by using its redundant electronic signal chain (ESA 2010a). In May 2010, almost 1 year into the mission, HIFI is again operational having been recommissioned and completed its Performance Verification Phase, and the instrument is delivering high-quality science data (cf. Sect. 4.3). The satellite is in good health, operating with full redundancy, except HIFI (ESA 2010b). After the recovery of HIFI, at the end of 2010, observations with all three instruments in most of their observing modes are routinely performed, and their science data products populate the *Herschel* Science Archive (HSA) for distribution to the observers as indeed was the case from the very beginning of the Science Demonstration Phase (SDP) (ESA 2010c). *Herschel* continues to perform nominally throughout 2011. A direct liquid helium content measurement indicates a likely end-of-life in spring 2013 (ESA 2011).

Throughout 2012, after successful improvements of the pointing performance to an accuracy of better than 0.9 arc-seconds, *Herschel* increases its observing efficiency to better than an average of 19 hours per day (the baseline was 18 hours) with a loss of observing time limited to 3% (ESA 2012a). An upgraded version of the HSA, with state-of-the art technology, is released in May, offering improved performance and a better user experience. All arrangements for the *Herschel* Post-Operations Readiness Review are put in place to assess the readiness of the *Herschel* Science Centre and Instrument Centres for the post-operations phase (ESA 2012b).

In February 2013, the spacecraft, the instruments and the ground segment continue to operate nominally. The science performance efficiency averages about 19.6 hours per day over the past 6 months. Helium exhaustion is foreseen for the second half of March 2013. In preparation for the post-operations phase, all *Herschel* centres successfully pass the Post-Operations Readiness Review (ESA 2013a). More than 99% of all priority observations (Key Programmes Guaranteed Time and Open Time observations) (cf. Sect. 4.2 for details) are completed (ESA 2013b).

On 29 April 2013, almost 4 years after launch, *Herschel* science observations come to an end when the helium cryogen runs out. Engineering tests are then performed to check the various subsystems of the spacecraft and the payload up to June. Since the *Herschel*'s orbit around L2 is unstable (as any such orbits are) requiring frequent maintenance operations (typically 4–6 times per year), a series of thruster burns are commanded, which move the spacecraft into a heliocentric orbit thereby avoiding all risk of collision with the Earth. On 17 June, a 'draining manoeuvre' is performed until fuel exhaustion followed by stabilisation of the spacecraft using the reaction wheels. The last command is uplinked to disable the downlink transponders, leaving *Herschel* into a safe 'disposal' orbit around the Sun where it will stay for centuries, predicted to return in 2027 to the vicinity of its former observation site, only to disappear again as a big and useless debris. A sad fate for such a glorious mission (ESA 2013c). It is worth remarking that there was a formal proposal submitted to ESA suggesting a disposal of *Herschel* by deliberately crashing it into the Moon. The impact, with an accurately known mass, velocity, trajectory and location, would have been observed by satellites in orbit around the Moon and also from observatories on Earth, providing useful scientific data on the lunar surface and subsurface. Sadly (from the point of those who wanted *Herschel* to go out in a blaze of glory and the planetary scientists who studied and made the scientific case), this option was rejected by ESA.

Post-operations phase activities focused on supporting the astronomical community exploiting *Herschel* data, including significant and progressive improvements of the interactive data processing software, successive upgrades of the HSA, repopulation of the archive with reprocessed data products using the latest reduction algorithms and calibration and ingestion of user-provided data products (ESA 2014a). Online data processing workshops brought improvements of processing and archive search capability to the astronomical users, with participants attending online and reacting very positively through the *Herschel* Science Centre Helpdesk (ESA 2014b). The last major version of the *Herschel* data processing software was rolled out in December 2015 (ESA 2016), and the HSA will be populated with the final pipeline (standard) data products in 2016.

Listings of point sources from the entire photometry sky coverages of PACS and SPIRE are being produced and will be made available online through archives like the *Herschel* Science Archive (HSA), the Infrared Science Archive (IRSA) and the Strasbourg Astronomical Data Centre (CDS). Other added value products are also in preparation concentrating on the building of the *Herschel* Explanatory Legacy Library (popularly known as HELL), the legacy documentation of the *Herschel* mission and on the collection and generation of added value science ready products, catalogues or interactively reduced data products. These enhanced data sets will eventually be ingested and served through the HSA with the intention to provide the best possible legacy products to the science community (ESA 2016).

By the end of August 2016, over 3400 scientists were registered as *Herschel* users. As of August 2016 a total of 1742 *Herschel* papers were recorded in the refereed literature (cf. Sects. 4.5.1 and 4.5.2). The *Herschel* and *Planck* project teams were the recipients of the American Institute of Aeronautics and Astronautics (AIAA) 2015 'Space Systems Award', 'for outstanding scientific achievements recognised by the worldwide scientific community and for outstanding technical performances of the two satellites (ESA 2015). Other awards included the 'Grand Prix 2010' award by the French Association for Aeronautics and Astronautics (AAAF) and the Group Achievement Award of the UK Royal Astronomical Society in 2014.

3.8 Evolution of the Cost Estimate Up to Completion of the Post-operations Phase

At this point it is of interest to consider the evolution of the cost estimate throughout the various stages of FIRST/*Herschel* from the System Definition Study to completion of the mission up to the end of the post-operations phase. This data is summarised in Table 3.2 and illustrated in the diagram of Fig. 3.6, taking into

Table 3.2 Evolution of the cost (million euros) of the various options of the *Herschel*/*Planck* mission (Bert Bastijns, private communication)

			Cost estimate (actual EC)	Backdating (1984 EC)
1985	H2000 approval. Cost envelope	CS mission		400
		M mission		200
1991	FIRST SDS study estimate	Industry	572 (1990 EC)	440
	(ESA 1992a)	ESA	630 (1990 EC)	485
		CS allocation	520 (1990EC)	400
1993	FIRST SDS rider estimate as	CS4	708 (1993 EC)	453
	(ESA 1993a)	CS allocation	625 (1993 EC)	400
1998	FIRST/*Planck* merger			
	(ESA 1998c)			
	Total cost target CS minus 10%		654 (1997 EC)	360
	FIRST/*Planck* on a common platform		690 (1997 EC)	380
	FIRST/*Planck* carrier option		757 (1997 EC)	417
	FIRST and *Planck* stand alone		837 (1997 EC)	461
2003	(a) *Herschel*/*Planck*/Eddington group		1051 (mixed 2003 EC)	493
	(ESA 2003a)			
	(b) Eddington cancelled → *Herschel*/*Planck*		860 (mixed 2004 EC)	400
	(ESA 2003b)			
2005	Various delays →*Herschel*/*Planck*		1038 (2004 EC)	470
	(ESA 2004b)			
2009	*Herschel*/*Planck* launch:		1176 (mixed 2009 EC	523
	(ESA 2009b)			
2017	*Herschel*/*Planck* to end of post operations		1190 (2016 EC)	526[a]
	(Bert Bastijns, private communication)			

[a]This cost for CS + M missions to be compared with cost target of CS mission minus 10%, i.e. 400–10% = 360 million euros (1984 EC) as agreed by SPC in 1998. It must be emphasised that the envelope for CS +M missions was set at 400 + 200 = 600 million euros in the original H 2000 plan (cost backdating to 1984 EC provided by Bert Bastijns)

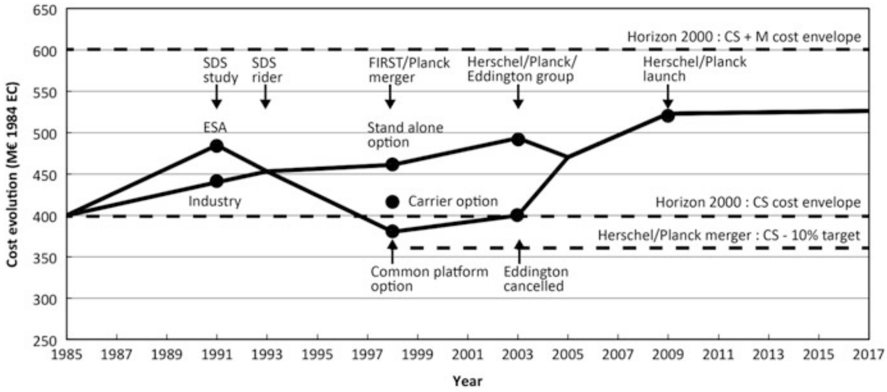


Fig. 3.6 Diagram of the cost evolution of the FIRST/*Herschel* mission showing the various options considered during the preparatory activities and the development phase. All cost estimates are backdated to 1984 EC for ease of comparison (cf. Table 3.2). © *ISSI (S. Volonte), ESA (B. Bastijns) & CEA (V. Minier)*

account the combination of the *Herschel* cornerstone with the *Planck* medium mission in a single industrial development project as decided in 1998. The cost estimates are given in actual economic conditions (EC) throughout the years. For ease of comparison, the costs are also backdated to 1984 EC, the time when Horizon 2000 was established. In 1984, the cost limit for a cornerstone (CS) was set at 400 million accounting units (MAU) equivalent to 400 million euros, and 200 million euros was the envelope for a medium (M) mission (cf. Sect. 2.2.6). Although the cost target agreed by SPC in 1998 (cf. Sect. 3.4.3) could not be fulfilled, the final cost given in Table 3.2 shows that the combination of *Herschel* and *Planck* within a single project dictated by programmatic and financial constraints led to the implementation of a Cornerstone and a medium mission within a total amount of 1190 million euros (2016 EC) up to the end of the post-operations phase. Backdated to 1984 EC this corresponds to a total cost of 526 million euros, some 13% (74 million euros 1984 EC → 210 million euros 2016 EC) *lower* than the total allocation of 600 million euros (1984 EC) for these two missions as set in the original Horizon 2000 plan (cf. Fig. 3.6).

3.9 Conclusion

Over 30 years separate the FIRST/*Herschel* proposal submitted to ESA in November 1982 and the completion of its spacecraft operations in June 2013 and the post-operations phase at the end of 2017. Those three decades and more witnessed a long series of crises; numerous technological, managerial and political difficulties; as well as great technical advances and were finally crowned with an outstanding scientific success. *Herschel* has delivered what it promised, and, more, thanks to a set of complex and advanced instruments with technology readiness that was not at all guaranteed when mission development started. Such a result is certainly one of the most positive and innovative effects of the Horizon 2000 concept based on the three main principles emphasised in Chap. 2: adoption by ESA of the design-to-cost approach for all elements of the mission under its responsibility but no compromise on scientific excellence and ensured security against mission cancellation. Abandoning any one of these principles would have led to the removal of the mission from the European space science programme at various stages in the development of the largest astronomical mission ever undertaken by ESA. As mentioned in Sect. 3.8, it is clear that the combination of *Herschel* and *Planck* within a single industrial project jointly launched by a single Ariane 5, allowed both missions to be carried out and indeed led to the implementation of a Cornerstone and a medium mission within a total cost some 13% lower than the total allocation set in the original Horizon 2000 plan for these two missions.

The following chapter offers a comprehensive summary of the scientific power of *Herschel* in exploiting an unexplored field of FIR and submillimetre imaging and spectroscopy.

Appendix 1

A.1.1 FIRST Assessment Study Team

J. Delanoy, Institut de Radio Astronomie Millimetrique (IRAM), St. Martin d'Héres (F)
R.D. Joseph, Imperial College London, London (UK)
E. Kollberg, Chalmers University of Technology, Göteborg (S)
G.F. Winnewisser, Physikalisches Institut, Cologne (D)

Assisted by

J. Baars, Bonn (D), D. Lemke, Heidelberg (D), S. Lidholm, Cork (EI), A. Moorwood, Garching (D), N. Panagia, Bologna (I), P. Salinari, Florence (I), B. Vowinkel, Cologne (D), M. Walmsley, Bonn (D), G. White, London (UK)

ESA/ESTEC personnel:

Th. de Graauw Study Scientist assisted by M. Kessler

Appendix 2

A.2.1 FIRST System Definition and Rider Study Teams

A.2.1.1 Science Advisory Group

R. Genzel (Chairman), MPI für extraterrestrische Physik, Garching (D)
E. van Dishoeck, Sterrenwacht Leiden, Leiden (NL) (from 1992)
D. Downes, Institut de Radio Astronomie Millimétrique, St. Martin d'Héres (F) (until 1991)
P. Encrenaz, Observatoire de Paris, Meudon (F)
R. Hills, Cavendish Laboratory, University of Cambridge, Cambridge (UK)
T. de Jong, SRON, Groningen (NL)
E. Kollberg, Chalmers University of Technology, Göteborg (S)
J. Martin-Pintado, Centro Astronomico de Yebes, Guadalajara (E)
V. Natale, CAISMI-CNR, Firenze (I)
J-L. Puget, Institut d'Astrophysique Spatiale, Université Paris XI, Orsay (F) (from 1992)
J.- P. Swings, Institut d'Astrophysique, Université de Liège, Liège (B) (until 1991)

A.2.1.2 Tiger Team

P. Encrenaz (Chairman), Observatoire de Paris, Meudon (F)
J. Baars, MPI für Radioastronomie, Bonn (D)
M. Griffin, Queen Mary & Westfield College, London (UK)
A. Harris, MPI für extraterrestrische Physik, Garching (D)
E. Kollberg, Chalmers University of Technology, Göteborg (S)
J-M. Lamarre, Institut d'Astrophysique Spatiale, Université Paris XI, Orsay (F)
N. Whyborn, Chalmers University of Technology, Göteborg (S)

A.2.1.3 Payload Working Group

I. Furniss, University College London, London (UK)
E. Kollberg, Chalmers University of Technology, Göteborg (S)
E. Kreysa, MPI für Radioastronomie, Bonn (D)
A. Poglitsch, MPI für extraterrestrische Physik, Garching (D)
G. Serra, CESR-CNRS/UPS, Toulouse (F)
G. White, Queen Mary & Westfield College, London (UK)
J. Wijnbergen, SRON, Groningen (D)
P. Zimrnermann, Physikalisches Institut, Universität zu Köln, Köln (D)

A.2.1.4 Telescope Working Group

J. Baars, MPI für Radioastronomie, Bonn (D)
J.-M. Lamarre, Institut d'Astrophysique Spatiale, Université Paris XI, Orsay (F)
D. Plathner, Institut de Radio Astronomie Millimétrique, St. Martin d'Héres (F)

A.2.1.5 ESA Personnel

U. Frisk/G. Pilbratt, Study Scientist (ESTEC) for SDS/RS, respectively
J.W. Cornelisse, Study Manager (ESTEC)
S. Volonte, Astronomy Missions Coordinator (HQ)

Appendix 3

A.3.1 FIRST/Planck

A.3.1.1 Tiger Team

C. Lawrence, Jet Propulsion Laboratory, Pasadena CA (USA)
B. Gray, Jet Propulsion Laboratory, Pasadena CA (USA)
R. Genzel, MPI für Extraterrestrische Physik, Garching (D)
A. Poglitsch, MPI für Extraterrestrische Physik, Garching (D)
H.U. Norgaard-Nielsen, Danish Space Research Institute, Copenhagen (DK)
J.M. Lamarre, Institut d'Astrophysique Spatiale, Université de Paris XI, Orsay (F)
J.L. Puget, Institut d'Astrophysique Spatiale, Université de Paris XI, Orsay (F)
M. Bersanelli, Universita' degli Studi di Milano, Milano (I)
N. Mandolesi, Istituto Tesre-CNR, Bologna (I)
M. Griffin, Queen Mary & Westfield College, University of London, London (UK)
R. Hills, Cavendish Laboratory, University of Cambridge, Cambridge (UK)
Th. de Graauw, SRON, Groningen (NL)

A.3.1.2 ESA/ESTEC Personnel

G. Pilbratt, FIRST Study Scientist
J. Tauber, *Planck* Study Scientist
J. Steinz, F. Felici and T. Passvogel, project team

Appendix 4

A.4.1 Herschel *Science Team*

A.4.1.1 Instrument Teams

Th. de Graauw, succeeded in 2008 by F. Helmich, SRON, Groningen (NL), HIFI PI
T.G. Phillips, California Institute of Technology, Pasadena CA (USA), HIFI Co-PI
A. Poglitsch, succeeded in 2014 by E.Sturm, MPI für Extraterrestrische Physik, Garching (D), PACS PI
C. Waelkens, Katholieke Universiteit Leuven, Leuven (B), PACS Co-PI
M. Griffin, initially at Queen Mary & Westfield College, University of London, and subsequently at Cardiff University (UK), SPIRE PI
L. Vigroux, Institut d'Astrophysique de Paris, Paris (F), SPIRE Co-PI

A.4.1.2 Mission Scientists

P.D. Bartel, Kapteyn Astronomical Institute, University of Groningen (NL)
J. Cernicharo, CSIC, Inst de Estructura de la Materia, Madrid (E)
P. Encrenaz, Observatoire de Paris, Meudon (F)
P.M. Harvey, Department of Astronomy, University of Texas, Austin TX (USA)
M. Harwit, 511 H Street, Washington DC (USA)

A.4.1.3 Optical System Scientist (from 2001)

J. Fischer (from 2001), Naval Research Laboratory, Washington DC (USA)

A.4.1.4 ESA Personnel

G. Pilbratt, Project Scientist, chairman
T. Passvogel, Project Manager, succeeded in 2002 by G. Crone
G. Crone, Payload Manager, succeeded in 2009 by J. R. Riedinger
J. R. Riedinger, Mission Manager, succeeded in 2012 by L. Metcalfe, succeeded in 2014 by P. García-Lario

In addition, personnel from the *Herschel* Science Centre, the NASA *Herschel* Science Center and NASA *Herschel* Project and Programme Scientists were invited to HST meetings for particular purposes.

The number of names quoted in the appendixes above is enormous, and they reflect the tremendous efforts provided by the scientific community to cooperate and help inventing one of the most impressive space missions ever invented in Europe. They could not be listed in the Acknowledgements at the beginning of the book, but the authors of the book admire and command all of them for their essential contributions to *Herschel*.

References

Aerospatiale (1987) SE/ES 03/87

ESA (1983) FIRST far infrared and submillimeter space telescope SCI (83) 1

ESA (1984) Space Science Horizon 2000, SP-1070

ESA (1986) In: Olthof H, Longdon N (eds) Proc. Space-borne sub-millimetre astronomy mission workshop: a cornerstone of the ESA long term space science programme, SP-260

ESA (1989) IPC (89)128

ESA (1990a) In: Kaldeich B (ed) From ground-based to space-borne sub-mm astronomy. Proceedings of 29th Liège international astrophysical colloquium, SP-314

ESA (1990b) In: Schwehm G, Stammes P, Battrick B (ed) Formation of stars and planets and the evolution of the solar system. Proceedings of 24th ESLAB symposium, SP-315

ESA (1992a) SPC (92)25

ESA (1992b) SPC (92)43

ESA (1993a) SPC (93)47

ESA (1993b) FIRST - far infra-red and submillimeter space telescope, SCI(93)6

ESA (1993c) SPC(93)40

ESA (1994a) SPC (94)39

ESA (1994b) IPC (94)1, add. 7, rev.1

ESA (1994c) IPC (94)152

ESA (1994d) IPC (94)1, add 19

ESA (1994e) SPC (94)24

ESA (1995) IPC (95)75

ESA (1996a) SPC/MIN/77

ESA (1996b) IPC (96)115, Annex 3

ESA (1996c) SPC/MIN/78

ESA (1997a) SPC/MIN/80

ESA (1997b) In: Wilson A, Rowan-Robinson M, Pilbratt G (eds) The far infrared and submillimeter universe, ESA SP-401

ESA (1997c) SPC (97)28

ESA (1997d) SPC (97)35

ESA (1997e) IPC (97)1, add.2

ESA (1997f) IPC (97)87

ESA (1997g) IPC (97)143

ESA (1997h) IPC (97)94

ESA (1997i) SPC (97)22, rev.1

ESA (1997j) SPC (97)8

ESA (1997k) SPC (97)34, rev.1

ESA (1998a) SPC (98)28

ESA (1998b) SPC (98)16

ESA (1998c) SPC (98)15

ESA (1999a) IPC (99)1, add.4

ESA (1999b) SPC (99)8

ESA (2000) IPC (2000)1, add.7

ESA (2001a) SPC (2001)23

ESA, (2001b) In: Pilbratt G, Cernicharo J, Heras AM, Prusti T, Harris RA, Bonnet RM (eds) The promise of the *Herschel* space observatory. Proceedings of ESA symposium SP-460, Toledo

ESA (2001c) IPC (2001)35
ESA (2001d) IPC (2001)1, add.19
ESA (2002a) SPC (2002)14
ESA (2002b) SPC (2002)26
ESA (2003a) SPC (2003)20
ESA (2003b) SPC/MIN/105
ESA (2004a) IPC (2004)23
ESA (2004b) SPC (2004)39, rev. 1
ESA (2005a) SPC/MIN/109
ESA (2005b) SPC/MIN/110 Annexe II
ESA (2005c) SPC/MIN/110
ESA (2005d) SPC (2005)23, rev. 1
ESA (2005e) SPC (2005)53
ESA (2006a) SPC (2006)10
ESA (2006b) SPC (2006)27
ESA (2006c) SPC (2006)38
ESA (2007) SPC (2007)24
ESA (2008a) SPC (2008)10
ESA (2008b) SPC (2008)23
ESA (2008c) SPC (2008)19, rev.1
ESA (2008d) SPC (2008)33
ESA (2009a) SPC (2009)8
ESA (2009b) SPC (2009)36
ESA (2009c) SPC (2009)17
ESA (2009d) SPC (2009)31
ESA (2010a) SPC (2010)5
ESA (2010b) SPC (2010)15
ESA (2010c) SPC (2010)28
ESA (2011) SPC (2011)48
ESA (2012a) SPC (2012)26
ESA (2012b) SPC (2012)36
ESA, (2013a) SPC (2013)9
ESA (2013b) SPC 2013)10
ESA (2013c) C (2013)93
ESA (2014a) SPC (2014)11
ESA (2014b) C (2014)93
ESA (2015) C (2015)114
ESA (2016) C (2016)10
NASA (1980) Large deployable reflector (LDR), Technical report CR 152402, W. H. Alff
NASA (1983) In: Leidich C, Pittman RB (eds) Proceedings of LDR workshop, CP 2275
NASA (1985) LDR system concept and technology definition study, NAS2–11861 and NAS2–11862
NASA (1988) In: Mahoney MJ (ed) Asilomar III LDR workshop, CR 184659
Pilbratt GL et al (2010) A&A 518, L1

Chapter 4
Herschel Science Evolution and Results

Abstract This chapter outlines the diagnostic importance of the far-infrared and submillimetre spectral range and presents the evolution of the mission's science objectives from the time of its original conception and goes on to describe the results and scientific achievements of *Herschel*. It shows how the opening of a spectral window, previously poorly studied, has changed our view of the most important components of our Universe, from galaxies and stars and the processes of their formation to planetary systems and the 'water trail'. The last section is a bibliometric analysis of the scientific return of *Herschel* and how it has interconnected the scientific community, illustrated by a new and original co-citation mapping. The quantity and quality of the data coming from *Herschel* during its almost 4-year observing lifetime will remain unique for a long time, casting new light on astronomical phenomena and processes never seen before.

Chapters 2 and 3 have analysed the history of the development of the *Herschel* Space Observatory and gave an overview of the conditions and constraints, which the FIRST Cornerstone of ESA's Horizon 2000 long-term plan was confronted to. Chapter 3 in particular offers an exhaustive description of the evolution of this large FIR and submm space observatory from the first idea proposed in 1982 to the launch in 2009. In the course of these 27 years, the technological and scientific context followed a considerable evolution through new advances in the manufacturing of large-space qualified telescopes and more performing detectors, with continuously improved characteristics, while new scientific discoveries were coming from the predecessors of *Herschel* such as IRAS launched in 1983, ISO in 1995, *Spitzer* (formerly called SIRTF) in 2003 and AKARI (formerly called ASTRO-F) in 2006 (see Table 2.2).

This chapter deals with the results and scientific achievements of *Herschel* and shows how the opening of a spectral window poorly studied before has changed our views of the properties and the evolution processes of the various objects, which populate our Universe, from galaxies through stars to planetary systems. The quantities and the qualities of the data coming from *Herschel's* observations during its more-than-nominal almost 4-year lifetime will remain unique for a long time and contribute to casting new light on astronomical phenomena and processes never seen before. The following sections offer a summary overview of the results obtained by *Herschel* and how they modify our understanding of the complexity

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_4

of our Universe and of its structures. The last section is a bibliometrics analysis of the scientific return of *Herschel* and how it has interconnected the scientific community.

4.1 The Diagnostic Importance of the FIR and Submillimetre Spectral Range

Approximately half of the electromagnetic energy released in the Universe today appears in the 8–1000 μm part of the spectral energy density (SED) function of the Universe (Fig. 4.1). It is referred to as the cosmic infrared background (CIB), and it peaks in the far infrared (FIR) at around 150 μm. The CIB is essentially due to the integrated star formation processes in the Universe, from the beginning until today. The infrared luminosity of a galaxy is linked to its star formation, and it is necessary to understand the formation and evolution of galaxies and thus of their star formation history over cosmic time to properly understand the origin of the CIB. Conversely, the observed value of the CIB provides an observational constraint on the history of the formation of stars and stellar systems.

The reason why the FIR/submm part of the spectrum is the repository of such a high proportion of the Universe's radiant energy is that much of the radiation

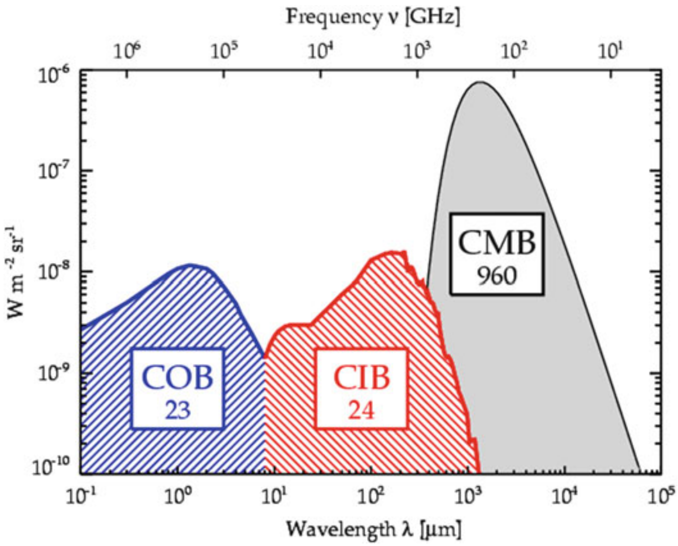


Fig. 4.1 Spectral energy density (SED) function of the Universe adapted from Dole et al. (2006), with the *Herschel* spectral coverage spanning over 60–600 μm approximately, corresponding to the CIB area of the figure. The cosmic microwave background (CMB) is the remnant of the Big Bang, while the cosmic infrared background (CIB) and the cosmic optical background (COB) dominate all electromagnetic radiation emitted later. © Dole et al. (2006)

Fig. 4.2 The galaxy M51 observed with *Herschel*/PACS in the far infrared (*top left*), with the *Hubble* Space Telescope in the visible and near infrared (*bottom left*). The *Herschel* image illustrates the distribution of dust mixed with gas in the interstellar medium of a spiral galaxy. The *Hubble* image is dominated by light from stars. The *right-hand side image* shows a combined far-infrared and visible image. *Red colour* indicates the interstellar medium area that fills in *dark area* in the *Hubble* image. Note that infrared image appears with less resolution than optical image. This is due to the fact that angular resolution decreases with the wavelength. © CEA (V. Minier, P. Abelleira) based on ESA/*Herschel* and *Hubble* Space Telescope data

produced at shorter wavelengths from luminous young stars, and from accretion of material onto protostars,[1] and onto black holes at the centre of galaxies has been absorbed by dust grains in surrounding clouds of gas composed of hydrogen, helium and small amounts of heavier elements (Fig. 4.2). Although the dust grains only constitute about 1% of the total mass of the interstellar medium (gas + dust), they have an important influence on the star formation process. If conditions are favourable, stars can form in such clouds when dense cores within them collapse under their own gravity. The collapse and accretion of matter onto the growing protostars is accompanied by the release of gravitational energy in the form of heat, and eventually the central condensation becomes dense and hot enough for thermonuclear fusion to commence, and a star is born. From the initial stages until the emergence of a fully formed star, no visible light emerges because the dust grains absorb it. But the release of energy during collapse, and later from the embedded young star, heats the gas and dust in the surrounding cloud to temperatures of a few tens of degrees above absolute zero, and the absorbed energy is reradiated at longer

[1] A protostar is a nascent or forming star.

wavelengths. The far-infrared emission from the dust and the gas helps the cloud to remain cool, assisting its collapse. Its detection and its study provide us with a means of investigating the physical and chemical conditions in star-forming clouds and how the star formation process takes place.

Black bodies with temperatures between 5 and 50 K peak in the FIR/submm and gases with temperatures between 10 and a few hundred K emit their brightest molecular and atomic emission lines in this wavelength range. Broadband thermal radiation from small dust grains is the most common continuum emission process in this band. It is a unique calorimetric tool for deriving the luminosities of a wide variety of astronomical sources that are obscured by, or embedded in, interstellar or circumstellar material. Superimposed on the dust continuum spectrum of such sources, radiation from the gas manifests itself also in the form of spectral features at well-defined wavelengths, which are characteristic of their chemical and physical conditions. They can be studied with high-resolution spectroscopic measurements from the detailed shape of the spectral features, revealing a great deal of information about the gas chemical composition, temperature and density and its state of motion such as infall, outflow or rotation. In studying obscured star formation going on today in our own or nearby galaxies, it is therefore essential to make both continuum and spectral observations in the far infrared. This is why the *Herschel* scientific objectives were so promising in going back to the formation of stars over a large time range and how they were able to nurture the birth of life. It has achieved these objectives, thanks to its large telescope and unique complement of direct detection photometers, spectrometers and heterodyne spectrometers.

4.2 Evolution of FIRST/*Herschel* Science Objectives

The ‘Far-Infrared and Submillimetre Space Telescope’ (FIRST) was proposed to ESA in November 1982 and introduced in the Horizon 2000 long-term science programme in 1984, eventually officially approved in January 1985 (see Chaps. 2 and 3) as a ‘High-Throughput Heterodyne Spectroscopy Mission’—one of the four so-called Cornerstones of this plan—based on an 8-m telescope feeding a set of spectrometers and photometers operating in the FIR and submm range. The FIRST proposal stressed that this range is the last major spectral window on the Universe to be opened (at the time, it was not even poorly explored), providing unique access to important phenomena. It listed the main science objectives as cosmology, including the cosmic microwave background (CMB) anisotropies at various angular scales, nearby active galaxies, the interstellar medium (ISM) and star formation, mass loss and Solar System studies, in particular the giant planets and comets.

On 4–7 June 1986, the ESA workshop on a ‘Space-Borne Submillimetre Astronomy Mission’ (see Chap. 3) was organised in Segovia, for the astronomy community to convey to ESA what it wanted the Cornerstone to be. It is interesting to compare the science objectives outlined in the justification of the Horizon 2000 Cornerstone with the output of the Segovia meeting and follow their evolution in the period 1982–1986, also noting that the science from the recently launched IRAS mission started to appear already in 1983 and that another infrared mission, ISO, was selected in March of that same year.

The Cornerstone emphasised both physics and chemistry of the ISM, through spectroscopy diagnostics and cosmology. But in Segovia, cosmology studies were restricted, focusing mostly on galaxy evolution, a re-orientation confirmed in 1987 by the Science Advisory Group (SAG) (see Chap. 3), in charge of defining a mission concept compatible with the overall scientific performance of a Cornerstone while at the same time remaining within the allowable financial envelope. The 29th International Liège Colloquium, 'From Ground-Based to Space-Borne Submillimetre Astronomy', held in July 1990 confirmed the Segovia science objectives and the change of scientific priorities.

As detailed in Chap. 3, the FIRST System Definition Study (SDS) and rider studies conducted in the period 1990–1993 to define an affordable Cornerstone mission resulted in the so-called Red Report which contained a list of six main objectives: stars and the interstellar medium; nearby galaxies; distant galaxies and quasars; the early Universe; studies of planets, their atmospheres and satellites; and observations of comets. It is on the basis of this report that the ESA SPC decided in November 1993 to approve FIRST as Cornerstone 4.

Subsequent to the series of difficulties the ESA scientific programme had to face in the period 1994–1996 (see Chap. 3), at the January 1996 meeting of the ESA Space Science Advisory Committee (SSAC), the request was made 'to review the science priorities of the Cornerstone in view of a foreseen delay of 12–18 month'. A 'science hearing' held in September 1996 involving invited experts identified the 150–500-μm-deep broadband (extragalactic) surveys and related research as being the most important overall key programme FIRST should address. Also included were detailed studies of the physics and chemistry of our own interstellar medium and their equivalent in other stars, including high-redshift objects, as well as spectroscopic studies of comets and of the outer planets in our Solar System, assuming that a number of 'bright source' science goals would likely be addressed by the NASA/DLR SOFIA airborne observatory before the launch of *Herschel*.[2] The summary conclusion of the hearing was that the Cornerstone was scientifically healthy, and its scientific strengths, then reaffirmed, allowed ESA to release the Announcement of Opportunities (AO) for payload proposals in October 1997. The three instruments eventually selected by the SPC in February 1999 are listed in Table 3.1 in their final flight configuration. Their science capabilities reflect these priorities, and the group photo of Fig. 4.3 presents the three principal investigators and the *Herschel* project scientist in 2011.

The Science Management Plan (SMP) describing the various elements of the scientific exploitation of the mission also foresaw a special AO for observing time in the form of 'Key Programmes' addressing the main objectives of the mission and

[2]Actually, in the end, SOFIA's first light occurred one year after the launch of *Herschel*.

Fig. 4.3 From left to right, the three *Herschel's* PIs: for HIFI, Frank Helmich (who in 2008 replaced Thijs de Graauw who occupied that function before); for PACS, Albrecht Poglitsch; and for SPIRE, Matt Griffin surrounding Göran Pilbratt, *Herschel* project scientist holding the A&A Special Issue '*Herschel*: the first science highlights' at a special event held in SRON, Ultrecht, on 8 March 2011. © ESA/*Herschel*

the areas of astronomy where the impact of FIRST would be the greatest. At a historical scientific meeting held in Toledo (Spain) in December 2000, the top-level science targets defined at the September 1996 'science hearing' just mentioned were confirmed as follows:

- Extragalactic photometric surveys in the 150–500 μm band, followed up with spectroscopic detailed studies of the most interesting objects
- Galaxies and galaxy clusters
- Star formation and detailed studies of the physics and chemistry of the ISM, both locally in the galaxy and in external galaxies, including high-redshift galaxies
- Evolved stars and dust
- Solar System studies, through high-resolution spectroscopy of comets and outer planets and searches for Kuiper-belt objects.[3]

At that meeting, FIRST was also officially renamed the '*Herschel* Space Observatory', following a proposal made by Göran Pilbratt, the project scientist.

These top-level objectives were further elaborated in view of helping the community to respond to the AO for Key Programmes, issued in February 2007. The observing time was awarded in two steps: first guaranteed time (GT—owned by

[3]The Kuiper-belt extends from the orbit of Neptune (at 30 AU) to approximately 50 AU from the Sun. Like the asteroid belt, it consists mainly of small bodies or remnants from the Solar System's formation. Although many asteroids are composed primarily of rock and metal, most Kuiper-belt objects are composed largely of frozen volatiles such as methane, ammonia and water.

contributors to the *Herschel* mission, mainly (90%) by the instrument PIs), then open time (OT—available on a competitive basis to the worldwide astronomical community). By coincidence, a total of 21 proposals each for GT and for OT (out of 62 proposals) were allocated observing time, altogether reaching a total of 42 Key Programmes. They collectively represented about 57% of the time available in the nominal mission, the rest being reserved for subsequent AOs spanning all facets of the science discussed in the previous 10 years.

Although the science accessible through IR, FIR and submillimetre astronomy evolved rapidly in the period from the FIRST proposal in late 1982 to the actual *Herschel* science operations in the period 2009–2013, about 30 years later, the actual science objectives of FIRST/*Herschel* had both persistent and evolving components. With no surprise the decrease in telescope size from the initially proposed 8 m, finally to the 3.5 m aperture flown (see Fig. 3.3), mainly affected the extragalactic part of the foreseen heterodyne spectroscopy observations. Nevertheless, extragalactic astronomy overall gained importance through the outcome of the 1996 'hearing' and was eventually facilitated by both the successful development of powerful direct detection instruments, a somewhat colder telescope (Chap. 6), and the scientific developments brought about by IRAS, such as the unexpected discovery of powerful infrared-dominated galaxies and later by ground-based submm observations, ISO and the *Spitzer* space mission. Eventually, the importance of water vapour studies both in the ISM and in the Solar System, though not foreseen in the original proposal, emerged already in Segovia in 1986 and gained importance gradually over time.

4.3 *Herschel* in Orbit

After the launch on 14 May 2009 at 13:12 UTC—together with the ESA *Planck* mission—on board an Ariane 5 ECA rocket from the Kourou Space Centre and both spacecraft (first *Herschel* followed by *Planck*) were released within 30 min after lift-off, very precisely on their initial transfer trajectories towards L2.[4] From then onwards, *Herschel* and *Planck* lived their lives as two physically separated missions. During the first weeks in space, from the initial 4 days of 'Launch and Early Operations' (LEOP), followed by the 'Commissioning Phase' (CoP), the telescope temperature was kept at 170 K using heaters and was then allowed to cool down to reach a predicted temperature below 120 K, permitting the cryocover protecting the cryostat from condensation of outgassed volatiles to be opened, precisely 1 month after launch. Following the end of the ground contact period, a daring attempt was made to produce an early observational result, an operation initially not planned. This 'sneak preview' was scheduled on 14–15 June 2009 on operational day (OD)

[4]L2 or second Lagrangian point in the Sun-Earth system is a virtual orbital point at 1.5 millions of km, moving with Earth around the Sun.

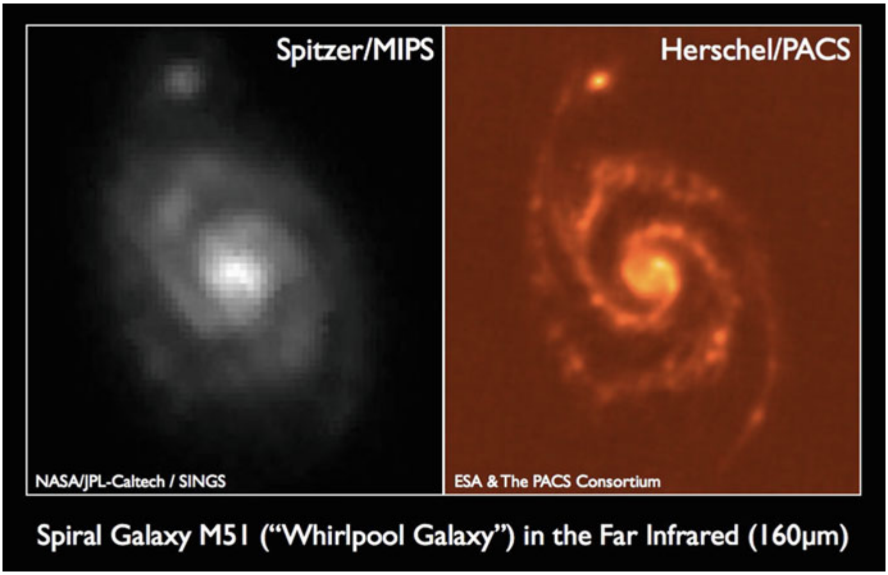


Fig. 4.4 On the *right* is the very first image obtained by *Herschel* at 160 μm. The *left image* is the best of M51, taken by NASA's 85 cm Spitzer Space Telescope, with the Multiband Imaging Photometer for *Spitzer* (MIPS) also at 160 μm. Because of the larger size of its telescope, *Herschel* reveals structures that cannot be discerned in the *Spitzer* image. © ESA/*Herschel*

#32 under unforeseen thermal conditions, consisting of repeated observations of the 'whirlpool' spiral galaxy Messier 51 (M51) using a range of instrument settings for the PACS photometer. In the end, the results were very encouraging as can be judged on Fig. 4.4 and then in the following observations on Fig. 4.5.

Not only did these observations verify that the optical performance of *Herschel*—including the crucial telescope focus and telescope-to-instrument alignment—obtained on the day after cryocover opening, just hours after receiving the very first observational data, was according to expectation, but also the ability to produce the very first maps just a few hours after offered an opportunity for an end-to-end test of the operation system, from observation to delivery of the data product. It is also worth pointing out that a complete optical alignment test could never be performed before launch and that there were no means of adjustments in flight. This was possible to achieve early in the mission, thanks to the overall vision and innovative implementation of *Herschel*, involving ground testing, simulations and preparations for both the space and ground segment (Fig. 4.6) in order to ensure that, when launched, *Herschel's* precious but limited helium cryogen resource would (and could!) be used sensibly and effectively from the very beginning.

'First light' observations involving all instruments were conducted on 21–24 June 2009 (see Fig. 4.5) followed by the successful 'In-Orbit Commissioning Review' (IOCR), marking the start of the 'Performance Verification Phase' (PVP) and opening the way to the 'Routine Science Phase' (RSP (see Chap. 3). These 'first

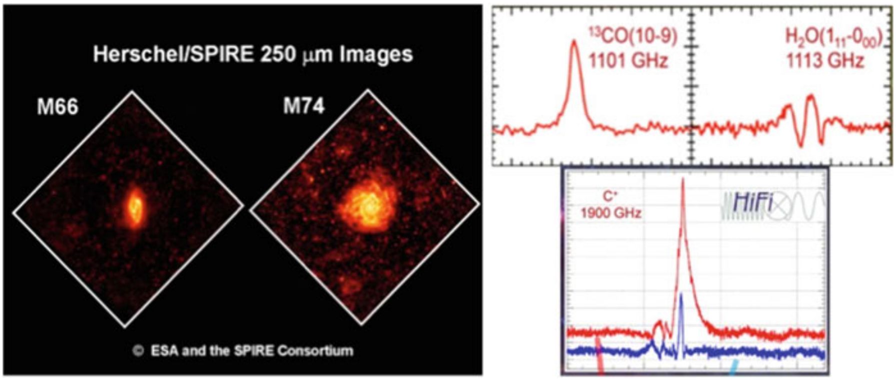


Fig. 4.5 *Left*: 'first light' *Herschel*/SPIRE images at 250 μm of the nearby galaxies M66 and M74, obtained on OD#42, the emission originates from the dust component in their respective interstellar medium. The apparent 'noise' visible as a background in the M74 image originates from a population of distant 'submm' galaxies opening up a new window of the Universe for systematic study and is anything but noise. *Right*: 'first light' *Herschel*/HIFI terahertz spectra of C^+, H_2O and CO towards the DR21, a huge star-forming molecular cloud located in the constellation Cygnus, 6000 light-years from Earth, obtained on OD#39, providing kinematical information implying massive outflows in the heart of DR21 and evidencing the high-resolution heterodyne unprecedented spectroscopy capability of HIFI over a very wide range of frequencies, most of which had never before been covered. © ESA/*Herschel*

light' observations demonstrated and confirmed most of *Herschel*'s major observing capabilities, including several 'firsts' as illustrated on Figs. 4.4 and 4.5, and opened the way to *Herschel* venturing into unchartered waters for real, with the promise of spectacular results. However, two major aspects to be achieved before were to formally release all instruments' observing modes as well as their associated uplink observation templates (AOTs) and downlink software components. This activity intended to ensure that all science observations would only be performed once the observing modes would have been validated end to end. Then, the 42 Key Programmes (each with more than 100 h of observing time) mentioned earlier would be 'validated' in the Science Demonstration Phase (SDP), in view of obtaining data and science results upfront early in the mission to present to the community and preparing for the first in-flight AO to be released in the spring of 2010, aiming at publishing a 'special issue' of the *Astronomy & Astrophysics* (A&A) journal (Fig. 4.3).

That plan, unfortunately, could not be followed. On 2 August 2009, the HIFI instrument malfunctioned and had to be taken out of operation altogether (Chap. 8). The impact was immense, not only on the HIFI team but across the management of the in-orbit technical and scientific operations. All related activities had to be urgently replanned with only two instruments (PACS and SPIRE) available, facing the need to take advantage of the integrity of all of the available observing time while analysing at the same time the evolution of the HIFI situation and identifying possible recovery measures. This meant shorter turnaround times for interpreting

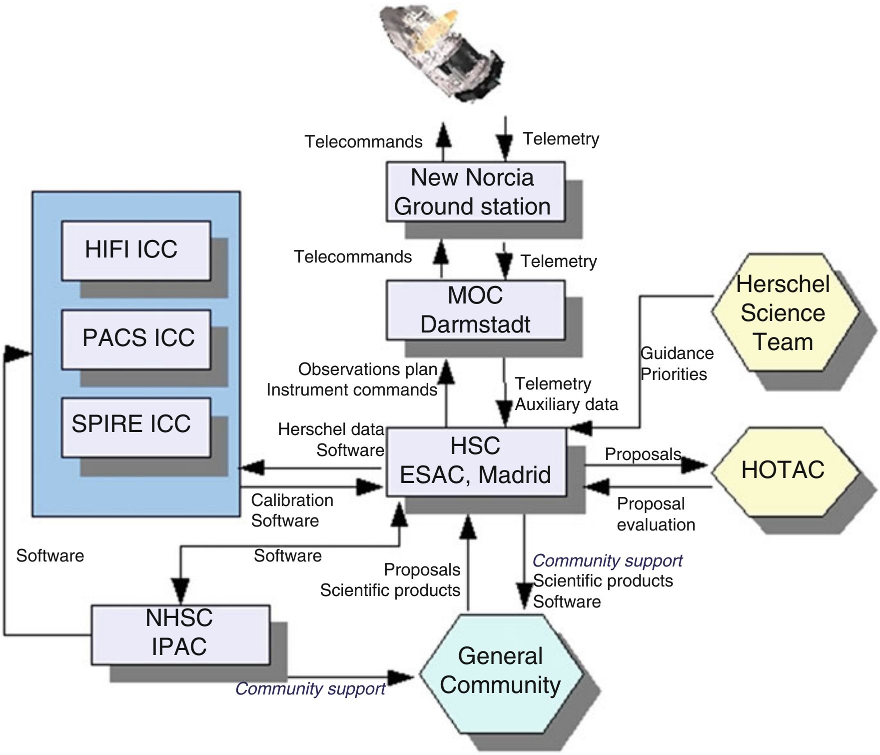


Fig. 4.6 The *Herschel* Ground Segment during flight operations included the *Herschel* Science Centre (HSC), provided by ESA, located at ESAC near Madrid, supported by the *Herschel* Science Team, for the maximisation of the scientific return of the mission and by the *Herschel* Observing Time Allocation Committee (HOTAC) for the selection of observing proposals during operations. The HSC and the NASA *Herschel* Science Center (NHSC), located in IPAC, at Caltech in California, acted as interface to the science community. Three dedicated Instrument Control Centres (ICCs) provided by their respective PI were responsible for enabling the operations and data reduction of each instrument. The Mission Operations Centre (MOC), provided by ESA, located at ESOC, Darmstadt, was responsible for the execution of all in-orbit operations. MOC's role has ended with the end of in-orbit operations, although some support has continued during a 3-month transition phase after switch-off, to ensure the final consolidation of data and the successful closure of outstanding issues. In 2010 a third yellow box for the *Herschel* Users' Group was added to the Ground Segment but is not shown here. © ESA/*Herschel*

obtained data and required the planning of new tests and new observations, putting an additional strain on the teams involved.

A possible—and attractive—solution to both easing that strain and using the observing time efficiently would be to release 'simple' observing modes for 'real' science exploitation as soon as possible. The photometric scanning modes were the obvious candidates, being the simplest ones, and with lots of time-consuming approved observations waiting to be executed. The 'SPIRE/PACS parallel mode', using both photometers simultaneously for mapping large regions of the sky

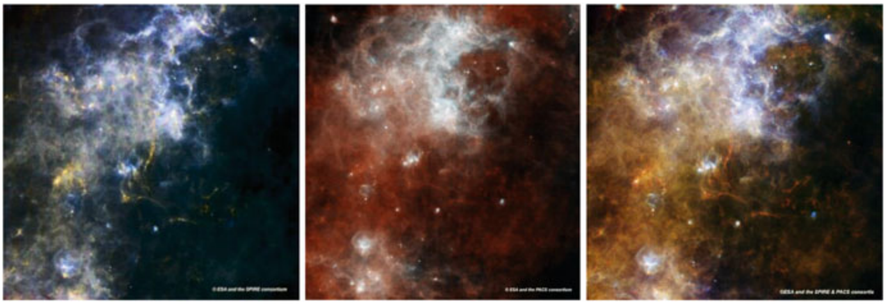


Fig. 4.7 SPIRE/PACS parallel mode observation of the galactic plane in the constellation of the Southern Cross obtained on 3 September 2009. *Left*: in this composite SPIRE image, *blue* denotes 250 μm, *green* 350 μm and *red* 500 μm emission. *Middle*: in this PACS image, *cyan* denotes 70 μm and *red* 160 μm emission. *Right*: in the composite image, *blue* denotes 70 μm and *green* 160 μm emission, while *red* is the combination of the emission from the 250/350/500 μm bands. The colour coding allows differentiating the material that is extremely cold (*red*) from that which is warmer. Material as cold as less than 10 K to few times that value has most of its emission in the *Herschel* bands. Since stars form in such cold and dense environments, the composite image easily locates the star-forming filaments that would be very difficult to isolate from a map made at a single FIR or submm wavelength. © ESA/*Herschel*

separated by some 18 arc-min in five colours simultaneously, was employed for the first time over ODs#111–112 on 1–3 September 2009. The observations were an immediate success, and the data were used to generate beautiful images and make an immediate scientific impact (Fig. 4.7).

Eventually, to the satisfaction of all involved, on 10–14 January 2010, HIFI could be brought back in a fully functional operation situation using its redundant warm electronics and a series of functional testing (Chap. 8). It was then decided to 'overweigh' HIFI observing schedule in the forthcoming couple of months in order to allow for it to catch up while at the same time safeguarding its scientific potential against further anomalies. The performance verification activities of the 'reborn' HIFI were conducted in February, and an identified set of science observations (referred to as the Priority Science Phase—PSP) were conducted in March. Initial results were presented at a workshop held in Leiden on 12–13 April 2010.

The A&A Special Issue mentioned above (Fig. 4.3) '*Herschel*: the first science highlights' (volume 518, July–August 2010) contained 152 papers, most of them based on PACS and SPIRE SDP observations as a consequence of the HIFI anomaly. A second special issue of A&A, with an additional 50 papers based on HIFI SDP/PSP observations, was then published later in October 2010 as part of volume 521. A year after launch, *Herschel* had produced more than 150 submitted papers and organised a major scientific result conference (the '*Herschel* First Results' symposium—aka ESLAB 2010), with the first AO of the two cycles of calls for new observing proposals being imminent. The '*Herschel* Observing Time Allocation Committee' (HOTAC) met on 11–14 October 2010 and concluded their evaluation of the 576 received OT1 proposals, of which 241 were awarded observing time in early November 2010. A second cycle, (OT2), followed a year later.

Overall, *Herschel* successfully executed ~23,400 h of HOTAC-approved science observations, almost 20% more than foreseen to be performed in the prelaunch 'nominal' mission. This 'additional' time was made available due to the slightly longer-than-nominal lifetime of the mission, and because more than nominal science observing time per day (18 h/day) was obtained, as a result of routinely using part of the ground-contact period for science observations. These successful observations covered five 'science areas' comprising galaxies and active galactic nuclei (28%), cosmology 22%, ISM and star formation (39%), star and stellar evolution (8%) and Solar System objects (4%). Overall, 52.6% of the observing time was dedicated to spectroscopy and the remainder to photometry. About 2600 h of scientific calibration observations in standard observing modes were executed on top of the HOTAC-approved science observing time. These science and calibration data are freely available, together with standard 'pipeline' products, being publicly available to download from the *Herschel* Science Archive.

4.4 A Selection of Scientific Results

Herschel has successfully executed the 42 'Key Programmes', together comprising about 45% of the executed science observing time and hundreds of smaller observing programmes. The scientific results span a wide variety of topics; here it will only be possible to highlight three areas that are (and expected to remain) particularly strongly associated with *Herschel* as examples of the prodigious scientific harvest:

- Observations of 'filaments' in detail in relatively nearby molecular clouds revealing a two-stage scenario for the birth of stars like our own Sun
- Observations of infrared-dominated galaxies and the resulting implications for galaxy evolution and the history of star formation across more than 90% of the age of the Universe
- Observations of water vapour and the associated 'water trail' from prestellar cores to planetary bodies in our Solar System and the origin of the water on Earth today

It needs emphasising that by necessity a plethora of other interesting and important scientific results thus have not been discussed. Many others, interesting and important, can be found in specific articles and reviews, referred to in the text and listed at the end of the chapter.

4.4.1 Star Formation in Our Milky Way Galaxy

Herschel has proven to be an essential tool for the study with unprecedented detail of the interstellar medium (ISM) and the earliest phases of star formation in our galaxy. The so-called Gould Belt—a ring of molecular clouds that contains the closest star-forming regions—has been mapped with the PACS and SPIRE instruments in the '*Herschel* Gould Belt Survey' Key Programme (HGBS—André et al. 2010) providing a new insight of the formation process for low-mass stars, including solar-type stars.

The most significant result from the images of these clouds is their highly complex structured extended emission showing ubiquitous networks of thin elongated structures referred to as 'filaments' (Fig. 4.8), which may extend to several light-years in length and tend to have a 'universal' width of around ~0.1 pc[5] or one third of a light year, even though their densities can vary over several orders of magnitude (Arzoumanian et al. 2011). This particular characteristic width is significant as it implies that the filaments contain similar amounts of both thermal and turbulent energy and thus that their formation process is not determined primarily by gravitational effects but rather by other ones such as possibly the dissipation of magnetohydrodynamic turbulence within the parent clouds.

Besides revealing this filamentary structure, the *Herschel* maps allow the detection of large numbers of prestellar condensations (cores) and protostars, two phases in the formation process of a star. They are found in many but not all of the fields. Wherever found, the majority of them are clearly overlapping the filaments. In fact, significantly the filaments are the structures in which most pre- and protostellar sources can be found, and the key physical properties of these—their masses, luminosities, temperatures and lifetimes—can be established. However, while the existence of filaments is ubiquitous, star formation is not taking place in all filaments suggesting that filament formation precedes star formation.

In several of the fields observed with PACS (Schneider 2013, 2015), the probability distribution functions (PDFs) of the column density (Fig. 4.9) reveal a characteristic pattern consisting of normal logarithmic distributions, either without a power-law tail, synonymous of a cloud structure governed by isothermal supersonic turbulence, or with a power-law tail indicating that gravitation is dominating, whereby parts of the gas are undergoing gravitational collapse. The studies of complete networks of filaments formed within molecular clouds reveal a two-stage scenario for the birth of stars. Filaments with densities of more than ~16 $M_\odot$/pc[6] become gravitationally unstable and fragment, while the filaments that produce stars do have critical surface densities of ~160 $M_\odot/pc^2$. These observations suggest that prestellar cores preferentially form in areas where the observed molecular hydrogen column densities are above ~7×10^{21} per cm^2, which corresponds to a visible light extinction of seven magnitudes ($A_v \sim 7$)[7] (Fig. 4.9 left),

[5]The parsec or pc is a unit of astronomical distances, equal to about 3.26 light-years (3.086×10^{13} kilometres). One parsec corresponds to the distance at which the mean radius of the Earth's orbit subtends an angle of 1 arcsec.

[6]One $M_\odot$ corresponds to one solar mass.

[7]$A_v \sim 7$ implies visible light is absorbed by a factor of ~600.

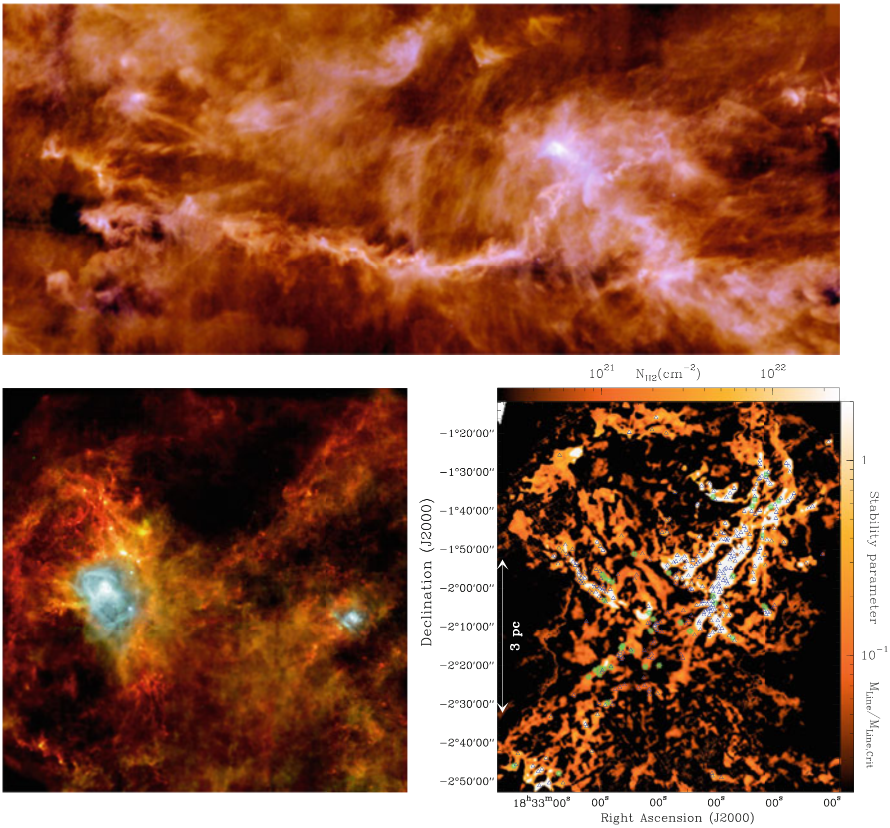


Fig. 4.8 *Top*: SPIRE image acquired with *Herschel* as part of the Gould Belt Survey Key Programme in 2010 and 2012, showing the distribution of gas and dust in the Taurus B211 Molecular Cloud, a giant stellar nursery about 450 light-years away in the constellation Taurus, spanning about 2°–5°. © ESA/*Herschel* and Palmeirim et al. (2013). *Bottom left*: composite image of filaments based on SPIRE and PACS data (in *orange* at 500 μm and 350 μm, in *green* at 250 μm and at 160 μm and in *blue* at 70 μm) in the Aquila molecular cloud located at a distance of about 850 light-years, in the Gould Belt. *Bottom right*: Aquila column density map analysed through a curvelet Fourier transform algorithm enhancing the contrast of filamentary structures. The prestellar cores (*blue triangles*) and protostars (*green stars*) in Aquila are clearly preferentially found overlapping filaments with densities over a certain threshold, shown white in the picture. © ESA/*Herschel* and André et al. (2010)

explaining that the star formation process is taking place within the dense cold parts of molecular clouds, thus remaining invisible to optical telescopes.

The process of forming high-mass stars of O and B spectral types,[8] of masses $>15\ M_{\odot}$ and 2–15 $M_{\odot}$, respectively, is more rare and more complex but also more

[8]Most stars are currently classified using the letters O, B, A, F, G, K, and M, a sequence from the most massive and hottest (O and B type) to the least massive and coolest (M type).

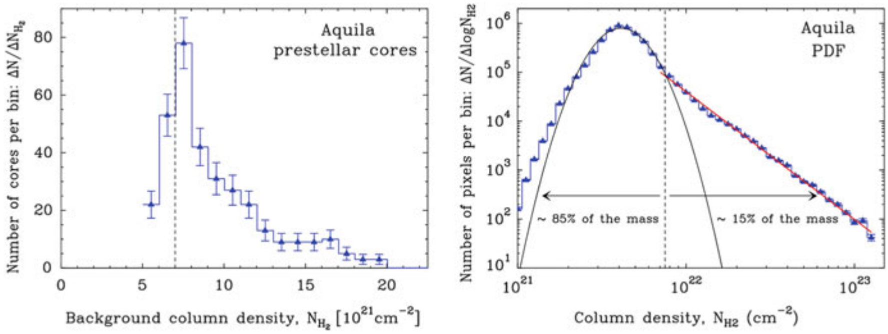


Fig. 4.9 *Left*: number of prestellar cores as function of the column density in Aquila and the corresponding column density probability distribution function (PDF) (*right*). The prestellar cores are preferentially found where the observed column density has a critical value of ~7 × 10^{21} per cm^2 or higher (cf. Fig. 4.8 *right*), which corresponds to where the power-law tail of the PDF emerges, indicating that gravity is dominating above this critical column density. © André et al. (2012)

rapid. It is found in massive filaments with extremely high visual extinction of $A_v > 100$—and in massive clumps and clusters, found at filaments' intersections. The '*Herschel* imaging survey of OB Young Stellar objects' Key Programme (HOBYS, Motte et al. 2010) imaged essentially all of the regions forming O- and B-type stars at distances less than 3 kpc from the Sun providing spectacular and beautiful images of molecular clouds such as the Rosette nebula at a distance of ~1.6 kpc in the constellation Monoceros shown in Fig. 4.10 (left) and the Eagle nebula (right), a young active high-mass star-forming region in the Serpens constellation at ~2 kpc away from the Sun. Although O stars are very rare and short-lived (the more massive the star, the rarer and more short-lived), thanks to their high surface temperatures, enormous luminosities and white/blue colours, they dominate the appearance of the entire galaxies, giving actively star-forming galaxies white/blue colours in visible light. In both cases, the ISM is sculpted by the harsh radiation from the O and B stars (outside of the imaged areas in the figure), which are warming up the surrounding gas and dust, showing up as blue in the figure, and appear to be eroding away material from the surface of the cloud creating the various sculpted features, such as the famous 'Pillars of Creation'[9] in the Eagle nebula (recognisable in the low-centre region of Fig. 4.10 (right)). Massive stars can also trigger the formation of new stars when their UV radiations ionise and compress the surrounding molecular gas. They can then produce bubbles of ionised gas with compressed shells (Zavagno et al. 2010) or transform a filament into a bipolar nebula (Minier et al. 2013).

[9] 'Pillars of Creation' is the name of structures of interstellar gas and dust reminiscent of elephant trunks revealed by the *Hubble* Space Telescope in the Eagle nebula, some 6500–7000 light-years from Earth, so-called because they are in the process of creating new stars, while at the same time being eroded by the light from nearby stars that have recently been formed.

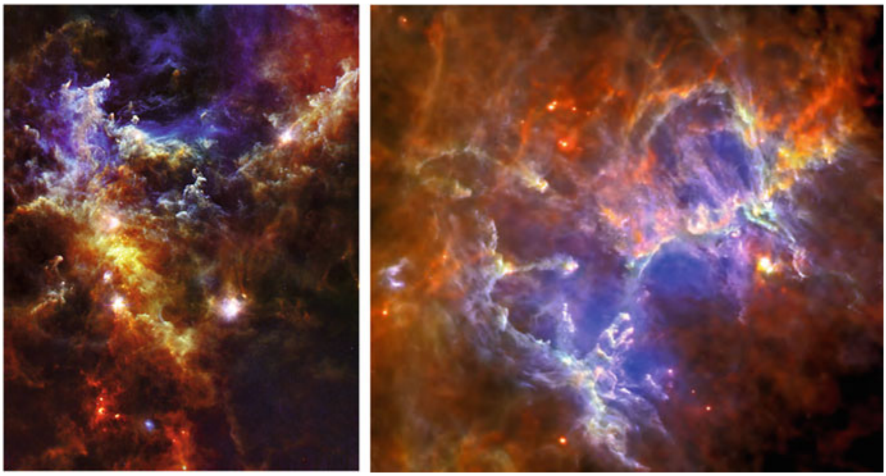


Fig. 4.10 *Left*: composite PACS/SPIRE images of the Rosette molecular complex (Motte et al. 2010) and the M16 Eagle nebula (*right*) (Hill et al. 2012). Both are three-colour images with 70 μm (*blue*), 160 μm (*green*) and 250 μm (*red*) and clearly display the complexity of the interstellar medium and interactions with stars. © ESA/*Herschel*

The '*Herschel* Infrared Galactic Plane Survey' Key Programme (Hi-GAL, Molinari et al. 2010a, b) survey is the largest *Herschel* observing programme of the mission both in terms of observing time and area of the sky covered. Using ~1000 h (~4% of the scheduled science time), it has observed the galactic plane in five photometric bands (100, 160, 250, 350 and 500 μm) covering 360 degrees in longitude of the entire galactic plane, within a range in latitude of ±1 degree. This huge database will be used for many years to study the formation of stars throughout the galaxy, including O and B stars, and their interaction with the interstellar medium, the timeline and history, trigger rates and efficiencies of high-mass star formation, as well as the lifecycle of dust, providing a definitive measure of the amount of cold material in our galaxy. In practice, identifying individual sources with their corresponding positions and fluxes in crowded regions like the galactic plane (Fig. 4.11) is a complex task prone to errors. Nevertheless, the first release of source catalogues covering the longitude range −70 to +68 degrees of the inner galaxy (Molinari et al. 2016) provides no less than 123, 308, 280, 160 and 85 thousand compact sources, in the 70, 160, 250, 350 and 500 μm bands, respectively.

Our galaxy has also been surveyed by HIFI (Langer et al. 2010, 2014) to study the gas composition directly along 500 lines of sight through the galactic plane (Fig. 4.12 left) in the fine structure line of ionised carbon [CII] at 1900 GHz covering the entire galaxy. This line acts as a major ISM cooling mechanism and is an important tracer of the properties of the diffuse atomic and molecular gas clouds. This dataset has been combined with carbon dioxide (CO) and atomic hydrogen (HI) datasets, with the objective of characterising the evolution from

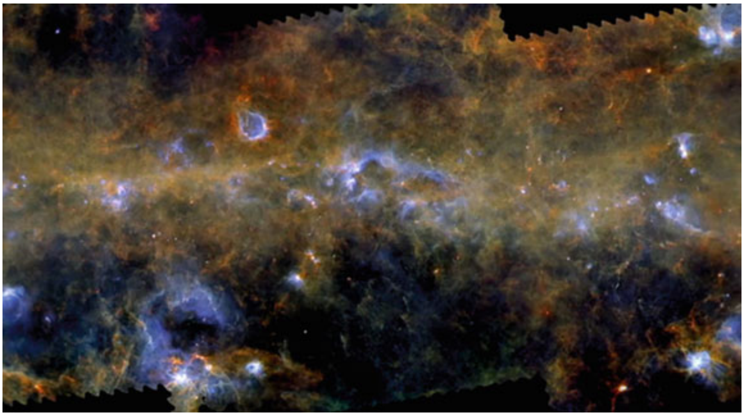


Fig. 4.11 This image of the galactic plane covers galactic longitudes from 10 to 15 degrees. It is a composite figure combining *Herschel* PACS 70 (*blue*) and 160 (*green*) and SPIRE 350 μm (*red*) data. It spectacularly illustrates the complexity of the interstellar medium evidencing a wealth of bright sources, wispy filaments and bubbling nebulas against the background of diffuse gas and dust, identifying areas where stars are being born in the galaxy. © Molinari et al. (2016)

atomic to molecular clouds, using [CII] as a probe of the intermediate cloud state with abundant molecular hydrogen (H_2) and little, or no, HI or CO showing that there is a significant 'CO-dark' gas component in the Milky Way with no detectable CO emission (Pineda et al. 2013). At distances greater than the Sun-to-galactic centre, this 'CO-dark' gas component is even dominating (Fig. 4.12 right). These datasets will be exploited for many years to come to study the formation of stars throughout the galaxy, the ISM itself and the interaction between both.

Because it is the most nearby region where high-mass stars formation is taking place, being close enough (~420 pc) to be studied in detail, the Orion Molecular Cloud (OMC) occupies a prominent position in this field. Consequently, it has been targeted by a number of different *Herschel* observing programmes, combining imaging surveys of the kind already described, as well as a very extensive spectral survey. The Orion Kleinmann-Low (Orion KL) nebula is the brightest region of the OMC in the far infrared and one of the chemically richest regions of our galaxy. It is a massive star-forming region and a nascent cluster of embryonic or very young stars embedded in their parental molecular cloud. It was the target of a 1.2 THz spectral survey using the HIFI instrument and spanning a frequency range from 480 to 1907 GHz at a resolution of 1.1 MHz. This largest spectral coverage ever obtained towards this high-mass star-forming region in the submm range with high spectral resolution was conducted in March 2010, soon after HIFI was back in operation (Sect. 4.3). A total of ~13,000 lines from 79 isotopologues of 39 molecules were identified (Crockett et al. 2014), with excellent agreement between data and modelling. The relative and absolute abundances derived from the *Herschel* data, and molecular fits, which extend from below 100 GHz to beyond 1.9 THz, represent a legacy for comparison with other sources and chemical models (Crockett et al. 2014).

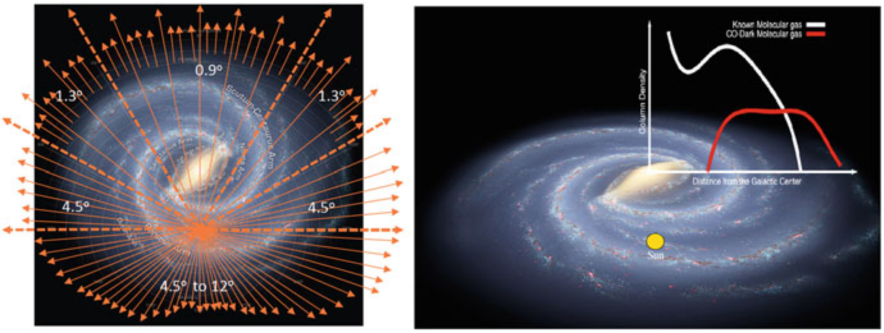


Fig. 4.12 *Left*: galactic observations of the Terahertz CII (GOT C+) 1900 GHz line survey, using a total of approximately 500 lines of sight. The outcome is shown on the *right image* where the small diagram compares the amount of known molecular gas (*white line*) with the amount of 'CO-dark' molecular gas in the galaxy (*red line*). The 'CO-dark' gas dominates the known molecular gas in our galaxy at distances larger than the Sun-to-galactic centre distance (~8 kpc). © ESA/*Herschel*

4.4.2 Evolution of Galaxies and Star Formation over Cosmic Time

Far-infrared observations of the interstellar medium (ISM) in galaxies beyond our own provide a tracer of their evolution. In 1983, the Infrared Astronomical Satellite (IRAS) detected ~75,000 galaxies extremely bright in the infrared. *Herschel* has enriched this view, by observing the far-infrared emission of the ISM in more, and more distant, galaxies. The intensity of that emission primarily depends on the temperature and on how much dust (and therefore how much gas) there is in the ISM. Infrared-dominated galaxies either have a high level of star formation or an active galactic nucleus (AGN[10]), or a combination of both. They radiate most of their energy in the 50–100 μm range, while more distant galaxies would have this range red-shifted by the expansion of the Universe[11] into the submm range, hence their appellation of 'submm galaxies'. One of the top-level *Herschel* science objectives was therefore to perform photometric extragalactic surveys in the 150–500 μm band, as well as spectroscopy for interesting objects formed early in the Universe.

At the time of the launch, about ~2000 submm galaxies were known. SPIRE's 'first light' observations of nearby galaxies revealed a multitude of 'background

[10]AGN emission is caused by accretion of matter by the central supermassive black hole in a galaxy.

[11]The redshift, normally denoted z, is the directly observable 'reddening' (lengthening) of radiation due to the expansion of the Universe. Given the redshift of a galaxy and by using the values of choice for the cosmological parameters, the look-back time can be calculated, or equivalently, for how long the observed photons have travelled since they were emitted in the galaxy.

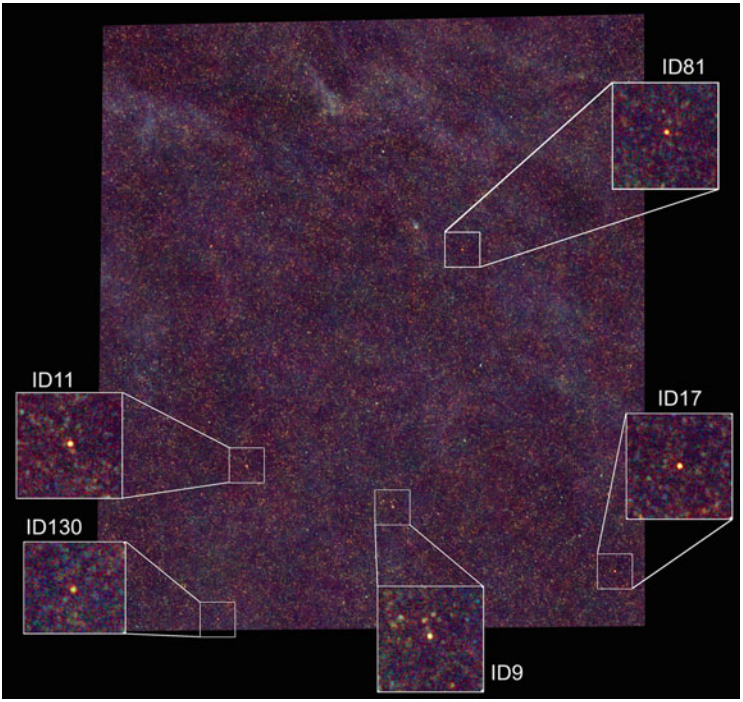


Fig. 4.13 *Herschel*-ATLAS ~4 × 4 deg^2 image of a region of the sky more than 60 times the area of the full Moon containing almost 7000 dusty galaxies. The *insets* are the galaxies detected, thanks to gravitational lensing. © ESA/*Herschel* and Maddox et al. (2010)

galaxies', which were identified to these hitherto elusive submm galaxies! Parallel observations performed in the *Herschel*-ATLAS Key Programme (the largest open time extragalactic survey awarded on *Herschel*), by SPIRE and PACS of a ~4 × 4 deg^2 field (only ~3% of the total field) obtained in ~16 h, revealed almost 7000 galaxies, some of them seen at a time when the Universe was only one fourth its present age, in at least one of the five 100, 160, 250, 350 and 500 μm bands (Fig. 4.13), implying that more than 200,000 galaxies could be detected in the full *Herschel*-ATLAS programme alone, increasing the pre-*Herschel* number by a factor 100!

Figure 4.13 offers a clear illustration of how the field of 'hidden' astronomy has evolved in the past 20 years since the *Hubble* Deep Field, the first area surveyed by a dust-sensitive camera. Five galaxies were found in the Field through 50 h of observation, meaning an average exposure time of 10 h per galaxy. The *Herschel*-ATLAS maps cover an area 100,000 times larger, and it took *Herschel* only 5 s on average to detect a galaxy in these images. The unprecedented number of galaxies detected by *Herschel* has made it possible to study galaxy evolution in the recent cosmic history. A subset of 1688 galaxies detected at 250 μm with optical counterparts showed that for a given space density of galaxies, the 250 μm luminosity

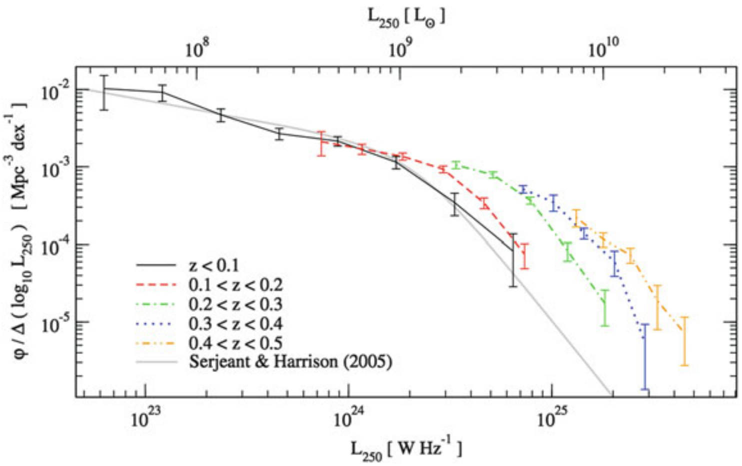


Fig. 4.14 The space density of galaxies on the *y*-axis versus their 250 μm luminosities on the *x*-axis, plotted in colour-coded redshift bins, reveals that the space density of galaxies with a given luminosity increases with look-back time, as computed from their observed redshifts. *Black* is the present, *dashed red* is 1 billion years ago, *green* is 2 billions years ago and *blue* is 3 billion years ago. Thus, galaxies were progressively brighter in the most recent few billion years. © Dye et al. (2010)

increases with redshift (Fig. 4.14). In other words, these galaxies were increasingly more luminous in the past few billion years. Importantly, the amount of dust in these galaxies also increased in the past few billion years (Dunne et al. 2011). As dust and gas—in particular molecular gas—associated with star formation are tightly correlated, these galaxies are also more gas rich with increasing redshifts in the range of $z = 0$ to 0.5 when the age of the Universe was 62% of its present age. It is thus tempting to conjecture that the increased luminosity for a given space density of these galaxies is simply due to the availability of more raw star-making material at earlier epochs. Even only 1 billion years in the past, a small fraction of the age of the Universe, galaxies were forming stars faster and contained more raw material for star formation than galaxies today.

Employing the deepest *Herschel* observations available in the GOODS field,[12] Elbaz et al. (2011) confirmed that over the redshift range $z = 0$, thus in the present Universe, to $z = 2.5$, corresponding to a look-back time up to 11 billion years when the Universe was only 2.8 billion years old, star-forming galaxies' infrared luminosity increases with redshift and argued that they form a 'main sequence' (Fig. 4.15). Thus, what a 'normal' star-forming galaxy is depends on its age. The 'main sequence' might be the best indirect evidence that galaxies are fed continuously with extragalactic matter, since a typical galaxy around 10 billion years ago would exhaust its gas reservoir in only ~500 million years unless replenished.

[12]The GOODS field is a 'popular' field for extragalactic surveys with ample multiwavelength data available from a large number of different observatories.

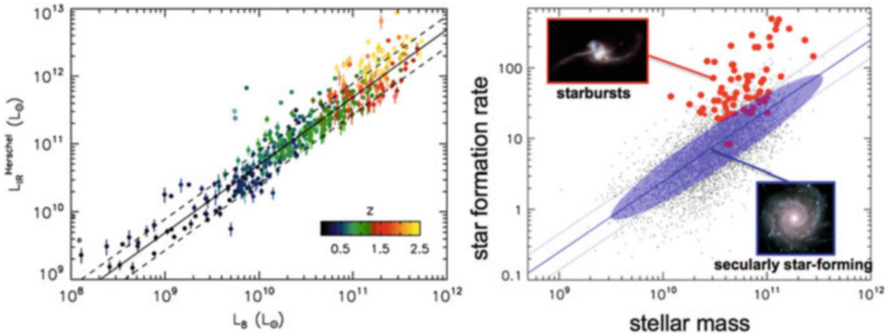


Fig. 4.15 *Left*: the "main sequence" of star-forming galaxies is a plot of the total infrared luminosity (an indicator of the amount of their star formation level) measured in units of solar luminosity versus the amount of 8 μm emission, itself an indicator of the star formation process. It forms a *straight line* up to a *redshift* of $z \sim 2.5$ corresponding to a 'look-back time' of ~11 billion years (Elbaz et al. 2011). *Right*: star formation rate versus stellar mass, identifying the presence of starburst galaxies forming stars at rates more than 10 times higher than the main sequence galaxies. © CEA/ Elbaz (2014)

A corollary to the existence of the 'main sequence' is that the vast majority of stars in the Universe must have been formed in secular mode rather than in starbursts,[13] much like stars are formed today in the galaxy but at a much lower rate. Using *Herschel*/PACS data, Rodighiero et al. (2011) concluded that for galaxies present in the cosmic time range of 9–11 billion years ago, when the peak of star formation rate occurred, starburst galaxies account for only 10% of the star formation rate (SFR) density at this epoch in cosmic history. Gruppioni et al. (2013) extended the analysis of Dye et al. (2010) for PACS-selected galaxies and confirmed that most of the infrared luminosity in the range $z = 0.8$–2.2 originates from 'main sequence' galaxies with masses in the range 10^{10}–10^{11} $M_\odot$. Taking this together, *Herschel* data thus tells us that most stars that were formed in the recent 11 billion years pertain to main sequence galaxies whose properties depend on the age of the Universe, and these galaxies formed more stars in the past because they simply contained more raw material to build up these stars. In that respect, it is actually a valid question to ask why star formation is still taking place at all in the Universe today. Why is there still raw material available, which has not been used up? In particular, in the starburst mode, why has not all of the molecular material been converted to stars?

[13]A starburst galaxy is a galaxy undergoing an exceptionally high rate of star formation, as compared to the long-term average rate of star formation in the galaxy or the star formation rate observed in most other galaxies. In a starburst galaxy, the rate of star formation is so large that the galaxy will consume its entire gas reservoir, from which the stars are forming, on a timescale much shorter than the age of the galaxy. As such, the starburst nature of a galaxy is a phase and one that typically occupies a brief period of a galaxy's evolution. The majority of starburst galaxies are in the midst of a merger or close encounter with another galaxy.

In the local Universe (look-back time less than 3 billion years ago), starburst galaxies are well off the main sequence and as a group appear to be essentially all interacting galaxies (Fig. 4.15, right). Spectroscopic observations with PACS revealed that essentially all of such galaxies do display massive molecular outflows (Fischer et al. 2010; Sturm et al. 2011; Veilleux et al. 2013) with median terminal outflow velocities of ~200 km/s but several with ~1000 km/s. These terminal velocities and the associated mass flows are such that the mass depletion timescales are short, ranging from a few to a 100 million years. Basically, these galaxies seem to be quenching themselves: the observed massive large-scale outflows of molecular gas are removing the raw material that would be used for further star formation, thus ending the starburst but at the same time sparing some of this material for later star formation. How are these outflows generated? A weak correlation noticed between outflow velocities and Active Galactic Nuclei (AGN) luminosity would suggest that the outflows are driven by AGNs. In the case of the nearby Ultra-Luminous Infrared Galaxy IRAS F111191+3257, a combination of observations obtained with the Japanese X-ray astronomy Suzaku spacecraft and with *Herschel*, Tombesi et al. (2015) were able to show that energetically the small-scale mildly relativistic accretion disk black hole-driven winds could be the source of the massive molecular outflows observed by *Herschel*. Besides outflows, *Herschel* spectroscopic observations mapped the emission from gas in galaxies close enough to be spatially resolved. Figure 4.16 associates a SPIRE FTS of carbon monoxide (CO) and ionised nitrogen (N^+) lines towards the centre of M83, a star-forming spiral galaxy (Wu et al. 2015), showing very different distributions. Since different lines trace the state of the gas (molecular, atomic, ionised) and its physical conditions (temperature, density, velocity, etc.), such observations can probe the global structure of the interstellar medium in galaxies and how and where star formation occurs.

Even further, spectroscopic measurement can identify 'normal' galaxies from 'starburst' ones. The *Herschel*-ATLAS field contains five very distant 'lensed galaxies' observable through gravitational lensing[14] whose 500 μm emission is amplified, typically by a factor of 10–30 (Fig. 4.13). Their spectrum was studied by SPIRE and HIFI out to redshifts that would otherwise be impossible without lensing and did not only showed that it was possible to study galaxies' ISM in the relatively early Universe but also provided independent support that a star-forming galaxy with a look-back time of 11.5 billion years has the properties of a 'normal' galaxy. A search over a limited field using the fluxes in the three SPIRE bands as an indicator for high-redshift galaxies produced a host of candidates including the spectacular massively star-forming galaxy-labelled HFLS3 (Riechers et al. 2013) with a redshift of $z = 6.34$ (a look-back time of 13 billion years), implying that we see this galaxy as it appeared when the Universe was less than 1 billion years old. HFLS3 is a major galaxy, with a mass of stars similar to that of the nearby

[14]Gravitational lensing results from the bending of light from distant galaxies by the relativistic attraction of a massive elliptical (submm faint) galaxy exactly located along the line of sight.

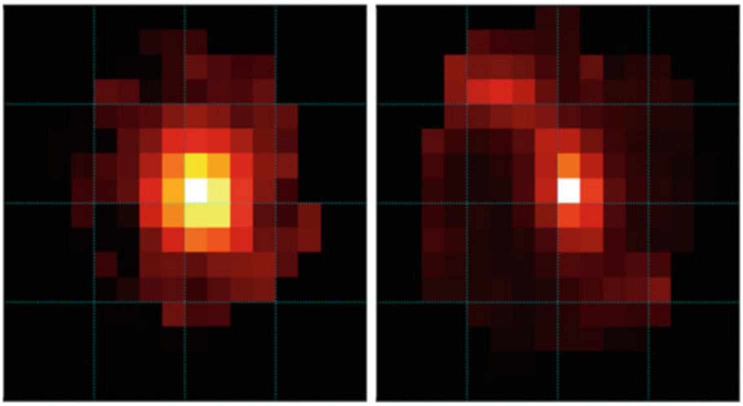


Fig. 4.16 Maps of the emission from CO molecule (*left*) and from the N^+ ion (*right*) in the M83 galaxy. Each pixel colour represents the intensity of the spectral line at this position. It does not represent the morphology of the galaxy but enlightens the distribution of gas (CO, N^+). © CEA/SPIRE

archetypical starburst galaxy Arp 220 but with 20 times the mass of gas and still with a dust mass over a gas mass ratio of ~0.6. With an initially inferred star formation rate in the range of 1000–5000 solar masses per year, HFLS3 challenges current galaxy formation theories.

4.4.3 'Water Trail' and Solar System Studies

Comet missions such as Giotto and Rosetta and space astronomy missions such as ISO, SWAS and *Odin* (see Table 2.2), but most importantly *Herschel*, have played a key role in the search for and study of water in the Universe. Today, water has been observed in gaseous and solid form in various environments under a range of physical conditions throughout our Solar System, in disks around young stars, in protostars, in the interstellar medium in star-forming clouds in our galaxy and in external galaxies, including high-redshift galaxies formed in the early Universe. Still, we are not sure of the details of the provenance of water on Earth, without which life, as we know it, would not have been possible.

In the cold conditions of molecular clouds, the bulk of the water in the prestellar stage is formed in the denser parts mainly on the surfaces of dust grains and is essentially kept frozen out from the gaseous phase as ice. Observationally the largest maps have been obtained in Orion, where water vapour emission is present on ~1 pc scales, more representative of the denser star-forming cores (Bergin and van Dishoeck 2012). Tracing the water abundance from the collapse in the molecular cloud to planet-forming disks—and ultimately to planetary bodies—for a wide range of sources and evolutionary stages was the objective of the 'Water in star-forming regions with *Herschel*' Key Programme (WISH, van Dishoeck et al. 2011).

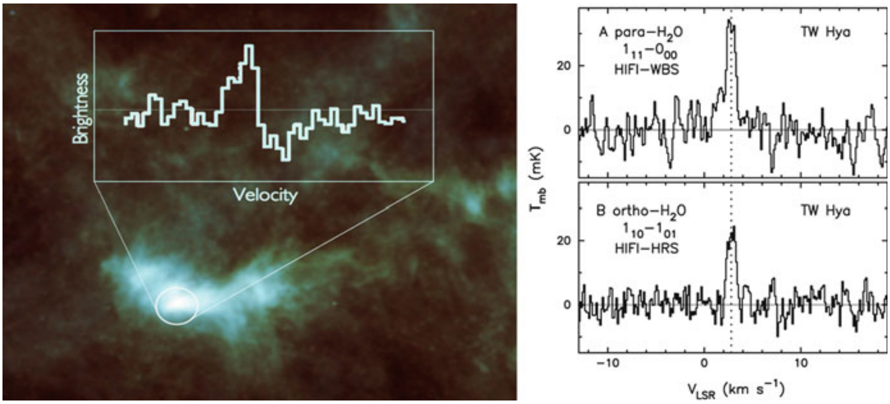


Fig. 4.17 *Left*: observed HIFI 557 GHz ground-state orthowater line (in *white*) overlaid on a SPIRE image of the prestellar core in L1544 (adapted from Caselli et al. 2012). *Right*: ortho- and para-H_2O spectra obtained with HIFI from the cold outer part of the proto-planetary disk around TW Hy (Hogerheijde et al. 2011). © ESA/*Herschel*

Herschel's large collecting area and the sensitivity of the HIFI instrument have allowed to detect for the first time water vapour in the 557 GHz ground state of the ortho-H_2O line[15] in the core of the L1544 molecular cloud in the Taurus constellation (Caselli et al. 2012). The resulting spectrum (Fig. 4.17 left) shows a profile characteristic of gravitational contraction (P-Cygni profile). L1544 shows no signs of stars but is on the verge of collapsing with the line profile indicating infall at ~1000 AU[16] from the prestellar core. The total mass of the observed water vapour is equivalent to that of 2000 Earth oceans. Both HIFI and PACS have also mapped prominent water spectra along protostellar outflows in a variety of different sources, where much more water vapour is liberated in shocks and mainly lost to space. For low-mass protostars, the 'shocked' water totally dominates that of the bulk of the collapsing envelope, even though it contains only a small fraction of the total mass of the system.

Figure 4.18 shows an imaging spectroscopic observation of spectral line emission from water at 179 μm (Nisini et al. 2010). Water is used as a sensitive probe of the bipolar outflow from a protostar, L1157, which is probably similar to the Sun during its formation. The spectral signatures are seen by PACS along the outflow lobes, showing how the outflow interacts with its surroundings. The water emission is most prominent in hot spots due to strong shocks as the outflowing gas from the protostar interacts with dense clumps in the parent cloud.

[15] Molecular hydrogen, H_2O, occurs in two isomeric forms: one with the two proton spins aligned running parallel, called orthohydrogen, and the other with the two protons running antiparallel. The water molecule made from orthohydrogen is called 'orthowater'.

[16] One AU or astronomical unit is a unit of length roughly the average Sun-Earth distance and equals 149.6 million km.

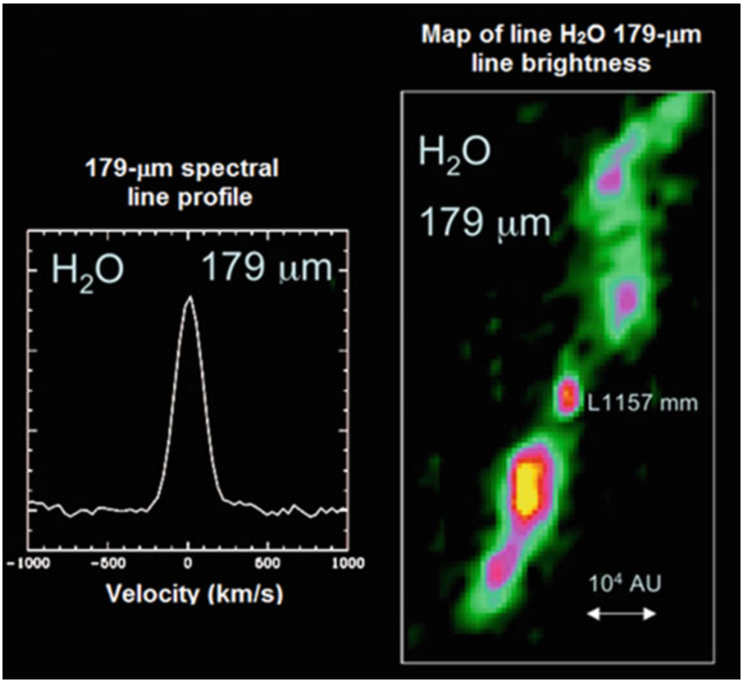


Fig 4.18 *Left*: Strong detection of emission from water vapour towards protostar L1157 at a wavelength of 179 µm by PACS. *Right*: a map of the intensity of the water emission around the source, showing a bipolar outflow. © ESA/*Herschel*

Herschel has also detected cold water vapour emission from the outer disk of TW Hydrae, a young star of order 10 million years, 0.6 $M_{\odot}$, ~55 pc away in the constellation of Hydra (Fig. 4.17 right) (Hogerheijde et al. 2011). With a gas-to-dust ratio assumed to be 100, it may already have, or is already forming, planets, but has no currently known planets. This emission has proven to be elusive, weaker than predicted, and thanks to very persistent time-consuming observations lasting 17 hours, it has been possible to detect both the ortho- and para-H_2O ground-state lines at 557 and 1113 GHz, respectively. Through modelling, it is estimated that the mass of this water vapour is equivalent to that of ~0.05 'Earth oceans', while the ice represents thousands of 'Earth oceans'.

Fomalhaut, a star twice as massive as our Sun, 25 light-years away, has been of keen interest to astronomers for many years. It is only a few hundred million years old, and in the 1980s, the IRAS infrared satellite has shown that it was surrounded by relatively large amounts of dust. *Herschel* with its unprecedented resolution produced images of the star in five colours, at 70 and 160 µm with PACS and 250, 350 and 500 µm with SPIRE (Acke et al. 2012), the best ever far-infrared images of the system (Fig. 4.19). The star itself is surrounded by hot gas and a belt of dusty material on the outer edges of the system, most likely the result of

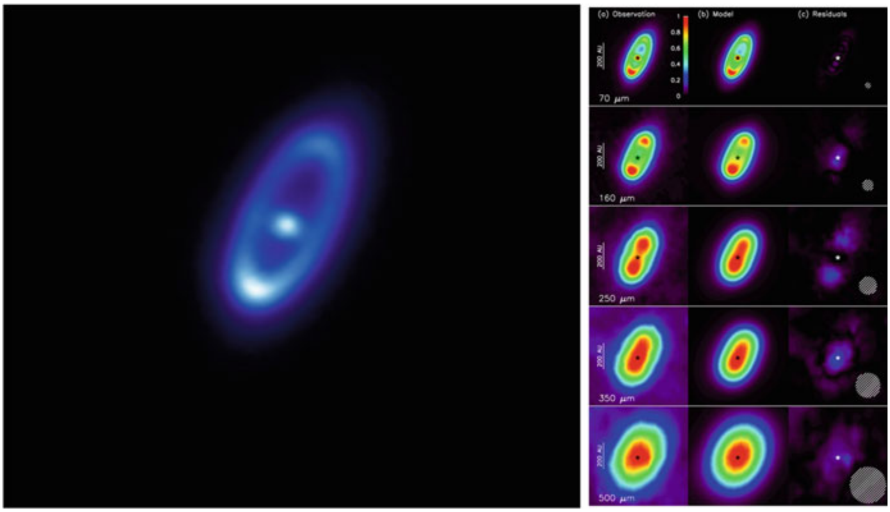


Fig. 4.19 *Left*: spectacular *Herschel*/PACS 70 μm image of Fomalhaut, with its debris disk at ~130 AU). *Right*: images, models and residuals at all five colours. © ESA/*Herschel* and Acke et al. (2012)

collisions between comets, and containing roughly 100 Earth masses, a factor a little larger than estimated for the primordial Kuiper belt[17] of our Solar System.

The observed existence of water ice in the outer disk of TW Hydra and possibly in the debris disk around Fomalhaut relates to the existence of water on Earth. *Herschel* has observed water in several places in the Solar System, causing a few surprises. In that respect, comet 103P/Hartley 2 triggered some excitation in the astronomy community when it was shown that its D/H ratio[18] (Fig. 4.20) matches almost exactly that of the ocean water on Earth (Hartogh et al. 2011b), an observation later confirmed for another Jupiter family class comet, 45P/Honda-Mrkos-Pajdušáková (Lis et al. 2013), which opens the possibility that not all of Earth's water originates from chondrites. *Herschel* also measured the D/H ratio for the Oort cloud[19] comet 2009P1/Garradd where it was found to be inconsistent with the Earth value (Bockelée-Morvan et al. 2012). These discoveries showed that the reservoir

[17]The Kuiper belt**,** named after Dutch-American astronomer Gerard Kuiper, is a circumstellar disk in the Solar System, extending from the orbit of Neptune (at 30 AU) to approximately 50 AU from the Sun. It consists mainly of small bodies or remnants from the Solar System's formation and is composed largely of frozen volatiles, such as methane, ammonia and water.

[18]The deuterium/hydrogen ratio or D/H is a valuable tracer of fractionation processes in hydrogen-bearing (water, hydrocarbon) systems. It has been measured in hydrogenated molecules preserved in planetary bodies such as meteorites and comets. It is used for determining the carrier of water on Earth.

[19]The Oort cloud named after astronomer Jan Oort, who first theorised its existence, is an extended shell of icy objects that exist in the outermost reaches of the Solar System. It is roughly spherical and thought to be the origin of most of the long-period comets that have been observed.

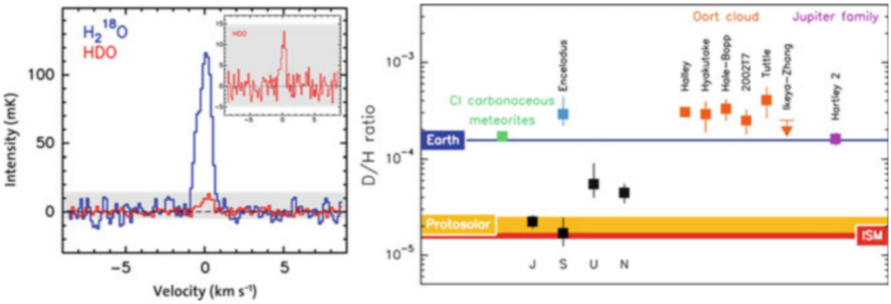


Fig. 4.20 *Left*: *Herschel*/HIFI spectra of comet 103P/Hartley 2 (Hartogh et al. 2011b) and D/H ratios for a number of Solar System bodies including Oort cloud and Jupiter family class comets. *Right*: note that before *Herschel's* results, only chondrites (CI carbonaceous meteorites) had a similar D/H ratio to that of the Earth (Altwegg et al. 2015). © ESA/*Herschel*

of Earthlike water in the Solar System includes both carbonaceous chondrite meteorites and (some) comets but also that there is a difficulty to explain isotopic diversity present in the comet's population. More recently the ROSINA mass spectrometer and pressure sensor on board the ESA Rosetta mission has measured the D/H ratio for the Jupiter family class comet 67P/Churyumov-Gerasimenko and has obtained a very high value (Altwegg et al. 2015), more than three times that of Hartley 2, yielding to the conclusion that D/H values for Jupiter family comets are highly heterogeneous. It is questioned whether the current understanding of the provenance of water from the Oort cloud and Jupiter family comets is correct and even whether the distinction between the two does make sense at all.

The assumption that water can be delivered to a planet by an impacting small body has been confirmed by *Herschel*'s observations of Jupiter, which in July 1994 was hit by the fragments of comet Shoemaker-Levy 9 whose single nucleus was torn to some 21 pieces by Jupiter's strong gravity during a close encounter with the giant planet in 1992 (Fig. 4.21 left). Three years later, ISO observed for the first time water vapour in the high atmosphere of Jupiter, and 16 years later, HIFI's raster mapping observations at high spectral resolution in the water line at 179.5 μm show that the bulk of the water is confined to pressures lower than 2 mbar. This is high in the stratosphere of the planet, well above the cold trap at the top of the troposphere—which does not allow much water vapour to travel past it—indicating an external origin of the water. Furthermore, the 66.4 μm PACS maps show that the water is 2–3 times denser in the southern hemisphere, with the maximum column density being observed at 44°S, the same as the comet's impact. Taken together these two facts suggest a 'single' event—the 1994 impacts—as to the origin of the water in the atmosphere of Jupiter rather than a continuing process emanating from the icy rings or the Moons (Cavalié et al. 2013).

HIFI also observed water lines at 557, 987, 1113 and 1670 GHz, in Saturn and its satellites, and showed that water emanating from deep within the Moon Enceladus as giant plumes end up in a giant torus around Saturn centred on the Moon's orbit (Hartogh et al. 2011a). Some of this water ultimately ends up in the upper

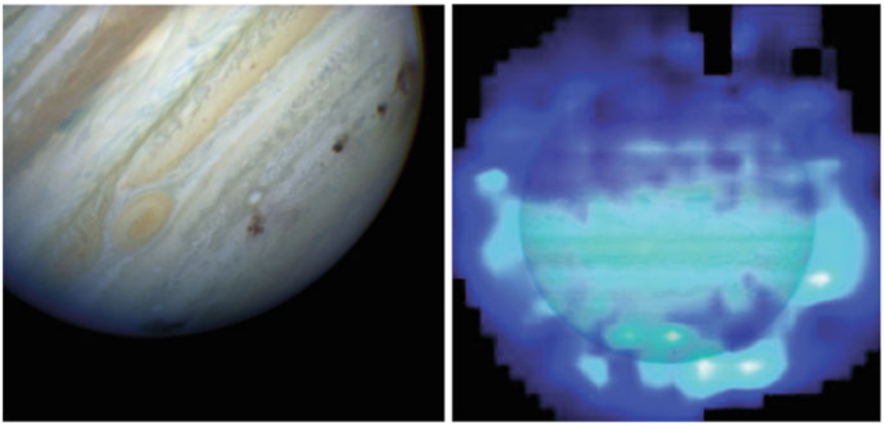


Fig. 4.21 *Left*: tracks left by the impacts of the pieces of comet P/Shoemaker-Levy 9 as observed by the *Hubble* Space Telescope in July 1994 (credit NASA-ESA). *Right***:** column density of water vapour in the stratosphere of Jupiter as observed across its disk by *Herschel*/PACS at 66.4 μm (Cavalié et al. 2013). © ESA/*Herschel*

atmospheres of Saturn and also of its giant moon, Titan. This is the first time a Moon has been observed to alter the composition of the atmosphere of its host planet. *Herschel* has also observed various members of the smaller bodies in the Solar System. The largest of these, the dwarf planet, Ceres, created surprise when HIFI's observations of the 557 GHz line revealed the presence of water vapour (see Fig. 4.22). Observations on four different occasions including repeatedly throughout a complete Ceres revolution indicate that the probable origin of the observed water is two localised sources on the surface of Ceres (Küppers et al. 2014). The water evaporation could be due to comet-like sublimation or to cryovolcanism, in which volcanoes erupt volatiles such as water instead of molten rocks. This is the first unambiguous detection of water vapour around an object in the asteroid belt.

4.5 Bibliometric Analysis of *Herschel's* Scientific Return

Bibliometric analyses may provide many insights into the impact of *Herschel* on the astrophysics field. First, one way to measure the return on investment for the scientific community is through the number of published refereed papers based on *Herschel's* data and their relative importance in the whole field of astrophysics as measured by the number of citations they have received. Second, bibliometrics analysis can provide information about the internal structure of '*Herschel's* community', i.e. about the different centres of interests of the astrophysicists who refer to *Herschel*'s data.

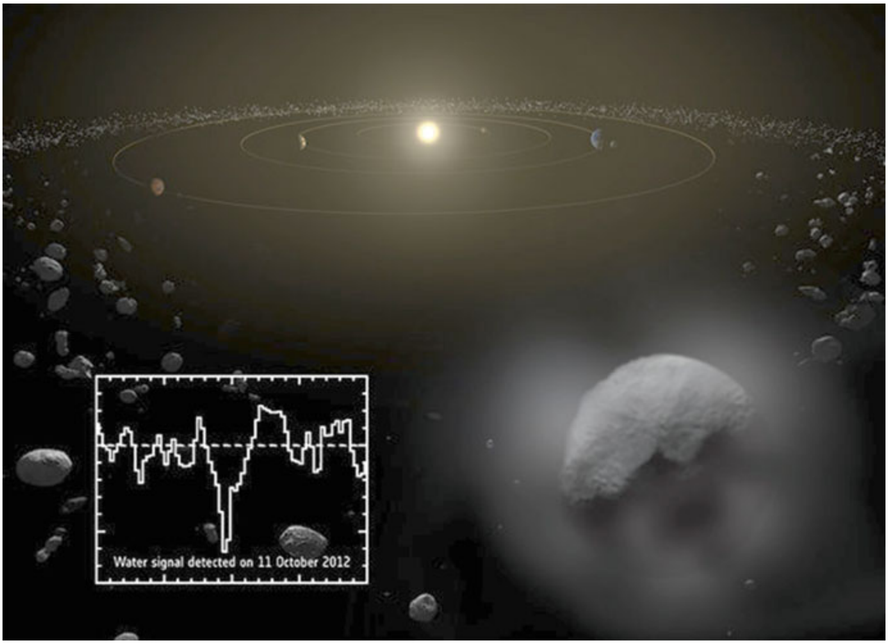


Fig. 4.22 HIFI detection of water on Ceres in the 557 GHz line, overlaid with an artist's montage of Ceres position in the asteroid belt between Mars and Jupiter. © ESA/*Herschel*

4.5.1 Scientific Return on Investment: ESA Key Performance Indicators

A *Herschel* publications list, currently referred to as the '*Herschel* Refereed Publications Library'[20] (HRPL), is maintained by the *Herschel* project scientist. It is publicly available in two formats from the *Herschel* Science Centre website, the Astrophysics Data System (ADS) format,[21] and through a web interface providing additional functionality to perform searches among the *Herschel* papers. A webpage with a listing of *Herschel*-related PhD theses is also provided.[22] The number of publications is reported to the ESA Science Programme Committee (SPC), since 2015 augmented with 'key performance indicators' (KPIs) defined by ESA for the purpose of providing the member states through their SPC delegations with quantitative measures of the performance of the science programme (ESA 2015).

As described in Sect. 4.2, a special effort was conducted under the framework of the Science Demonstration Phase (SDP) to generate and provide the maximum

[20] http://www.cosmos.esa.int/web/herschel/scientific-publications

[21] The URL http://adswww.harvard.edu/

[22] The URL is http://www.cosmos.esa.int/web/herschel/theses

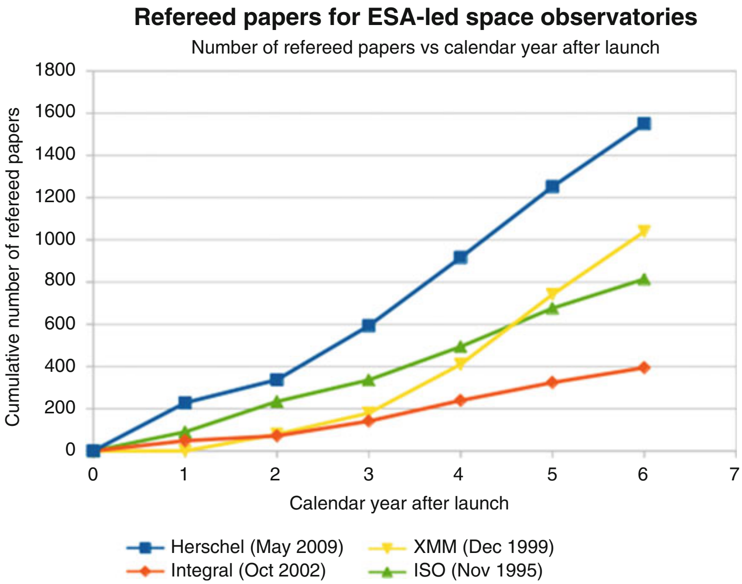


Fig. 4.23 Commencing with the *Herschel* A&A Special Issues in 2010, *Herschel* has maintained a healthy publication rate in comparison with all previous ESA-led space observatory missions (ISO, XMM-*Newton* and Integral). Note that while ISO was performing observations for less than 3 years and *Herschel* for less than 4 years, both XMM-*Newton* and Integral are still operational. © ESA

amount of information to the users' community in the context of the first inflight call for open time observing proposals. This included the *Astronomy & Astrophysics* Special Issue papers, also posted on arXiv[23] after having been accepted for publication. That generated 152 papers, which provided a good start in terms of *Herschel* publications. The rate of publications based on *Herschel* data has continued to be very healthy. In fact, when plotting the number of publications (as taken from the SPC reports) for all the four ESA-led space observatories ISO, XMM-*Newton*, Integral and *Herschel* for the six calendar years immediately following their respective launch year, *Herschel* comes out very impressively (Fig. 4.23).

The curves in Fig. 4.23 should optimally be compared with those related to similar NASA astronomy missions such as the *Hubble* Space Telescope,[24] *Chandra* or *Spitzer*. Unfortunately, such a comparison is difficult because the US publication numbers are defined differently to those of ESA. The only inter-calibration possible is for XMM-*Newton* for which both ESA and US numbers are available. The *Hubble* Space Telescope of course has a very poor publication record for the first

[23]A repository of electronic preprints of scientific papers in the fields of mathematics, physics and astronomy.

[24]Available on http://archive.stsci.edu/hst/bibliography/pubstat.html

5 years after launch because of the primary mirror problem and therefore cannot be used in this comparison. According to the NASA numbers as published by ESO, after the first few years after launch, *Chandra* has recorded a rate of about 400–500 publications per year, comparable to the NASA numbers for XMM-*Newton* and *Spitzer* 500–700 per year.

There are many different measures of scientific productivity, none of which is considered to be the obvious one to use for whatever purpose. Included are the numbers of publications (the simplest and perhaps most widely used), citations and various indices that attempt to provide a fuller picture by combining and normalising these different indicators.

The *Herschel* Refereed Publications Library as of 9 August 2016 lists 1742 papers with a total number of citations in excess of 45,000 (as provided by ADS) and an h-index of 87[25]; for reference the top ten most cited astronomical and technical/descriptive papers, respectively, are listed in Table 4.1. The number of PhD theses listed was 91 (as of July 2016); however, it is not known to what level of completeness this number refers or how it compares with what could or should be expected.

There are five KPIs defined by ESA and reported to the SPC. The first two refer to management (cost and schedule) of the programme. The last three are based on the output of refereed scientific publications from the various missions in the programme; they are:

- KPI-3: Number of refereed papers per year (as reported to the SPC)
- KPI-4: 'Impact factor' of ESA-led missions.[26] For *Herschel*, the number reported to the SPC for 2014 was 8.9, the highest of all the ESA science missions. The number for *Planck*, 6.7, was the second highest. Further analysis indicates that this may be due—at least partly—to the fact that these missions were 'young' in 2014, and from comparisons with other missions, the impact factors are likely to decrease with time. For example, in 2015, *Herschel* has decreased to 6.9, which is still a high number (ESA 2016).
- KPI-5: Fraction of refereed papers from all ESA missions, over the total number of refereed astrophysics papers in a given year. This refers to the ESA science programme as a whole, and the number reported to the SPC for 2014 was 11.9%.

Interestingly, at the time when the first *Herschel* data went public in early 2010, the data retrievals from the ISO archive increased, which suggests that availability of the *Herschel* data increased the demand for the 'old' ISO data. A possible conclusion is that the 'new' investment in *Herschel* increased the value of the 'old' investment in ISO or that *Herschel* also increased the science return of ISO. This leads to the hypothesis that a future follow-up mission to *Herschel* could

[25]Defined as follows: for h papers published, each has been cited in other papers at least h times.

[26]Defined for a given year as the total number of citations in this year to the papers published in the 2 previous years, divided by the number of these papers. It is thus the average citation number in a given year of the papers published in the 2 previous years.

Table 4.1 Top ten cited (Source: ADS August 2016)

(a) Science papers
• Star Formation in the Milky Way and Nearby Galaxies, Kennicutt & Evans, ARAA, 50, 531, 2012 (519)
• GOODS-*Herschel*: an infrared main sequence for star-forming galaxies, Elbaz et al. A&A 533, A119, 2011 (496)
• From filamentary clouds to pre-stellar cores to the stellar IMF: Initial highlights from the *Herschel* Gould Belt Survey, by André et al. A&A 518, L102, 2010 (484)
• The Lesser Role of Starbursts in Star Formation at $z = 2$, Rodighiero et al. ApJ 739, L40, 2011 (327)
• Clouds, filaments, and protostars: The *Herschel* Hi-GAL Milky Way, Molinari et al. A&A 518, L100, 2010 (290)
• Galaxy Structure and Mode of Star Formation in the SFR-Mass Plane from $z \sim 2.5$ to $z \sim 0.1$, Wuyts et al. ApJ 742, 96, 2011 (270)
• The *Herschel* Multi-tiered Extragalactic Survey: HerMES, Oliver et al. MNRAS 424, 1614, 2012 (257)
• Massive Molecular Outflows and Negative Feedback in ULIRGs Observed by *Herschel*-PACS, Sturm et al. ApJ 733, L16, 2011 (232)
• Characterizing interstellar filaments with *Herschel* in IC 5146, Arzoumanian et al. A&A 529, L6, 2011 (219)
• The Detection of a Population of Submillimeter-Bright, Strongly Lensed Galaxies, Negrello et al. Science 330, 800, 2010 (193)
(b) Technical/descriptive papers
• *Herschel* Space Observatory. An ESA facility for far-infrared and submillimetre astronomy, Pilbratt et al. A&A 518, L1, 2010 (1619)
• The Photodetector Array Camera and Spectrometer (PACS) on the *Herschel* Space Observatory, Poglitsch et al. A&A 518, L2, 2010 (1177)
• The *Herschel*-SPIRE instrument and its in-flight performance, Griffin et al. A&A 518, L3, 2010 (1059)
• The *Herschel* Data Processing System—HIPE and Pipelines—Up and Running Since the Start of the Mission, Ott, ASPC 434, 139, 2010 (481)
• The *Herschel*-Heterodyne Instrument for the Far-Infrared (HIFI), de Graauw et al. A&A 518, L6, 2010 (362)
• Scanamorphos: A Map-making Software for *Herschel* and Similar Scanning Bolometer Arrays, Roussel, PASP 125, 1126, 2013 (183)
• In-flight calibration of the *Herschel*-SPIRE instrument, Swinyard et al. A&A 518, L4, 2010 (166)
• In-orbit performance of *Herschel*-HIFI, Roelfsema et al. A&A 537, A17, 2012 (145)
• The *Herschel*-PACS photometer calibration. Point-source flux calibration for scan maps, Balog et al. ExpAstron 37, 129, 2014 (63)
• Flux calibration of the *Herschel*-SPIRE photometer, Bendo et al. MNRAS 433, 3062, 2013 (55)

increase the science return also of *Herschel*. This is not necessarily limited to space astronomy. Although we lack numbers, it is already evident that ALMA observing projects that build on *Herschel* observations and results are not uncommon.

4.5.2 *Structure of* Herschel's *Community: Co-citation Mapping*[27]

Since many science cases are addressed thanks to *Herschel* data, different scientific groups gather to exploit them. Indeed, Chaps. 2 and 3 have discussed the fact that *Herschel* was the instrument of a scientific community consisting of 'millimetre astronomers' using molecular spectral lines to study star formation and 'infrared astronomers' using the far infrared to map galaxies' interstellar medium. Thereby, *Herschel* has induced the constitution of a 'technico-instrumental community' (Shinn 2000; Shinn and Ragouet 2005) in the field of astrophysics. In order to map this community, we use an author co-citation analysis based on papers referring to *Herschel* in their title, abstract or keywords.

A co-citation measurement is defined as the frequency with which two papers are cited together by papers belonging to the corpus. The more co-citations two papers receive, the higher their co-citation strength and the more likely they are semantically related. This bibliometrics technique has been introduced by Henry Small (1973), and it has quickly been recognised as an effective tool for the mapping of communities in scientific fields (Small and Griffith 1974). Co-citation analysis is based on the assumption that there is an intellectual relationship between each pair of documents in a bibliography. This method has spread in the bibliometrics studies and become a standard. It should not be confused with bibliographical coupling: papers are bibliographically coupled when they refer to the same papers.

While co-citation analysis is based on pairs of cited papers, the author co-citation analysis is based on identifying pairs of cited authors. This variant has been introduced by Howard White (White and Griffith 1981), who has demonstrated that the mapping of a particular area of science can be done using authors as units of analysis and the co-citations of pairs of authors as the variable that indicates their scientific proximity from each other. The more two authors are cited together, the closer the relationship between them. The raw data are co-citation counts and the resulting map shows identifiable author groups (at least for an expert from the field). This method had an enormous impact on the field of the sociology of science and on other fields whose scholars have used as a tool for mapping their disciplines (Wallace et al. 2009). However, as far as we know, it is the first time that this method is used in order to map the community induced by a large scientific mission such as *Herschel*.

---

[27]This second part of our bibliometrics analysis has been conducted in collaboration with Pauline Huet and Yves Gingras who both work at the Centre interuniversitaire de recherche sur la science et la technologie (CIRST). This Canadian cluster of researchers studies the historical, social and economic dimensions of science and technology and is associated with the Observatoire des sciences et des technologies (OST), which is devoted to the bibliometrics measurement of science, technology and innovation.

Our corpus of raw data has been extracted from the Web of Science database (WoS)[28] due to methodological purpose rather than using the *Herschel* Refereed Publications Library. First, we have selected in the WoS all the documents from 2000 to 2014 whose titles or abstracts include the word '*Herschel*', suppressing all papers, which referred to the physicist William Herschel or to the William Herschel ground-based astronomical telescope. This operation provided us a corpus of 1597 documents. Then we have compared this corpus with all the papers in the *Herschel* Refereed Publications Library (from 2010 to 2014), which contains 1259 papers at the end of 2014.[29] Ninety-three percent of the papers in the *Herschel* Refereed Publications Library were retrieved in the WoS corpus. This difference between the corpuses has no significant impact on the mapping, especially as our author co-citation analysis takes into account only authors whose papers have been co-cited at least 40 times.

The mapping of these data forms a network of approximately 200 nodes as presented in Fig. 4.24. The respective positions of authors' names depend on map's axes, which were arbitrarily set spanning the most divergent groups in order to aid the interpretation of the network. As expected, Göran Pilbratt, as ESA *Herschel* project scientist, is in its centre because his descriptive paper is presumably cited by most of scientists using the data provided by *Herschel*. The mapping shows half a dozen different groups of authors around him. A 'community detector', i.e. an algorithm based on the Louvain method (Wallace et al. 2009) that we have run, has showed that the biggest group (in red) is a subnetwork of nodes, a lot more connected with each other than with the rest of the whole network. It has been identified as a community of researchers working on extragalactic star formation and high-redshift galaxies (high-z galaxies) in the distant Universe. It is connected to the 'submillimetre galaxies and lensing' group (in violet) that is strongly linked to the *Herschel*-ATLAS Key Programme (a galaxy survey led by Eales) including submillimetre galaxies identified through gravitational lensing effect. It means a contrario that the other groups are working on objects in our galaxy or in nearby galaxies. For example, the group circled in green can be identified as researchers using data from HIFI through spectroscopy of molecular clouds, having no connection with the extragalactic star formation. Thus, the locations of the different groups with respect to each other are significant. The 'star formation' group is, for instance, connected to the 'molecular spectral line' group because they share common source targets (e.g. molecular clouds, protostellar cores) and might be engaged together in open time programmes including photometric and spectroscopic observations (e.g. an open time project on radiation impacts of massive stars named 'pillars'). The hierarchy based on the 'degree of centrality'[30] of the authors

[28]Formerly called 'Web of Knowledge' is today's premier research platform helping scientists quickly find, analyse and share information in sciences and humanities.

[29]This preliminary work was achieved in early 2015 and does not contain the data between 2015 and 2016. In November 2016, 1832 articles were listed in the HRPL.

[30]A node's degree of centrality is defined as the number of nodes that are connected to that node.

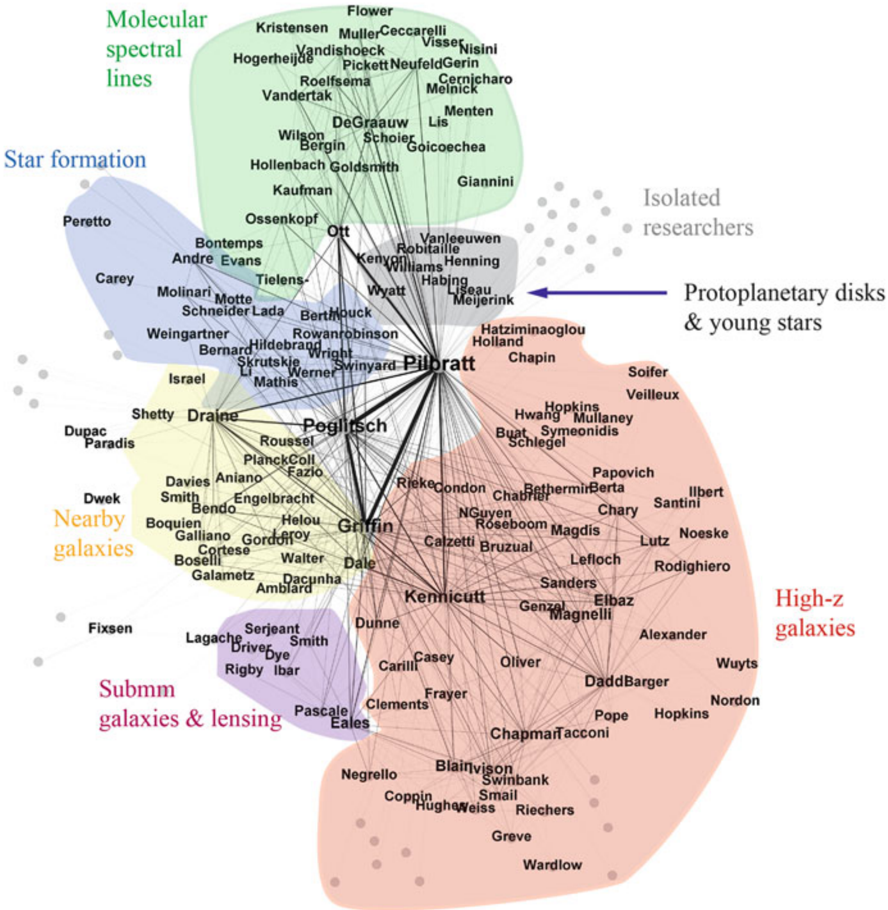


Fig. 4.24 Co-citation map showing authors' co-citation network. Each node represents a *Herschel* paper labelled with the first author name. For clarity purpose, nodes below a certain level of co-citation are not labelled. The overall node distribution is based on the scientific proximity between papers. The closer the nodes are from each other, the closer they would be scientifically linked. Note that the positions of authors depend on map's axes, which were arbitrarily set spanning the most divergent groups in order to aid the interpretation of the network. The size of name label is proportional to the degree of centrality of the node. Six different *Herschel* scientific sub-groups are then identified and indicated in colours. Papers relevant to most sub-groups occupy the central part (e.g. Pilbratt node), and hence the 'descriptive papers' (nodes: Pilbratt, Griffin, Poglitsch, Ott, Roussel) naturally connect them, e.g. SPIRE (Griffin) links galaxies to star formation research and PACS (Poglitsch) star formation to young stars and disks. See main text for further details. © CIRST/OST (P. Huet, Y. Gingras) and CEA (V. Minier, V. Bontems)

within the global network and within the groups is also important. The mapping exhibits therefore significant proximities of authors within groups but also across group boundaries or 'border authors' who seem to connect various areas of research. It is to be noted however that an artefact of the mapping process identifies a group of isolated researchers' nodes, which have not enough links with others to be attached to any other group.

As the co-citations are derived from the top-cited papers and since a lot of these cited papers are recent papers on the same subject, it is without surprise that we observe a partial correlation between the hierarchy in our network and the ranking of the HRPL database. Pilbratt is at the same time the most cited in HRPL database (1530) and predictably the most co-cited with the highest degree of centrality in our corpus (163 connections). Also predictable is the very central positions of A. Poglitsch (with 110 connections) and M. Griffin (106), first authors of the second and third most cited descriptive papers in HRPL database, respectively, about PACS and SPIRE instruments. It confirms that most scientists respect the norm to cite them when they use the data provided by these instruments. Their proximity indicates also that these authors are frequently co-cited, which means that scientists are inclined to use information provided by both instruments, in particular both bolometer imagers that could be used in parallel mode. This compares with the relative eccentricity of the nodes corresponding to the first author of the third HIFI descriptive paper, T. de Graauw, which ranks at the 13th place in terms of centrality because the spectroscopic data of HIFI are less likely to be compared with the data of SPIRE and PACS in the same papers. Ranking at the 12th place, S. Ott's node between the green and grey areas is primarily due to his paper describing the *Herschel* data reduction system (HIPE), which illustrates the importance of the science ground segment as a complementary scientific instrument.

As is also the case in the ranking of the HRPL database, the fourth rank of co-citation is occupied by R. Kennicutt (in the red group), but it is for partially different reasons. Kennicutt's most cited paper in the HRPL database is a science review published in 2012. He is also the most co-cited author from our corpus mostly because he is the author of a very highly cited 1998 science paper about star formation rate in galaxies (Kennicutt 1998). This paper is an obligatory reference for all papers on extragalactic star formation for having modified the so-called Schmidt Law.[31] The centrality of Kennicutt corresponds to the importance of this science case in the whole *Herschel* instrumental community. It also indicates that the red group in Fig. 4.24 is focused on extragalactic star formation. The importance of some seminal papers previously published is confirmed by the B. Draine fifth central node (between blue and yellow areas) who has published several papers based on *Herschel* data since 2010, none of them appearing in the top ten of the HRPL database. His high rank in the co-citation network is due to reference to his older works on 'interstellar dust' modelling. In particular, his paper on 'Optical-Properties of Interstellar Graphite and Silicate Grains' (Draine and Lee 1984) also is an obligatory reference, especially for research on interstellar dust properties through nearby galaxy observations. On the contrary, the low rank (188) of M. Schmidt (red group, one of the nodes without name) in terms of degree of

[31] An empirical relation between the gas density and star formation rate, first identified by M. Schmidt (1959).

centrality, despite the fact that a lot of papers refer to the 'Kennicutt-Schmidt Law', means that the astrophysicists do not often refer to his paper anymore. It is a good illustration of the well-known phenomenon called 'obliteration by incorporation' by the sociologist of science Merton (1949), referring to the tendency for ground-breaking research papers to fail to be cited after the ideas they put forward are fully accepted.

4.5.3 Identification of the Most Important Groups Linked to Science Cases

As mentioned earlier, the 'community detector' has identified the biggest group coloured in red in Fig. 4.24. All the nodes of this subnetwork are strongly connected with Kennicutt's node, which is an obligatory reference for everyone working on star formation in galaxies. The other two most important nodes of this group (S. Chapman and D. Elbaz) are researchers who study the extragalactic star formation in the distant Universe. We can thus presume that the red group encompasses these astronomers working on extragalactic star formation or, more exactly, on high- redshift galaxies. The green group identifies astronomers using the HIFI data and is naturally centred on T. de Graauw as the first author of his instrument in his descriptive paper and is working on 'molecular spectral lines'. Just below, circled in grey, is a dense but little group of astronomers studying 'protoplanetary disks and young stars' in our galaxy. T. Henning is a representative of this group: his papers exploit data provided by PACS combined with data on young stellar objects. On its left, in blue, is a group of astronomers mainly working on star formation in our galaxy with data from PACS and SPIRE. P. André is a typical representative of this group as co-investigator of SPIRE, leader of the SAG 3, a scientific group of SPIRE consortium studying star formation in molecular clouds. Below, in yellow, is a group connected to the nearby galaxy observations and very close to M. Griffin, which means that they are obviously using SPIRE data. Below again, in violet, the last group was also identified by the 'community detector' (see Sect. 4.5.2) as a subnetwork of the red group where the connection is a lot stronger inside than with the rest of the group nodes. These authors are related to the *Herschel*-ATLAS Key Programme that produced a large survey of galaxies including ones detected through gravitational lensing effect. On the left, the 'community detector' has also identified authors more connected together than with the others. This is because they are involved with ISM data from *Planck* used in *Herschel* data analysis, which is also true for the three nodes on the left of Draine's node (below the blue group).

4.6 Conclusion

The long-lasting history of *Herschel* extending over more than 30 years has led to the development and scientific exploitation of a unique astronomical space telescope (Fig. 4.25), providing major breakthroughs, thanks to a newly explored spectral window and a set of particularly innovative and challenging instruments combining both photometry and high-resolution spectroscopy.

In our own Milky Way, *Herschel* has conducted surveys of the interstellar medium and discovered the ubiquity of star-forming 'filaments' in molecular clouds, which are considered as a trademark of *Herschel*. It has also been revealed that these filaments need to have a mass per unit length greater than a certain critical value in order to produce stars; thus, although filaments are ubiquitous, star formation is limited to the most massive ones.

Herschel has also studied the processes of star formation in a huge part of the life of our Universe back to about ~11 billion years before the present time, revealing that for most galaxies, the amount of raw material available for star formation is the constraining factor. Nevertheless, in the local Universe and also at high redshifts, interacting galaxies can trigger massive starbursts displaying huge outflows of molecular gas first seen by *Herschel*, effectively quenching star formation while saving the raw material for future star formation. *Herschel* has extended these observations to when the Universe had less than 7% of its present age, identifying galaxies that challenge the present understanding of galaxy formation and evolution and even including a very high-redshift galaxy exhibiting prodigious star formation when the Universe was less than 1 billion years old.

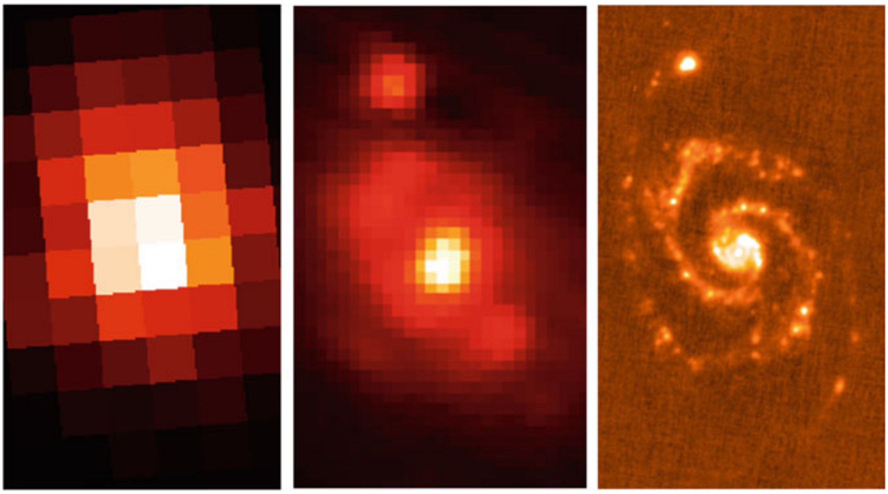


Fig. 4.25 M51 galaxy seen at 60–70 μm with IRAS, ISO and *Herschel*. Resolution and sensitivity increase mainly with the telescope diameter (from 0.5 m to 3.5 m). © CEA based on IRAS, ISO, *Herschel* datasets

The high-resolution spectral capabilities of *Herschel* have made it possible to study the presence of water in all parts of the Universe, importantly 'following' the water along the 'water trail' from prestellar core stage, through circumstellar disk stage and infrared excess of main sequence stars, and confirmed that water on Earth has not one single origin but can result from the bombardment of our planet by both asteroids and comets and suggested that such a bombardment by the debris of comet Shoemaker Levy 9 with Jupiter in 1994 has left the atmosphere of the giant planet with 'packets' of water whose scars may be still 'visible' more than 22 years after the impacts.

Bibliometrics analyses have provided many insights into the impact of *Herschel* on astrophysics. In terms of published papers (ESA space astronomy missions only), the citation and impact of *Herschel* ranks this mission, together with *Planck*, as the most prolific one up to date. The scientific community involved in *Herschel* encompasses thousands of scientists all around the world, and the number of unique names[32] on the papers in the *Herschel* Refereed Publications Library approaches 7000, corresponding to roughly 4000 different authors. Using the bibliometrics method of co-citation measurement, *Herschel* appears as more than a science machine: it brings together and structures the young submillimetre astronomy community founded on scientific and technical heritages from radio millimetre and infrared astronomy.

References

Acke B et al (2012) Astron Astrophys 540:A125
Altwegg K et al (2015) Science 347:1261952
André P et al (2010) Astron Astrophys 518:L102
Arzoumanian D et al (2011) Astron Astrophys 529:L6
Bergin EA, van Dishoeck EF (2012) Philos Trans R Soc A 370:2778
Bockelée-Morvan D et al (2012) Astron Astrophys 544:L15
Caselli P et al (2012) Astrophys J Lett 759:L37
Cavalié T et al (2013) Astron Astrophys 553:A21
Crockett NR et al (2014) Astrophys J 787:112
Dole H et al (2006) Astron Astrophys 451:417
Draine BT, Lee HM (1984) Optical properties of interstellar graphite and silicate grains. Astrophys J 285:89–108
Dunne L et al (2011) Mon Not R Astron Soc 417:1510
Dye S et al (2010) Astron Astrophys 518:L10
Elbaz D (2014) In: Presentation in Star formation across space and time symposium. http://www.cosmos.esa.int/web/herschel/star-formation-across-space-and-time
Elbaz D et al (2011) Astron Astrophys 533:A119
ESA (2015) SPC (2015) 14
ESA (2016) SPC (2016) 13

[32]Note that a unique person could have more than one unique name and that conversely – but less commonly – a unique name could apply to more than one unique person.

First Light. http://www.cosmos.esa.int/web/herschel/first-light
First Results Symposium. http://www.cosmos.esa.int/web/herschel/first-results-symposium
Fischer J et al (2010) Astron Astrophys 518:L41
Gruppioni C et al (2013) Mon Not R Astron Soc 432:23
Hartogh P et al (2011a) Astron Astrophys 532:L2
Hartogh P et al (2011b) Nature 478:218
Hill T et al (2012) Astron Astrophys 542:A114
Hogerheijde MR et al (2011) Science 334:338
Kennicutt RC Jr (1998) The global Schmidt law in star-forming Galaxies. Astrophys J 498:541–552
Kennicutt RC Jr, Evans NJ (2012) Star formation in the milky way and nearby galaxies. Annu Rev Astron Astrophys 36:189–231
Küppers M et al (2014) Nature 505:525
Langer W et al (2010) Astron Astrophys 521:L17
Langer W et al (2014) Astron Astrophys 561:A122
Lis D et al (2013) Astrophys J Lett 774:L3
Maddox SJ et al (2010) Astron Astrophys 518:L11
Merton RK (1949) Social theory and social structure. Free Press, New York
Minier V et al (2013) Astron Astrophys 550:50
Molinari S et al (2010a) Pub Astron Soc Pacific 122:314
Molinari S et al (2010b) Astron Astrophys 518:L100
Molinari S et al (2016) Astron Astrophys 591:A149
Motte F et al (2010) Astron Astrophys 518:L77
Nisini B et al (2010) Astron Astrophys 518:L120
Palmeirim P et al (2013) Astron Astrophys 550:A38
Parallel Mode. http://www.cosmos.esa.int/web/herschel/first-parallel-mode-images
Pineda JL et al (2013) Astron Astrophys 554:A103
Riechers DA et al (2013) Nature 496:329
Rodighiero G et al (2011) Astrophys J 739:L40
Schmidt M (1959) Astrophys J 129:243
Schneider N et al (2013) Astrophys J Lett 766:L17
Schneider N et al (2015) Astron Astrophys 575:A79
SDP Initial Results. http://www.cosmos.esa.int/web/herschel/sdp-initial-results-presentations
Shinn T (2000) Revue Française de Sociologie 41:447–473
Shinn T, Ragouet P (2005) Controverses sur la science. Editions Raison d'agir, Paris
Small H (1973) J Am Soc Inf Sci 24:265–269
Small H, Griffith BC (1974) Sci Stud 4:17–40
Sneak Preview. http://www.cosmos.esa.int/web/herschel/sneak-preview
Spectroscopy Galore. http://www.cosmos.esa.int/web/herschel/first-spectroscopy
Sturm E et al (2011) Astrophys J Lett 733:L16
Tombesi F et al (2015) Nature 519:436
van Dishoeck EF et al (2011) Pub Astron Soc Pacific 123:138
Veilleux S et al (2013) Astrophys J 776:27
Wallace ML, Gingras Y, Duhon R (2009) J Am Soc Inf Sci 60:240–246
White HD, Griffith BC (1981) J Am Soc Inf Sci 32:163–171
Wu R et al (2015) Astron Astrophys 575:A88
Zavagno A et al (2010) Astron Astrophys 518:L101

Chapter 5
Innovation in Technology and Management

Abstract The implementation of *Herschel* as one of the most powerful space telescopes ever built was possible thanks to key inventions and innovations. They result from a long-term process, leading to the required scientific performances and an efficient mission management. The innovation process started during the design of the mission and finished with the final breath of the cryostat. This chapter examines the concept of innovation and apply it to the system design process and management that shaped the success of the *Herschel* mission.

5.1 What Is Innovation?

The word 'innovation' is on the agenda of every research policy. The European Commission set the objective, in 2010, of developing an 'Innovation Union' by 2020. The policy document asserts that competitiveness, employment and standard of living in Europe depend primarily on its ability to promote innovation, which is 'also our best means of successfully tackling major societal challenges, such as climate change, energy and resource scarcity, health and ageing, which are becoming more urgent by the day' (European Commission 2010). Tasked both with stimulating the economy and improving living conditions, innovation, as omnipresent as it may be in that document, is not defined anywhere. Its meaning is taken for granted and evident.

The verb 'innovate' was borrowed in the fourteenth century from the low Latin 'innovare' meaning 'to renew'. It was first used in legal jargon to mean transitioning something new into something established. The first use of the term in relation to the progress of science and technology can be found in the writings of the British philosopher and scientist Francis Bacon: 'On Innovations' is the title of one of the many chapters of his *Essays or Counsels, Civil and Moral*, published in 1625. As a pioneer of the scientific methodology, he was aware of the resistance to progressive ideas and keen to overcome them. In his *Essays*, he presented innovation through two analogies. The first one excuses the innovations for their nonconformity and weaknesses by comparing them with newborns. The second underlines the risky nature of innovation by expounding a tendency to deterioration. But Bacon's main argument is to relate this tendency not to the innovation itself but to the corrupting nature of time, insisting on the fact that innovation is therefore not only desirable,

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_5

but essential: ‘Surely every medicine is an innovation; and he that will not apply new remedies, must expect new evils; for time is the greatest innovator; and if time of course alters things to the worse, and wisdom and counsel shall not alter them to the better, what shall be the end?’ (Bacon 1985).

The most troubling thing in this text is that it barely seems to have aged: we could easily find Bacon’s phrases, almost word for word, in current considerations on the necessity for ‘responsible innovation’. The analogy with a newborn is a *leitmotiv* of reflection on responsibility in respect of new technologies, which we must protect and supervise during their turbulent youth before they reach maturity. The swing between recognising the uncertainty inherent in innovation and the certainty that never innovating might involve an even greater risk is found in every text addressing the ‘acceptability’ of new technologies.

However, despite the surprising contemporary relevance of Bacon’s reasoning, we must emphasise that, nowadays, innovation takes on specific meanings and acquires operative value because of its importance in the economic field. Innovation was defined in economics by Joseph Schumpeter as the ‘creative destruction’ of value: innovation consists in any new combination of production means, which increase the value of production and, as a result, profit. Schumpeter identifies five possible areas of innovation: modification of products (or services), the opening of new markets, change in production processes, the use of new energy resources or new raw materials, and, lastly, changes to a company’s organisation (Schumpeter 1939). Obviously, the modification of products and processes could refer to innovation originating in science (OECD 2005). But we may also think of changes in organisations, discoveries of new materials and energy sources and even the creation of new needs and markets, which occurred through scientific research or, more precisely, because of *inventions*.

In this respect, scientific research and technical invention are often the underlying conditions of innovation. Nevertheless, Schumpeter insists on the fact that ‘Innovation is possible without anything we should identify as invention and invention does not necessarily induce innovation’ (Schumpeter 1939). That said, he did not ignore the importance of invention. He even stressed that technological progress is the driver of creative destruction: ‘the impact of new things—new technologies for instance—on the existing structure of an industry considerably reduces the long-run scope and importance of practices that aim (. . .) at conserving established positions and at maximizing the profits accruing from them’ (Schumpeter 1998). In particular, some innovations induced by inventions stand out by the fact that they occur discontinuously, in a spate, and these ‘innovation clusters’ disrupt the production cycle: some products are eliminated (destruction) while others trigger new demands (creation). Eventually, all the different areas of innovation (products, processes, markets, resources and organisation) are modified by these disruptive innovations. Thus the Schumpeterian definition of innovation does not necessarily tie in with scientific research, but it acknowledges the importance of scientific and technological inventions to stimulate the innovation process.

That is why we must turn towards the history and philosophy of technical inventions if we want to better understand the issues of innovation in scientific research. Although a philosopher of technologies like Gilbert Simondon does not use the word 'innovation', he describes accurately and explains the mechanisms of 'invention', a word he uses to denote creation or global reorganisation of an object as compared to simple adaptations of pre-existing objects: 'Quite generally, relational progress are gradual and continuous improvements, obtained through trial and error during use; they result from experience and accumulate; they retain the pace of the relationship between the operator and its environment. Instead, the progress of the autocorrelation require problem solving, an invention that creates a synergistic system of compatibility' (Simondon 2005). This general and very robust definition of invention, which implicitly characterised the technological origin of disruptive innovation, is specified in the case of scientific inventions: 'The inventions led by an already constituted or in progress science are not adaptations nor only combinations of functions aimed to create a network or a machine in which each organ is multi-functional and subject to a cyclic functioning [like is the case in industrial inventions optimizing the yield of machines.] The critical issue now is the fidelity of the transfer of information, between the different organs of a machine or at very long distance; it imposes no distortion by transformations, and also the absence of interference signals; energy efficiency is subject to the rigor of the proportionality between input signals and output variables, whether within a transducer, an amplifier or even a computer' (Simondon 2005). That is why Simondon chose progress in the field of scientific instrumentation as the paradigm to think of invention. He notices, about detectors, in 1968, that 'one produces radiation receivers whose background noise is reduced as much as possible by lowering the temperature (operating near absolute zero) of the input stage' (Simondon 2005). Indeed, this short citation would nicely apply to the long-term effort that has led since early 1960s to the development of cryogenically cooled bolometers and heterodyne receivers.

At this stage, we may summarise what is relevant about innovation to address the outcome of the research process in the field of space science instrumentation:

1. Innovation is (always) a *risky change*, because it means giving up something that used to work, for the promise of something more or less *unknown*, but it is a necessary change, not only because of competition, but because it is the aim of research always to push further the horizon of technology and knowledge.
2. Innovation has a purpose, which is to *create wealth*, it being understood that the principal outcome of scientific research is not money—there are a lot of different 'symbolic capitals' (Bourdieu 1976)—but scientific knowledge, technological progress, social networks, symbolic images, cultural values, etc.
3. Innovation occurring in a scientific field like astrophysics has to be understood as a process of invention in the sense of Simondon: an innovation is the result of a problem-solving, which represents *major progress* in the concretisation of a technical lineage, which has a scientific finality. In order to understand innovation, we must then go through the structural analysis of the instrument itself,

through the historical analysis of the progress of the different technical lineages converging in the instrument and also through a conceptual analysis of its design.

But, before examining these different technological aspects in the case of *Herschel*, we must understand the social and institutional conditions of such an innovation.

5.2 Setting the Condition for 'Coopetition'

As described in Chaps. 2 and 3, *Herschel* is one of four Cornerstone missions of the Horizon 2000 programme that was implemented by the European Space Agency (ESA) from 1985 to 2009 with its launch with *Planck*. Horizon 2000 strongly structured space science in Europe by encouraging and coordinating scientific and industrial collaboration and therefore offered an international framework for inventing the most advanced space missions. Another key aspect of this programme was its conservative and reasonable approach of budget expenses. Clearly, it was not about inventing new missions at no cost, but rather designing to a well-defined cost. Horizon 2000 could be seen as the framework of innovation for European space science in the 1980s and 1990s with large missions flying since the launch of the first Cornerstone of the programme in 1995 (see Chap. 2). By virtue of its emphasis on 'scientific excellence, schedule respect, and design-to-cost' (Chap. 2 and Chap. 3), Horizon 2000 fulfilled one of the major characteristics for innovation projects, which must reach a clear objective defined in terms of delays, costs and performance (Middler and Lenfle 2003).

Within this programme, *Herschel* was managed in different ways: top-down with ESA as the agency in charge of the mission and bottom-up with the three consortia of research laboratories in charge of delivering the payload consisting of three instruments (HIFI, PACS and SPIRE) and of ensuring the maximum scientific return. Despite a common ESA framework, the consortia were not organised in an identical fashion. However, they all had to manage both collaboration for bringing together a large team—a sort of 'interstitial community' gathered around their respective instrument (Shinn and Ragouet 2005)—and competition between the scientists within the team. All consortia devised their own way of organisation in order to make instrument design choices among many concepts in competition and to ensure the most beneficial return to the scientific community and to the national agencies and ESA.

The sociologist Robert Merton used to call 'communalism' the norm of the scientific community prescribing that all scientists should have equal access to scientific goods and collaborate to a common goal (Merton 1942). But the sociology of science has since shifted towards analogies more inspired by the free market model, which are supposed to be less idealistic and more realistic (Ziman 2000). After the predominance of a precapitalist model, where scientists were supposed to be engaged in a sort of potlatch (Hagström 1965), the sociology of science has turned towards a capitalist model (Bourdieu 1976), in which scientists are competing to accumulate both symbolic and institutional capitals. Some sociologists of science argued for a new model that could explain the propensity among researchers to cooperate by networking in order to win competitions (Latour and Woolgar 1979). Yet their model does not explain how some disinterested actions and mutually benevolent collaborations happen in such a complex and highly competitive environment. Since some occurred during the *Herschel* mission, we have to propose another model to understand the innovation management.

The management of the *Herschel* mission was indeed characterised by a process of international *coopetition*. This term has been coined by specialists in game theory (Nalebuff and Brandenburger 1996) to designate a strategy which combines cooperation and competition. Most economists find this type of paradoxical behaviour in business rather strange and unusual, but acknowledge that coopetition is optimal when the actors must adapt to a very competitive environment but still have to collaborate in collective action (either to prevent other competitors from emerging or to pool their resources and exchange knowledge) which is a very general situation in the scientific field. A coopetition situation has three characteristics (Bengtsson and Koch 1999):

1. It involves *complex relations* between actors who are potential collaborators and competitors.
2. It occurs only when there is a relative *separation between the phases* of cooperation and competition through a sharing of tasks and responsibilities and/or with an alternation of the phases.
3. It requires a *regulatory body* whereby the actors can confront their claims and can develop a code of conduct (see Chap. 10).

All these characteristics are easy to find in a case such as *Herschel*: the research teams have to cooperate in order to pool resources which otherwise would be difficult to assemble while also maintaining a comparative advantage and protect themselves from the opportunism of their partners-competitors; the mission emulates competition at some stage, e.g. before the selection of the instruments, and coordinates also the division of work at some others; the governance of the ESA, the explicit and implicit rules of the scientific field (Bourdieu 1976) and the many feedback mechanisms between all the different agents during the project are aimed at securing confidence between these agents and arbitrating their possible conflicts. Thus coopetition appears to be a compromise between the extreme economic and scientific norms, i.e. between the ruthless struggle for life and the irenic 'union of workers of the proof' (Bachelard 1949), which guarantee that innovation will be rewarded.

5.3 Promoting Sociality Between Science and Industry

According to Merton, another norm of the scientific community is ‘universalism’, meaning that nobody should be discriminated against by unscientific criteria so that everybody is allowed in principle to be a member of the scientific community (Merton 1942). There are obvious counterexamples in the history of science, but the point is that when such arbitrary discriminations interfere with the normal functioning of scientific research, they are harmful to the progress of knowledge. The autonomy of a scientific field is always relative, but it is a necessary social condition of its effective functioning (Bourdieu 1976).

However, the norm of universalism only applies to exchange *within* the scientific field. From the moment when scientists have to collaborate with people outside of the scientific field, they must deal with different values and take into account other requirements than their own. This is the case when scientists have to collaborate with industry in order to achieve the implementation of instrument such as *Herschel*. Scientists and industrialists have to learn how to deal with people who are different from them and have to manage different kind of relationships to work together.

The norm of ‘sociality’ has been defined by the philosopher and sociologist Richard Sennett to describe this way of collaborating for a common goal with people who have different values and norms (Sennett 2012). Hence sociality is different from solidarity which applies only when people share the same values and norms. It implies in particular the development of a form of ‘dialogical communication’ instead of the ‘dialectical communication’, which prevails in the scientific field: when scientists discuss together, they try to prove their point and to discard other points of view because truth is going to emerge from this conflictual process, but when scientists and industrialists have to interact, first, they must translate the arguments of the other side in their own words in order to avoid misunderstandings, and, second, they must set the condition for compromises. This process is therefore rather complex, recurrent and potentially frustrating, but it is essential to the success of the social construction of the mission. The understanding of this dialogical process between the scientific field and the industrial field requires us to identify and study some key-actors, which were able to represent their pairs' interests and, at the same time, to cross boundaries in order to circulate between these two different social worlds and to translate the expressions of these two ‘universes of discourse’ (Clarke and Leigh-Star 2008).

In the *Herschel* mission, such a complex and difficult role has been played notably by the Project Manager, Thomas Passvogel. According to the PACS PI, Albrecht Poglitsch, Passvogel understood how the scientific and the industrial fields work and ‘how [he] had to talk to them to get what [he needed] and what [he wanted], i.e. how to put pressure on different parties. For example, with industry, [he] could use money or contractual arguments, but with the instrument teams it did not work the same way. He said things like “OK, if you are not ready on time, then we will fly without this part of your instrument”. So, he was prepared to use all kinds of threats, but only to the extent where it did not do any real damage to the science or to the mission. Then I think his judgement was really good. I think that the important point was that he was in constant communication: he could ask questions again and again and find out where the real issues were and what one could accept and what one could not. He had the right combination of the understanding of the scientific side but also the habits of dealing with the "real world"' (Poglitsch 2015).

5.4 Concepts and Knowledge Leading to Innovation

The first steps of the coopetition process in the *Herschel* mission occurred during the early design phase (see Chap. 3), to be understood as the phase during which many mission concepts are proposed by the leaders of the original idea to the scientific community in order to gain their support. This community is not of one mind, but rather diverse with many different interests, for instance, radio astronomy for interstellar molecule detection or infrared astronomy for imaging interstellar dust (see Chap. 4). Many mission concepts were proposed in the 1980s and early 1990s, such as the Large Deployable Reflector in the USA, which consisted of 20-m diameter dish that was planned to be deployed in orbit by astronauts (and was abandoned after the Challenger accident in 1986). The Submillimeter Explorer was a less ambitious concept with a 4 m dish. It was supported by a collaboration led by Tom Phillips at the California Institute of Technology (Caltech) including French laboratories and the Centre National d'Études Spatiales (CNES). In the framework of ESA, FIRST, the Far-Infrared and Submillimetre Space Telescope, was initially conceived in the early 1980s. It was originally planned to have an 8-m class deployable dish as its main telescope mirror. We may consider that this ambitious project was an *undecidable proposition* relative to the knowledge available at that epoch. Indeed it was not possible to prove that this proposition was true or false if confronted with all currently scientific, technical, economic or social constraints.

The notion of 'undecidability' originates in mathematical logic (Gödel 1931) and is well defined in a number of theories. We borrow the concept from the C-K design theory, which is schematically illustrated in Fig. 5.1: 'Design can be modelled as the interplay between two interdependent spaces with different structures and logics: the space of concepts (C) and the space of knowledge (K). The structures of these two spaces determine the core propositions of C-K theory' (Hatchuel and Weil 2003). Here the space of concepts means the propositions and ideas that are undecidable. Space K (Knowledge) contains all established propositions: 'the pragmatic view of design describes a dynamic mapping process between specifications and solutions. However, it is clear that this approach fails to account for the expansions occurring in space C and in space K during the actual process' (Hatchuel and Weil 2009).

The C-K theory applies very well to large science instruments in particular because it takes into account the moving environment during their long-term design process. Science evolves with new discoveries (re)defining instrument performances or even making acceptable some descoping in ambition. Thus technological research is modified by external inputs into the K space. But conceptual researches in the C

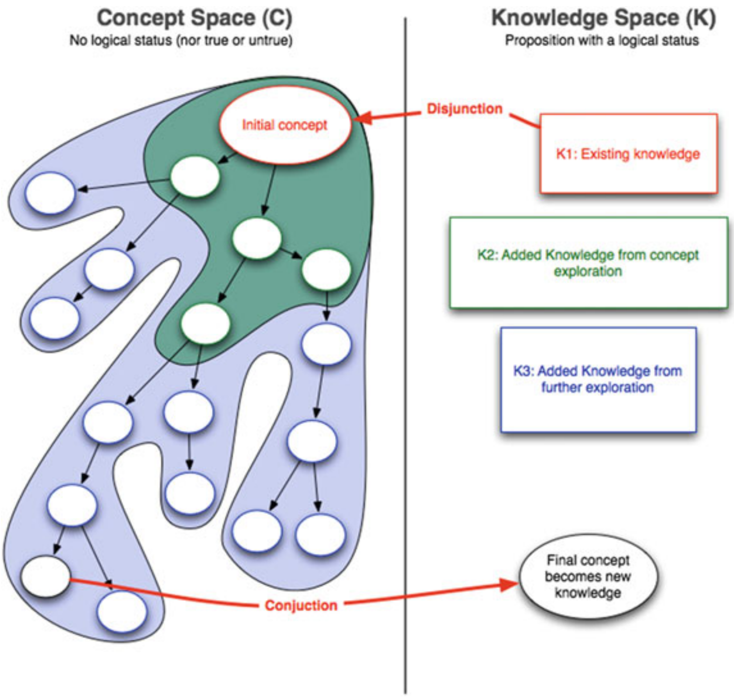


Fig. 5.1 The C-K dynamics: a space of concept starting with the initially undecidable concept and evolving through its interactions with a space of knowledge. © Ingi.b via Wikipedia under CC Attribution-ShareAlike 3.0 license

space also induce in return the need for new scientific or technological researches in the K space. European space policy and budget tend to be stable while aerospace companies may change their priorities or strategy. Against a product design theory that gives recipes to develop a product in conformity with the dominant design, the C-K theory investigates the risky and undecidable nature of an innovation project at the beginning, the irreversible choices as necessary conditions of its achievement and the open and moving configuration of the project with numerous insiders' and outsiders' inputs at all times. C-K theory can be used to study both the past realisations and research development on which *Herschel* was based and more importantly the dynamic evolution of these scientific, technical and social bases.

5.5 Mechanology and the Pace of Innovation

In *On the Mode of Existence of Technical Object*, the philosopher of technology Gilbert Simondon elaborates a methodology for the study of machines and their evolution, namely, the 'mécanologie génétique' that is translated as the 'genetic mechanology'. Mechanology is the science of machine (Lafitte 1932), and

Simondon defined it as *genetic* because the analysis of the technicity should be based on the genesis of each technical object within a technical lineage. This means that one should define technical objects and lineages according to the invention of their functioning. Most of the technical objects are usually named with reference to their function and therefore these designations refer to practical purposes. Their names do not reflect their operations, which define their true technicity. Therefore the ordinary classification of technical objects according to their functions is illusory because no fixed structure corresponds to its defined use. The same outcome can be obtained from very different operations and structures. For example, a steam engine, a petrol engine, an electric engine and a spring engine are named 'engine' because they are all defined by their function (to set in motion) and, yet, there is more apt analogy between a spring engine and a crossbow than between a spring engine and a steam engine. Usage brings together heterogeneous structures and operations in artificial classes that get their meaning not from the technicity of the objects but from the relationship between their function and another external function, usually that of the human operator. We identify four principles that help understanding the method of mechanology:

1. The first principle of genetic mechanology is to identify technical objects according to their *functioning* rather than their function.
2. The second principle of genetic mechanology is always to consider a technical object in relation with its 'associated environment', i.e. the external conditions of its functioning.

The associated environment is both natural and artificial. For example, *Herschel* was designed to operate in space, more precisely at a specific position in space (an orbit around the second Lagrange point), and also needed to be cooled in order to perform its task. But to simply define *Herschel* as a 'machine' with only one associated environment is too vague.

3. The third methodological principle is to distinguish between different levels of organisation within a technical system: the elements, the individuals and the sets (ensembles). These different levels of the technical systems are defined according to the type of relationship they have with the associated environment.

Elements are rather indifferent to the nature of the environment and may therefore be integrated within any machine: they can be spare parts that you find in cars or in boats or, in the case of *Herschel*, some components that can be used in different parts whether these parts are cryogenically cooled or not. On the contrary, individuals are precisely the technical systems which function only in a specific environment. They can even be in a situation of 'circular causality' with their associated environment when their functioning contributes to produce the environment, which is at the same time the condition of their functioning. For example, *Herschel*'s instruments are individuals: they perform an operation that requires a very specific environment. Finally, the sets are the devices whose functioning implies a relative separation between several different environments. As a matter of fact, *Herschel* is a technical set: for example, its functioning required that some of its individuals are cryocooled while others are not.

4. The fourth methodological principle is to identify the *technical lineage*, which the technical objects belong to, and to study the progress of its *concretisation*.

The technicity emerges within an evolution, which started with the first invention of a specific technical object. For example, the technicity of the petrol engine cannot be understood only through the study of any particular engine ever produced but by considering the sequence or continuity from the first engines to those we know and to those that are still in evolution. Consequently, as in a biological lineage, a defined stage of technicity contains within itself structures and dynamic schemes that evolve since the beginning of the technical lineage. For example, *Herschel*'s instruments are technical individuals belonging to different lineages. A technical lineage evolves by convergence and by adaptation of the object to itself; each object is unified according to a principle of internal resonance. This evolution process is called 'concretisation' because the first objects were abstract in the sense that their parts were operating independently and thus their unity was an abstraction while the most evolved objects are concrete in the sense that there are strong synergies between all their parts and that they operate as a quasi-organic whole. In the next chapters, we study some of the technical lineages whose concretisation has made *Herschel* possible.

To sum up, the technicity of a technical object is defined by its *functioning*, by the relationship of its different level of organisation with its *associated environments* and by the degree of *concretisation* within its *technical lineages*.

To bring about this understanding of machines, it is necessary to perform a structural analysis of their internal organisation and insertion within the environment and also a historical analysis of the process of concretisation of their technical lineages. The technical lineages proceed with ramifications and selections: technical evolution is sometimes proliferating and sometimes restrained. All these technical evolutions are not necessarily concretisations: they sometimes lead to 'hypertely' (a mechanological term borrowed from biology meaning an over-adaptation to a single task in a very specific associated environment), but in general the more evolved a machine is the more 'generic' it is, i.e. the better it adapts to variations of the environment because it has integrated its own associated environment within itself. Typically, technical lineages show phases of continuous progress alternating with other phases of 'saturation' during which it is impossible to improve the performance in one way without diminishing them in another way (Simondon 2012). To overcome this phase of saturation, major improvement must emerge as a global reconfiguration. For example: 'Contradictory requirements were set forth in the design of the steam boiler (a boiler must have a spherical, or cylindrical, shape in order to provide the needed strength under high steam pressures; however, these shapes create a minimal heating surface) that therefore lower the production of steam. To satisfy these requirements the cylindrical shape was preserved, but its length was increased. The boiler-vessel slowly transformed into a system of pipes with a large total heating surface' (Altshuller 2007). These thresholds and ruptures of gradual evolution give a punctuated pace to the historical evolution of technical lineages. Thus the modelling of technical lineages may provide a better understanding of the pace of innovation.

This modelling implies the identification of the origins of each lineage, i.e. the first object, which is ‘abstract’ in the sense that its components are functioning independently. This abstract object is followed by major steps of concretisation, which means that the functioning of the technical object becomes more and more ‘concrete’ as its structures evolve towards synergy, self-regulation and convergence of functions. This process can result from minor progress, gradual enhancements and adaptations of technical elements; but the true major improvements occur during inventions or reconfigurations of the global structure, which do not represent compromise, but rather resolutions of incompatibilities between subsystems. True technological progress does not follow from optimisation but from problem-solving, implying a recombination of the subsystems in a way that they collaborate instead of opposing each other: ‘The technical problem is thus one of the convergence of functions into a technical unit, rather than one of seeking a compromise between conflicting requirements’ (Simondon 2012). Besides that, the technical lineage has an end, i.e. a final saturation when concretisation has led to a totally unified technical system.

The concretisation of technical lineages proceeds level by level, from one systematic configuration to another, and gradual evolutions may appear during the stable periods at each level. Recurrent transformations between the levels give information about the logic of the progress and suggest that a specific pace of evolution should be determined for each lineage (Simondon 2012; Altshuller 2007; Bontems 2009). This pace is not the direct result of a law of nature since technical evolutions are not natural evolutions but artificial ones. So the modelling of a regular pace of evolution is a tool for forecasting, not a rule for prediction (Bontems and Le Roux 2009): it is a probabilistic foresight that can be modified by human decisions, including those based on this foresight, but not a deterministic prediction similar to the evolution of a closed physical system. However it may take the form of a specific relaxation law: ‘Such a pace of relaxation has no equivalent anywhere else; neither the human world, nor the geographic world can produce such an oscillation with successive crises and emergence of new structures’ (Simondon 2012). The new structures are not only the reconfigurations inside the technical lineages but also the results of bifurcations and substitution between them. When a saturation occurs, a technical lineage may be transformed, but it may also be abandoned and replaced by new ones with different functioning processes (see Chap. 7).

5.6 Typology and Census of Innovations

Together, the C-K theory describing the conception phase, Simondon’s mechanology analysing the machine through its invention and concretisations and the sociological concepts of coopetition and sociality characterising the innovation framework of a space mission are the three conceptual pillars of our analysis of the *Herschel* Space Observatory. This allows us to describe various types of innovation.

As discussed, innovation is a very loose concept when it is not further defined according to the type of invention involved in the progress of a technical lineage. In effect, since innovation may designate any modification, the first type of innovation could be even a regression, i.e. the replacement of a technology by a less advanced but probably more robust one. However, as no regression was identified in *Herschel*, we can focus our analysis on real technological progress such as concretisation of the technical object, i.e. when 'the technical object progresses by interior redistribution of functions into compatible units, eliminating antagonisms that may occur in the initial distribution; this specialisation does not occur function by function but synergy by synergy; it is the synergetic group of functions and not the individual function, which constitutes the real sub-set in the technical object' (Simondon 2012). Thus we define different types of innovations occurring in the lineages within *Herschel* according to the degree of synergy between structural elements performing several functions and also between the object and its associated environment.

1. *Customisation*: This is the adaptation of a pre-existent technology to a new environment. An over-adaptation induces a loss of genericity, which means that the instrument will not work as well or not at all in another environment or if a strong variation of the environment occurs. For example, the cryostat of *Herschel* is a customisation of the cryostat developed for the Infrared Space Observatory launched in 1995.
2. *Synergy*: This is the coupling of two different pre-existent technical lineages to perform a common task, sometimes in a new environment. It is more than just a customisation because there is also an increase of synergy inside the object but without any real invention since the two subsystems pre-exist. For example, the heterodyne receivers were first designed for ground-based application, but then subparts were specifically qualified for use in space.
3. *Incremental innovation*: Incremental innovation is a gradual progress of a technical lineage. It is an improvement which diminishes the harmful consequences of residual antagonism and increases performance without modifying the division of functions inside the object. For example, the SPIRE bolometer camera is an incremental step with respect to the feedhorn coupled bolometer cameras previously deployed on Earth-based telescopes.
4. *Disruptive/breakthrough innovation*: Disruptive innovation is a major progress of a technical lineage, either realised by a global *reorganisation* of the object or by a *substitution* of a new lineage to another to perform the same task. It usually occurs when a certain organisation of the object has previously reached a point of *saturation*, which means that it was impossible to increase the performance of the object in one sense without decreasing it in another. For example, the PACS bolometer camera with pixelised arrays proposed a new detector architecture with better instantaneous coverage of the image plane.
5. *Radical innovation or creation*: This is the invention of a new technical lineage. For example, silicon carbide as new material for the *Herschel* mirror overcame the strong specification on weight, resistance and performance.

5.7 Plurality of Functionalities and Relativity of Criteria

It should be noted that the benefits of an innovation are always relative because there is no single criterion to measure the performance of a technical object. Indeed the choice of the relevant criterion depends on the functionality considered. Functionality is not the same as function; it characterises the different relations between the internal functioning of a machine and its associated environment. To understand this point, we must go back to the work of the French architect and philosopher Jacques Lafitte, who introduced a distinction between *passive* machines that function without any movement, such as architectural buildings; *active* machines, which move and function with some energy transfer, such as engines; and *informational* machines that transfer information and often involve a form of feedback, such as a self-guided torpedo (Lafitte 1932). But a certain machine may belong at the same time to different types. For example, in the resonant cavities of PACS, the indium bumps have three functionalities: they are *passive* as elements of its design, but they are also *active* insofar as they evacuate the heat and even informational because they conduct the signal.

Simondon shows that passive functionalities rely on the absorption of perturbations from the environment that an active functionality is linked to the yield of transmission of energy and, finally, that informational functionalities depend on the amplification of information (Simondon 2005). Engineers implicitly use a similar typology when they test a mechanical model and some operational models. Simondon has also combined these passive, active and informational functionalities with the analysis of the relations between orders of magnitude involved in the working of a machine: some machines work because their functioning depends only on structural relations between their parts at their own scale (simple scale relation); some others work insofar as they rely for their operation on a relation between two scales, i.e. their functioning implies a relation between their own structure and the microstructure of their material (median scale relation); finally, some machines work thanks to an operation between many scales, i.e. when they establish a relation between their microstructure, their structure and the associated environment at a bigger scale (major scale relation).

The analysis of the complexity of the different functionalities of a machine may help to define the relevant figure of merit for evaluating the performance of each object.

1. A simple passive functionality is defined by a degree of *solidity*.
2. A median passive functionality is defined by some *qualities of the material*.
3. A major passive functionality is defined by the *stability* (resistance to perturbations).
4. A simple active functionality is defined by the yield of the *mechanical work*.
5. A median active functionality is defined by the yield of the *energetic conversion*.
6. A major active functionality is defined by some aspects of *motion*.
7. A simple informational functionality is defined by the *fidelity and speed* of transmission.

8. A median informational functionality is defined by the *sensitivity and precision*.
9. A major informational functionality is defined by the *complexity* of the data processing.

Using this classification, we can examine the specifications of a technological device in terms of its functionalities and try to determinate which ones are the more relevant to characterise its functioning, in order to choose the right criterion or combination of criteria to measure the pace of concretisation. Based on all these concepts, different types of innovation within *Herschel* will be described and analysed according to different types of functionalities: the telescope mirrors, the cryostat, the bolometer arrays for PACS and SPIRE and the heterodyne receivers in HIFI (Fig. 5.2).

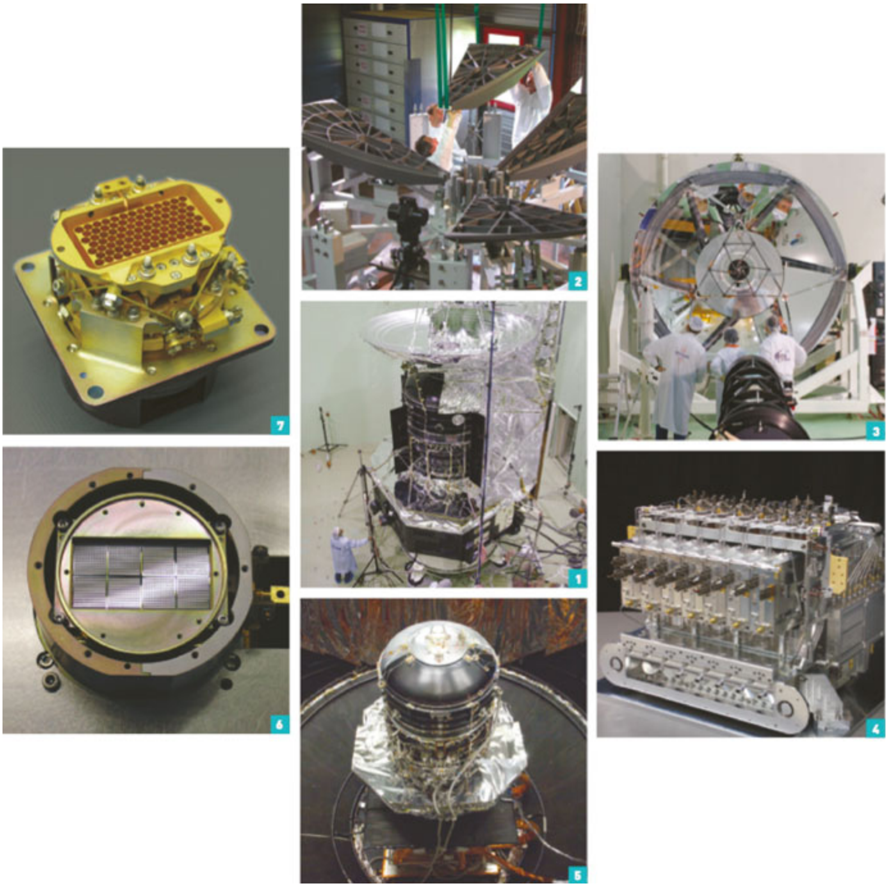


Fig. 5.2 *Herschel* and its innovative systems. (**1**) The *Herschel* Space Observatory. (**2**) The telescope SiC segments. (**3**) Telescope optical alignment. (**4**) The HIFI instrument. (**5**) The cryostat. (**6**) One of the two PACS bolometer arrays. (**7**) The bolometer feedhorn array of the SPIRE imager. © CEA (V. Minier, S. Leroy)

References

Altshuller G (2007) The innovation algorythm. TRIZ, systematic innovation and technical creativity. Technical Innovation Center, Worcester

Bachelard G (1949) Le Rationalisme appliqué. Presses universitaires de France, Paris

Bacon F (1985) The essays or counsels, civil and moral, of Francis Lord Verulam Viscount Saint Albans. Clarendon Press, Oxford

Bengtsson M, Koch S (1999) Cooperation and competition in relationships between competitors in business networks. J Bus Ind Mark 14(3):411–426

Bontems V (2009) Gilbert Simondon's genetic "mécanologie" and the understanding of laws of technical evolution. Techne 13(1):1–12

Bontems V, Le Roux R (2009) Objectivité et normativité de l'évolution technique chez Gilbert Simondon. In: Aït-el-Hadj S, Boly V (eds) Les Systèmes techniques: lois d'évolution et méthodes de conception. Hermès, Lyon, pp 23–35

Bourdieu P (1976) The specificity of the scientific field and the social conditions of the progress of reason. Soc Sci Inf 14(6):19–47

Clarke AE, Leigh-Star S (2008) The social worlds framework: a theory/methods package. In: Hackett E, Amsterdamska O et al (eds) The handbook of science and technologies studies. MIT Press, Cambridge, MA, pp 113–137

European Commission (2010) Europe 2020 Flagship Innovative Innovation Union. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Brussels

Gödel K (1931) Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte für Mathematik und Physik 38:173–198

Hagström WO (1965) The scientific community. Basic Books, New York

Hatchuel A, Weil B (2003) A new approach to innovative design: an introduction to C-K theory. ICED'03. Stockholm, 14 Aug 2003

Hatchuel A, Weil B (2009) C-K design theory: an advanced formulation. Res Eng Des 19:181–192

Lafitte J (1932) Réflexions sur la science des machines. Bloud & Gay, Paris

Latour B, Woolgar S (1979) Laboratory life: the social construction of scientific facts. Sage, Beverly Hills

Merton RK (1942) Science and technology in a democratic order. J Leg Polit Sociol 1:115–126

Middler C, Lenfle S (2003) Gestion de projet et innovation. In: Mustar P, Penan H (eds) L'Encyclopédie de l'innovation. Economica, Paris, pp 49–69

Nalebuff BJ, Brandenburger AN (1996) Co-opetition 1. Revolutionary mindset that redefines competition and co-operation 2. The game theory strategy that's changing the game of business. Doubleday, New York

OECD (2005) Oslo manual, guidelines for collecting and interpreting innovation data, 3rd edn. OECD & Statistical Office of the European Communities, Paris

Poglitsch, A (2015) Interview of Albrecht Poglitsch by Vincent Bontems and Vincent Minier, 17-12-2015

Schumpeter JA (1939) Business cycles: a theorical, historical and stastistical analysis of the capitalist process. McGraw-Hill Book, New York

Schumpeter JA (1998) Capitalisme, Socialisme et Démocratie. Payot, Paris

Sennett R (2012) Together: the rituals, pleasures and politics of cooperation. Yale University Press, New Haven

Shinn T, Ragouet P (2005) Controverses sur la science. Pour une sociologie transversaliste de l'activité scientifique. Raisons d'agir, Paris

Simondon G (2005) L'Invention dans les techniques. Cours et conférences. Seuil, Paris

Simondon G (2012) Du Mode d'existence des objets techniques. Aubier, Paris

Ziman J (2000) Real science. What it is, and what it means. Cambridge University Press, Cambridge

Chapter 6
Silicon Carbide Telescope: Radical Innovation

Abstract This chapter reviews the development of the *Herschel* telescope and shows how the initial concept of FIRST (renamed *Herschel* in 2000) was 'undecidable' (not capable of being told feasible or unfeasible) due to the poor knowledge of how to build it following its introduction in the Horizon 2000 long-term plan. It describes the competition between various industrial concepts, the abandoned NASA concept and how the development of silicon carbide (SiC) mirror technology made it possible to build the largest telescope then ever flown, just fitting in an Ariane 5 fairing. It offers a reflection on the need of major innovations requiring advanced material technologies and innovative mirror structures to go from *Herschel* to the next generation of high-resolution far-infrared space telescopes.

Much of astronomy is driven by the quest for higher sensitivity and improved angular resolution in order to detect fainter or smaller objects. These requirements lead to the design of large telescope mirrors ensuring high angular resolution (which improves in proportion to the telescope diameter) and light-collecting power (which improves in proportion to the telescope area). Ground-based optical/infrared telescopes now have primary mirrors up to 10 m in diameter, and new observatories are being built with telescopes up to around 40 m in size. But space telescopes are comparatively smaller because of the limitations on the size and mass of payloads that can be launched. *Herschel* with a monolithic primary mirror of 3.5 m in diameter protected by a sun-shield just fitted inside an Ariane 5 payload fairing. For larger sizes, the telescopes must be deployable. Launch mass and loads are also important drivers and impose significant constraints on the kind of orbit into which a satellite can be placed. Space telescopes therefore require advanced material technologies and innovative mirror structures in order to satisfy scientific requirements and launch specifications.

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_6

6.1 Telescope Principle

A telescope focuses electromagnetic radiation onto instruments, such as photometers, spectrometers or radio receivers. To do so, astronomical telescopes consist of a series of mirrors, the main ones being the primary mirror which faces the sky and intercepts the incoming radiation, and the secondary mirror which directs it towards the scientific instrument(s). The light may then be guided by additional mirrors prior to and inside the instruments before reaching the detector(s). However, increasing the path length and the number of mirrors induces some losses due to the imperfect reflection of the mirrors. In the case of the *Herschel* telescope, a Cassegrain telescope design (Fig. 6.1) was adopted where a hyperbolic secondary mirror reflects the light collected by a parabolic primary to the instruments through a hole in the primary (Toulemont et al. 2004). The overall telescope transmission reached well above 99% for the longest wavelengths.

The sensitivity[1] of the complete system (telescope plus instrument) is a measure of the minimum signal that can be distinguished above the random background

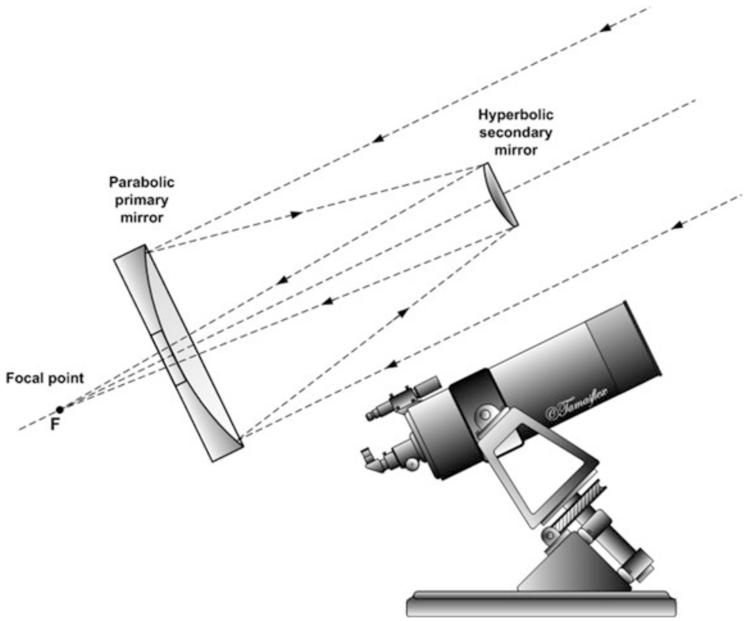


Fig. 6.1 Typical astronomical telescope: Cassegrain reflector consisting of a parabolic concave primary mirror and a hyperbolic convex secondary mirror. Detectors are located at the focal point or focal plane in case of detector arrays. The telescope can be placed within a tube to protect it from side light sources. © Szőcs Tamás Tamasflex via Wikipedia under CC Attribution-ShareAlike License

[1]To derive the full system sensitivity, many other parameters must also be taken into account such as the intrinsic sensitivity of the instrument or background emission from the instrument and natural environment (atmosphere, space), and (square root of) integration time that will improve it.

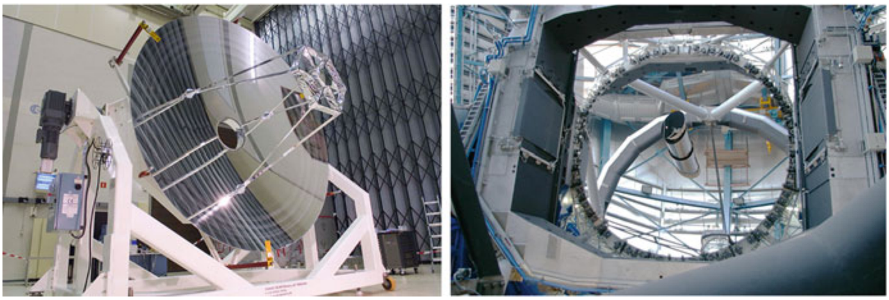


Fig. 6.2 *Left*: the *Herschel* telescope displaying its primary and secondary mirrors. © ESA/*Herschel*. *Right*: One of the four VLT telescopes based on the ESO Paranal site in Chile. © ESO

noise. The diameter of the primary (D) is a key parameter in determining the sensitivity, since light-gathering power, and so overall sensitivity, improves proportionally with D^2. Angular resolution is another key characteristic of a telescope. It represents the ability to distinguish between different nearby celestial objects on the sky or, alternatively, the size of the region of sky which a detector sees. The angular resolution is proportional to λ/D, where λ is the wavelength of the electromagnetic radiation being detected.

These basic telescope principles apply to optical, infrared and radio telescope designs. Depending on their size, the mirrors may be protected by a tube. For large telescopes, such as *Herschel* or the Very Large Telescope in Chile, the mirrors are 'naked' (Fig. 6.2).

6.2 FIRST: An 'Undecidable' Initial Concept

In 1982, the astronomical community supporting FIRST (ESA 1982) proposed to equip a satellite dedicated to the far-infrared observation of the universe with a telescope including an 8-m diameter and self-deployable primary mirror (see Chaps. 2 and 3). At that time, it was an 'undecidable' concept as there was no immediate knowledge of how to build it (see Chap. 5 for a definition of this concept). The 8-m diameter was a desirable size for the new and exciting study of the cold interstellar material (ESA 1983), and FIRST was expected to open the last major electromagnetic window (wavelengths between 100 μm and 1 mm) onto the universe and allow key investigations of the origins and evolutions of planets, stars and galaxies (see Chap. 4). It was the main scientific motivation.

Providing the European astronomers with a great observatory was another motivation. At the same time in the USA, NASA planned an impressive series of space experiments for infrared astronomy and cosmology with estimated launch

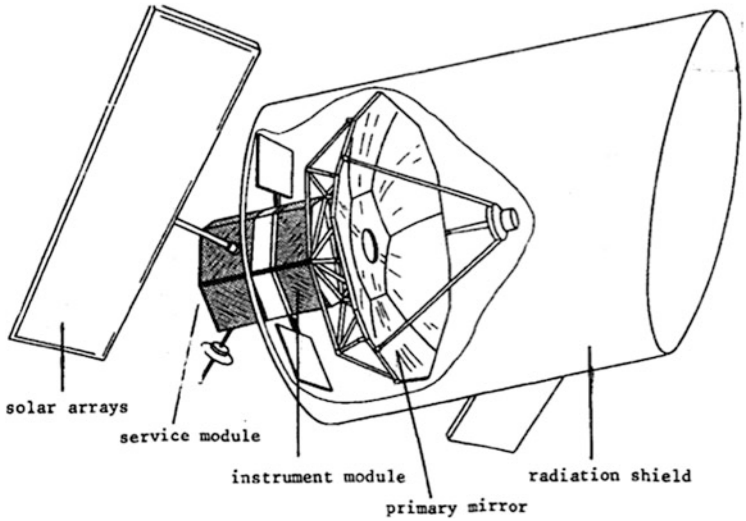


Fig. 6.3 FIRST satellite in 1982 from the FIRST mission proposal to ESA. It consisted of an 8-m diameter deployable telescope placed inside a giant inflatable radiation shield. © ESA

dates in parantheses: IRAS (1983), IRT/Spacelab 2 (1984), COBE (1987), SIRTF (1990) and LDR (1990s) (Fazio 1983). With the exception of IRAS, many missions were either postponed or dismissed. Nonetheless to some extent, they paved the way for *Herschel* and the NASA contribution to it. In Europe, similar ambitions required bringing together the scientific community, the space industries and using the growing capability of the Ariane rocket. The original FIRST proposal was very ambitious (Fig. 6.3). The proposed operating wavelength range implied a surface accuracy of about 8 μm rms ($\lambda/13$) for the completely deployed primary reflector.[2] Only rigid-panel reflectors could have achieved this specification. State-of-the-art technology was supposedly available in Europe, within Dornier in collaboration with the Max Planck Institut für Radioastronomie (Bonn) and the Optical Sciences Center (Tucson, USA). They manufactured graphite/epoxy panels with a surface area of 1.5 m^2 and surface accuracy of 4 μm. However, as mentioned in the 1982 proposal (ESA 1982), the transferability of this accuracy to larger panels had to be verified, but would be feasible. The telescope deployment concept appeared as well as an unknown parameter of the FIRST telescope because it had never been achieved in space astrophysics before. Previous studies under NASA contract

[2]The root-mean-squared (rms) accuracy of a mirror provides a statistical measure of the departure of the surface from the ideal shape. Note that the FIRST proposal indeed stated 8 μm rms surface accuracy; however this will give you almost 16 μm wave front error (WFE) rms. This does not provide good performance and it was an error in the proposal.

backed up the idea, proposing a tetrahedral truss support framework for the main reflector. It could fit within Ariane 3-4 payload fairing, assuming a central panel of 3.6-m diameter with the outside triptych panels hinged onto the central one.

The proposed telescope concept evolved in the assessment study in 1983 and finally became the reference concept in the Horizon 2000 programme. New technical studies constrained it. The telescope, described as the antenna,[3] still included an 8-m self-deployable primary reflector with 12 outer panels that could be folded to fit in the Ariane fairing. The panel design consisted of an aluminium honeycomb sandwiched between two thin layers of carbon-fibre reinforced plastic (CFRP). It was motivated by ‘good experience of this technology in Europe, also for space application, since such panels are being used for antennae for communications and application satellites’ as described in the assessment study. Again Dornier was mentioned as the expert company as well as other European companies like Messerschmitt-Bölkow-Blohm GmbH (MBB) in Munich. Other materials were considered such as glass and ceramics, but were eliminated due to their higher areal density[4] (>30 kg/m^2 compared with the approximately 7 kg/m^2 for CFRP). Beryllium technology, later used for the James Webb Space Telescope deployable mirror, was judged attractive as an alternative to glass, but was not developed in Europe.

Based on this description, four types of factors and chronological events constrained the earliest telescope design. These four types of factors/events will be described as ‘knowledge’ as defined in the C-K theory (see Chap. 5) and be represented as four timelines summarising the key milestones constraining the evolution of the telescope design:

- Scientific and technical motivations (with a promising electromagnetic window to study the infrared universe at high angular resolution)
- Technological feasibility (with emerging capabilities in designing large space mirrors both in Europe and the USA)
- Emblematic space projects (bringing new expertise in space industries)
- A programmatic framework (with the European space community willing to organise itself through call for proposals and the forthcoming Horizon 2000 programme)

In the next sections, the tortuous road to the *Herschel* telescope is summarised and how the industrial competition regulated by ESA and science objectives (see Fig. 6.4) led to the selection of a backup technology over all established technologies that has ultimately become a radical innovation.

[3]A term used more commonly in radio astronomy than in optical/infrared astronomy.

[4]The areal density is the ratio of the weight to the surface area of the primary mirror.

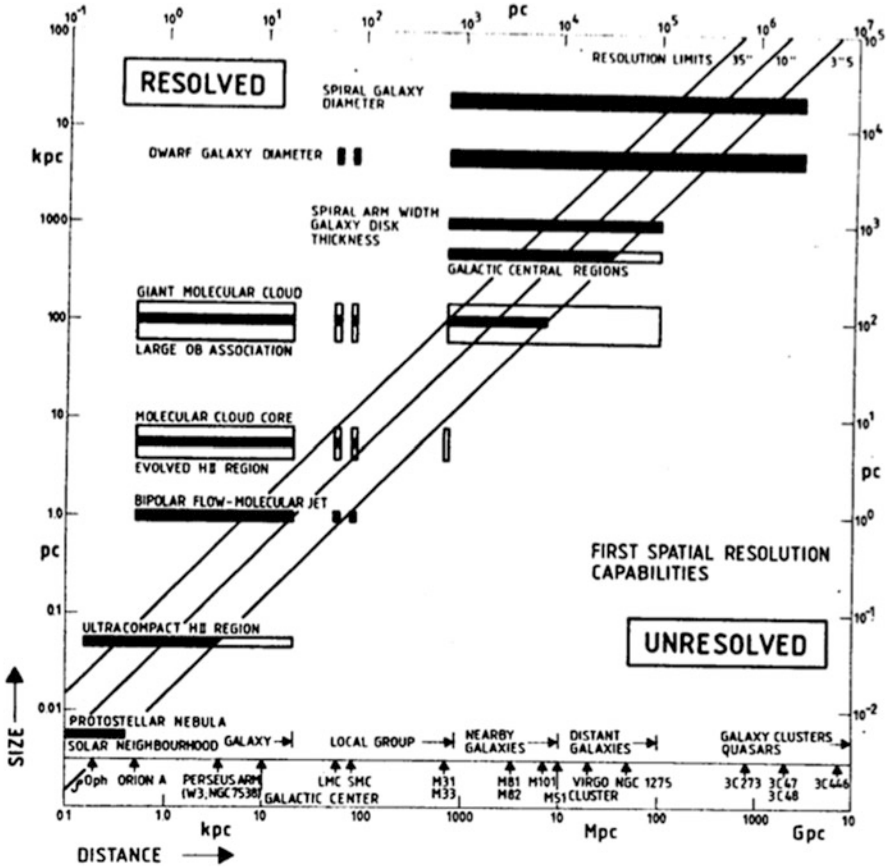


Fig. 6.4 Examples of astronomical objects that could be observed by FIRST as a function of distance and size. The *diagonal lines* indicate typical FIRST resolutions in 0.1–1 mm wavelength range. At the *bottom* of the diagram, typical representatives of each class of objects are given. Source: Fig. 2.1 in FIRST assessment study (ESA 1983). © ESA

6.3 Competition Between Various Industrial Concepts

Between 1985 and 1993, various studies and meetings modified the initial telescope concept, with the major change being a decrease of the primary mirror diameter from 8 m via 4.5 to 3 m. The Segovia meeting in 1986 confirmed the deployable concept, but then the Science Advisory Group (SAG) studied the potentiality of FIRST to remain a Cornerstone mission of the Horizon 2000 programme leading to trade-off between scientific requirement, technical complexity and cost (see Chap. 3). Two concepts were discussed and analysed: the initial 8-m telescope and a 4.5-m single-panel design. In December 1990, the SAG recommended on the basis of technical risk, total programme cost and scientific priorities that the telescope should have a diameter of 4.5 m with a passive surface and a diffraction limit at 150 μm. Two potential concepts were considered, a multi-panel mirror studied by Matra Marconi Space (MMS)[5] and a single-panel version studied by Dornier as a backup. Further cost and surface accuracy estimates led to a further reduction in the size of the primary mirror to 3 m in 1992–1993 (Rider Study by Tiger Team, see Chap. 3). SAG chair Reinhard Genzel (priv. comm. 2013) mentioned that a 2-m telescope was proposed and immediately rejected as becoming a threat for the scientific requirements. The key specifications were then a 3-m diameter Cassegrain monolithic telescope with a total wave front error lower than 11 μm rms, resulting in an effective collecting area of at least 5 m^2, and spatial resolution ranging from about 7 arcsec (FWHM[6]) at the shortest (85 μm) to about 70 arcsec at the longest (900 μm) wavelengths.[7] Three telescope options were identified:

- The multi-panel concept consisted of a set of mirror segments supported by a rigid backing structure
- The single-panel concept based on a monolithic panel design in CFRP, with no separate backing structure
- The thin (2 mm) shell concept made in CFRP supported by dense mesh of fixations organised in a triangular pattern

The performance of the multi-mirror concept was expected to be the best, especially with active control of the panels, but cost was high. The Dornier single-panel concept came second in terms of performance and appeared as the simplest and cheapest one. It was then selected as the baseline for the FIRST telescope under the overall coordination of Matra. A demonstration model of 1.1-m diameter was manufactured by Dornier and submitted to vacuum tests at orbital temperatures at the Centre Spatial de Liège (CSL). The material proposed for the secondary mirror was lightweight glass (Zerodur®[8]) fixed to the primary mirror by a CFRP hexapod (Fig. 6.5).

In mid-1995 a contract was placed with Matra for the development and qualification of the FIRST telescope with Dornier as subcontractor for the main reflector to be made of gold-coated CFRP. The first phase of the study encountered a number of development difficulties and was delayed by six months. Subsequent reassessment by Matra and Dornier of the development work to achieve a qualified telescope projected a schedule delay of more than a year and an increase of the cost at completion by a factor of two. Finally in November 1996, it was decided to abandon the CFRP technology development as it was concluded that the

[5]Matra Espace and Marconi Space System merged in 1990 to become MMS.

[6]The technical term full-width half-maximum, or FWHM, is the angular size in an image of a point source on the sky measured at half maximum of its flux intensity. It is approximately equal to the observing wavelength divided by the telescope aperture.

[7]The 'Red Book' FIRST, Sci (93)6, ESA, 1993.

[8]Zerodur® is a low expansion glass ceramic from Schott company. It is an inorganic, non-porous lithium aluminium silicon oxide glass ceramic. It can be provided as small finished components, as ultra-lightweight mirror substrates, but also as mirror substrates for large segmented and monolithic astronomical telescopes.

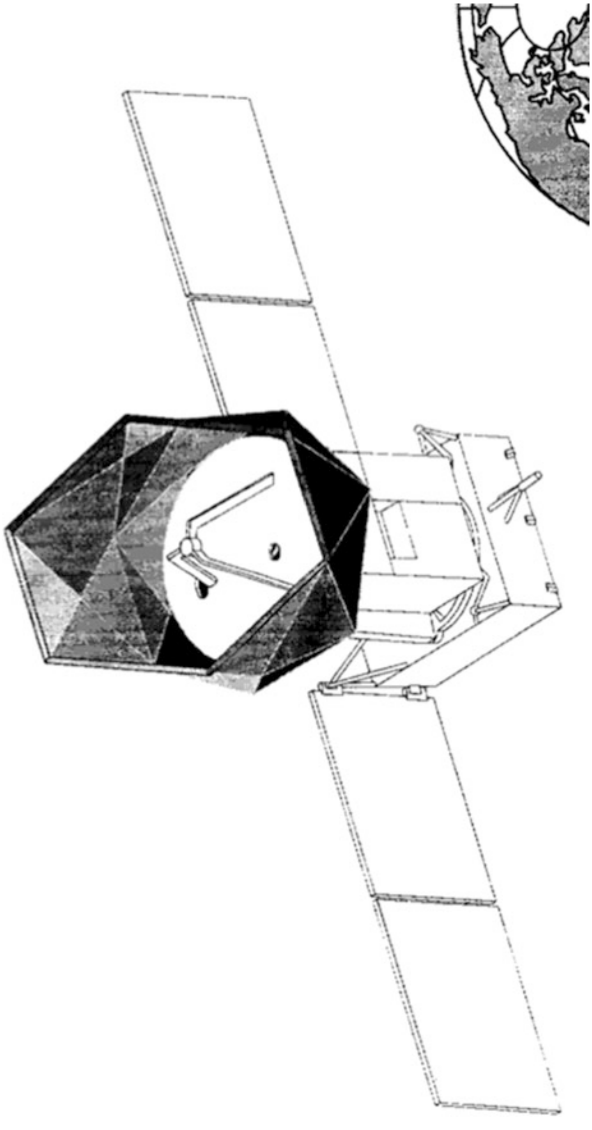


Fig. 6.5 The FIRST satellite design in 1993. Source: ESA 1993. © ESA

specification could not be met and that there were still some fundamental technology concerns. However, for risk reduction purposes, alternative technologies were studied as backups, and ESA had implemented a technology development programme towards proving these critical technologies. The programme included the development and test of large infrared telescopes. Within the same FIRST telescope development contract, both Matra and Dornier had been given the mandate to study alternative materials as possible backups. Each of them studied separately silicon carbide-type materials (SiC and C/SiC, respectively) for the telescope. Matra (renamed Astrium in 1998[9]) with the Céramiques & Composites company (which later became Boostec, in 1999) produced a large SiC piece similar to the one needed for the FIRST reflector.

In parallel to this process, ESA had placed a small technology contract with Aerospatiale and Neyrpic Framatome Mécanique (NFM) to investigate the feasibility of using a whole aluminium telescope for FIRST. The study showed that thermal distortion could be kept sufficiently small by active control of temperature gradients. It also showed that the aluminium surface could be machined with sufficient accuracy. So this presented an option with no major development risk, a simple concept and limited cost, as claimed by Riti and Singer in a draft paper from 1997.[10] The contract was subsequently extended to build a full-scale 3-m diameter flat surface demonstrator mirror to perform machining trials. This work was to be completed by early 1997. The technical and development proposals with cost estimates were submitted to ESA for both the silicon carbide types and for the aluminium telescope. A comparative evaluation concluded that the aluminium telescope with active thermal control was the preferred solution from all points of view: technical, development risk, schedule and cost. The aluminium solution was therefore chosen as the baseline but the SiC technology studied by Matra was kept as a backup. The C/SiC technology option was therefore abandoned in this process. Note that the aluminium option was inspired by the success of the ISO mission, launched in 1995, which had an aluminium telescope structure and a fused silica mirror (Singer 1995). Many members of the ISO project managing team were then reallocated to the FIRST project.

The rescheduling of FIRST in the ESA Horizon 2000 plan led to an advance of the main target dates by approximately one year, with a launch in late 2005/early 2006 rather than late 2006 as previously planned (see Chap. 3). Given that all FIRST's critical technology developments, especially the telescope, had to be started before the start of the development of the satellite, it was important that the study of the silicon carbide backup technology be initiated immediately so that a 'choice'[11] between SiC and aluminium could be made by spring of 1998. As a result, the work on the silicon carbide backup by Matra/Astrium was carried out in parallel to the work on the baseline aluminium material. Following the aluminium machining tests by NFM on two reduced scale (1.3-m diameter) mock-ups and a full-scale mock-up

[9]Matra merged with DASA (Deutsche Aerospace until Daimler Chrysler Aerospace AG), DASA following MBB, in 1998 to become Astrium, which in turn became EADS in 2000.

[10]This paper was never published and only a draft was available.

[11]Note that a clear choice was not made, but aluminium was tried as a preferred option than SiC.

(3-m diameter), ESA considered that the machining demonstration had failed, despite Aerospatiale's view that incorporating the experience gained in the development plan would allow the requirement to be reached. Another important factor in the decision was that the aluminium telescope was considered to be too heavy.

6.4 JPL/NASA Concept

The aluminium telescope concept having been discarded, the FIRST telescope suddenly went to NASA, while the Europeans concentrated on the payload and the service module. Initially, US research laboratories had studied their own concepts for far-infrared astronomy and large space telescopes. The Submillimeter Explorer, proposed in 1986 by a team led by Tom Phillips of Caltech, received some support from France, especially from Ecole Normale Supérieure with many coinvestigators.[12] Another ambitious concept was known as the large deployable reflector (LDR) for wavelengths between 30 µm and 1 mm (Swanson and Kiya 1983), which adopted several shapes over the years. Its primary mirror, for the LDR design with a diameter of 20 m, was composed of 37 segments manually deployable by astronauts once in orbit and thus subject to review every two years through the Space Shuttle as for the *Hubble* Space Telescope (see Fig. 6.6). An R&D project, called precision segment reflector (PSR), was then undertaken to develop relevant new technology and make large mirrors based on panels of 0.9 m, the deformation of which at low temperature (170–120 K) in space was predictable and controllable (Lehman and Helou 1990). The areal density goal was around 5 kg/m^2, far below *Hubble* glass mirror (around 200 kg/m^2). Other work on large mirrors was tackled in the mirror lab of Richard F. Caris at the University of Arizona, leading to the famous honeycomb approach to giant telescopes, with an attempt in 1983 by Roger Angel's team to use silicon carbide (SiC) in the mirror structure. Although Europe proposed new technologies, the US tradition and research in building large telescope mirrors remained very strong and trusted.

In May 1997, NASA/JPL formally announced its interest in building the FIRST telescope as a contribution to the mission. This possibility was developed for over a year with primary mirror diameters varying between 3 and 3.8 m. It involved Composite Optics, Inc. (COI) in San Diego, California, which was awarded a contract in June 1998 as reflector developer. A major milestone for the NASA/JPL technology was the completion of a demonstrator. A feasibility study started within COI with the design, manufacturing and testing of a subscale 2-m diameter primary mirror. A prototype was built (Catanzaro et al. 2001), and several materials were tested in Zerodur® and CFRP. The material selected for this demonstrator was CFRP using a different process (high-temperature curing) than in the early ESA development and already applied in a similar smaller-scale technology (as for the NASA Microwave Limb Sounder mirror). By mid-2000, the activities and the tests performed with a 2-m spherical demonstrator had shown some of the capabilities of the material and

[12] Submillimeter Explorer, JPL proposal to NASA, July, 1986.

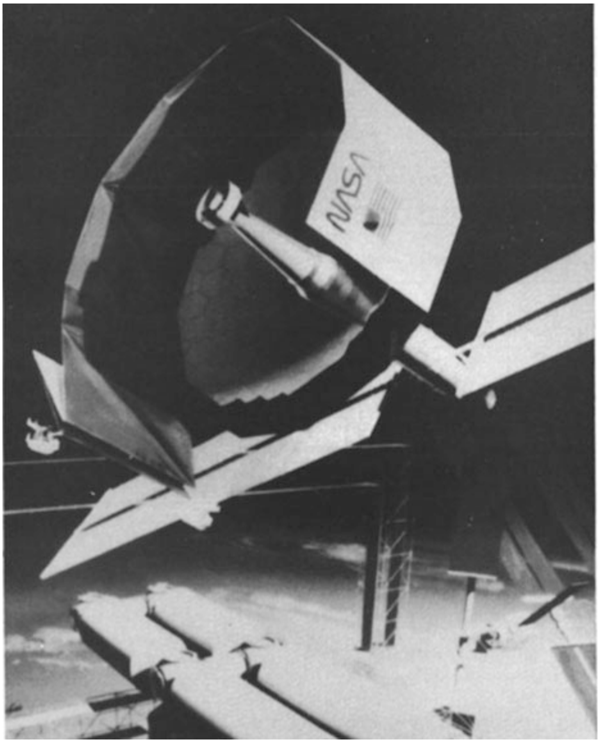


Fig. 6.6 LDR project planned for launch in 2010. © NASA

characterised errors on the surface at a temperature of 70 K. Following the path of the PSR project, JPL had then tried to build a carbon fibre mirror by weaving it on a mould. Demoulding conducted to small deformations due to the hydrophilic nature of the carbon fibre. It was then attempted to distort the mould inversely to offset the distortions produced by the mould. The ability to correct errors using a secondary or tertiary mirror was also proposed (Catanzaro et al. 2001). Improvements were identified at that stage that needed to be implemented with the JPL/COI technology in order to meet the stringent FIRST telescope requirements. However, it was recognised that they could not be achieved. Furthermore, by the end of 2000, and due to budgetary reasons, NASA was forced to stop all activities on the work while waiting for budget approval for the flight model, due to be announced in 2001. Finally, in May 2001, NASA informed ESA officially that they were unable to deliver the telescope due to tight budget constraints, technology development problems and cost overruns in NASA's contributions to the *Herschel/Planck* scientific instruments.

What went on behind the scenes is not fully clear, but in addition to the officially declared reasons, different managerial approaches at ESA and NASA had probably created some lack of confidence and misunderstanding. NASA might have been under the impression that ESA did not really count on them to deliver a telescope for FIRST, or at least that not delivering a telescope would not be a problem for ESA. JPL had strongly advocated building a telescope as large as could be accommodated that was a 3.8-m aperture. ESA was more concerned with the feasibility

Table 6.1 Telescope characteristics

Year	Diameter	Material	Structure	WFE (rms)	Industry
1982	8 m	Graphite	Deployable	>16 μm	Dornier
1983	8 m	CFRP	12 deployable panels	>16 μm	Dornier MMB
1990	4.5 m	CFRP	Single panel	11 μm	Dornier
1993	3 m	CFRP	Single panel	11 μm	Dornier
1997	3 m	Al	Single panel	11 μm	Aerospatiale
1999	3.5/3.8 m	CFRP	Single panel	10 (goal 6) μm	COI
2001	3.5 m	SiC	12 brazed panels	6 μm	Astrium Boostec

The wave front error (WFE) budget is driven by the primary mirror. WFE in 1982 and 1983 are deduced from a 8-μm rms surface error

than the size, preferring rather a good 3.5-m telescope than no 3.8-m telescope. The work done at COI was highly regarded and appreciated, but there was always a doubt in the background as to whether CFRP was a viable technology at all. When the NASA telescope work ran into both technical and cost problems, at the same time the SiC back-up technology was looking more and more realistic. The pendulum began to swing towards the SiC option, paving the way for the formal reason for dropping the US option: that there was not enough money for the model philosophy, including a flight spare, which ESA insisted on.

At that stage, many telescope options had been proposed and are summarised in Table 6.1.

6.5 Plan B: The SiC Telescope

When ESA had chosen NASA/JPL and COI to build the *Herschel* telescope, a senior staff member in Matra is reputed to have said 'one day, you will come back to us'. Indeed this finally happened after the abandonment of the NASA concept. ESA decided to continue funding the SiC backup technology at a low level at the time of choosing the aluminium option. And Matra had certainly pursued its research and development on lightweight SiC space mirrors with the French company Céramiques & Composites (as well as SiCSPACE, a merged entity) based in Tarbes near Toulouse, an industrial ceramics specialist for space and military industries. In 1997–1998, a 1.35-m spherical mirror prototype was assembled with nine SiC segments—the available ovens were not large enough to accommodate a complete monolithic mirror, and this demonstrator qualified the brazing technology to assemble together these segments (Safa et al. 1997). Tests at 150 K were conducted in a cryogenic tank in a vacuum chamber in Toulouse. To switch from 1.35 to 3.5 m, a bigger brazing oven was needed. The larger mirror would comprise twelve SiC petals brazed together. A first segment was produced between 1999 and 2001. During that time, Matra became Astrium Toulouse and Céramiques & Composites became Boostec.

Made of silica crystals, the sintered SiC ceramic was prepared at high temperature and high pressure, providing a very uniform internal structure and avoiding heterogeneities in pressure within the material. As a consequence, the ceramic was extremely hard and had to be machined with diamond tool heads. Between 2001 and 2003, at least twenty-six panels were produced in total, offering some petal spares. The twelve segments were assembled and brazed to form a single mirror. Mechanical assembly and gluing with epoxy had also been considered, but not implemented. Astrium preferred brazing and used the BraSiC® patent of CEA DTEN (now CEA Liten). The French Alternative Energies and Atomic Energy Commission (CEA) had conducted research in the field of metallurgy since the 1990s, with a focus on implementing specific component fabrication and assembly processes. This research was not targeted to a specific industry. Brazing is used to bond metal parts by melting a filler material into the joint. Unlike welding, brazing is carried out at a temperature lower than the melting point of the parts to be assembled. Only the filler metal melts while the surrounding material remains solid. The technique is suitable for joining and sealing a broad range of components. Like diffusion welding, brazing forms a very 'tight' joint. Work at CEA had focused on high-temperature brazing of ceramics for extreme industrial environments like those common in the chemical and aerospace industries, and a proprietary process and the associated filler materials (silicon-based brazing alloys), called BraSiC®, had been developed. This process was used to assemble silicon carbide and other similar ceramics and to make telescope mirrors for space applications. It involved adding a material between two mirror segments with the thermal expansion coefficient matched to that of SiC (approximately 0.1 ppm/K). The joints were very thin, about 300 μm. Tests performed at 77 K with liquid nitrogen showed that the rigidity of the joints was not affected by cooling to that temperature and was equal to that of the SiC material. Air bubbles that might still appear inside the mirror were removed by sonication (Fig. 6.7).

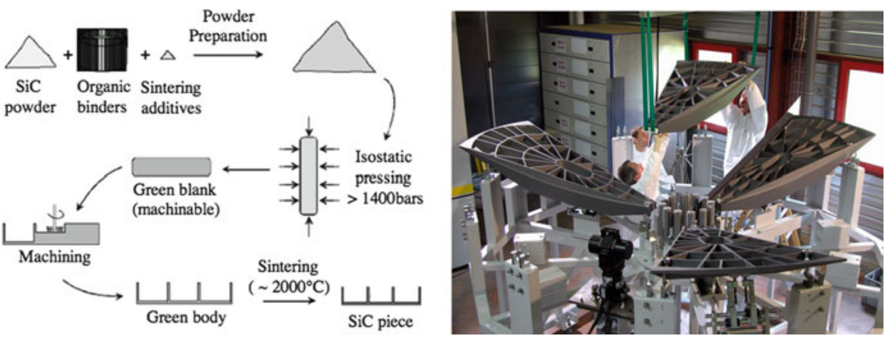


Fig. 6.7 *Left*: major manufacturing for sintered SiC pieces (reproduction from Safa et al. 1997). Silicon carbide fine powder is mixed with organic binders resulting into a 'green material'. The powder is isostatistically pressed at a high pressure (>1400 bars). The resulting green body is machined to the desired shape. For reflectors, the rear face lightweighting is performed on the green body. The machined green body is pressureless sintered at high temperature, about 2000 °C. *Right*: SiC panel assemblage before brazing. © BOOSTEC, ESA/*Herschel*

Unfortunately, the first brazing of the primary mirror in 2003 was a failure. Several vacuum cavities formed on the first frame of the mirror, and six new petals were brazed to replace those damaged by the cut. The new frame was then machined to lighten the inner layer of the mirror by reducing it down to a 2-mm thick skin. The hardness of the ceramic mirror prevented the diamond machining from going below 3-mm thickness, resulting into a mirror weight of 260 kg as specified by ESA. For comparison, the primary mirror of the *Hubble* Space Telescope made of an epoxy structure and covered of fused silica is 2.4 m in diameter and weighs 830 kg (hence an areal density of about 180 kg/m^2). A similar glass mirror for *Herschel* would have weighed 1.7 tonnes. The company Opteon, from the observatory of the University of Turku in nearby Tuorla, Finland, carried out the polishing of the SiC surface to a final roughness of 30 nm. The mirror accuracy was such that the completed telescope would have a total wave front error of less than 6 μm rms. This was a lengthy operation lasting several months. Once polished, the mirror had to be aluminized to provide its necessary reflectivity. This work was carried out using the coating facilities at Calar Alto Observatory in Spain. A thin silicon oxide layer of was then added to protect the mirror to enable cleaning. The secondary mirror and its supporting hexapod were also made of SiC, to ensure homogeneous thermal change effects at telescope level. A second replacement mirror (a blank mirror never polished) was produced in 2005 in case of damage. The full flight telescope was delivered to ESA on 1 February 2008.

6.6 Invention of a Technical Lineage: From *Herschel* to *Euclid*

In Chap. 5, the technical lineage concept and the progress of its concretisation, the C-K design theory and the coopetition management were introduced to characterise innovation and its typology. How do they apply here? Can innovation be represented and measured (quantified and represented)?

First, the SiC telescope defines a novel concretised stage in the development of space telescopes. This radical innovation or invention of a technical lineage has managed to become the European prime technology for space telescopes and has continued its development in the ESA Gaia mission or in the ESA *Euclid* mission. SiC was also used to construct the mirrors and optical bench of the NIRSpec instrument on the *James Webb* Space Telescope. The telescope within the OSIRIS NAC instrument on-board Rosetta, which was designed to produce high-resolution images of the comet nucleus, is also 100% in SiC. From an invention to a radical innovation, the SiC mirror technology has established itself in the face of competition from other technologies in Europe and in the USA. The Japanese Akari mission also made use of SiC mirrors, based on a different technology with respect to the Boostec technology used in Europe (Takashi and Salama 2009). Thanks to SiC the Airbus Company is leading the development of Earth observation satellites.

Secondly, the technical and social process leading to Boostec SiC technology was not straightforward. Figure 6.8 represents the history of the *Herschel* telescope

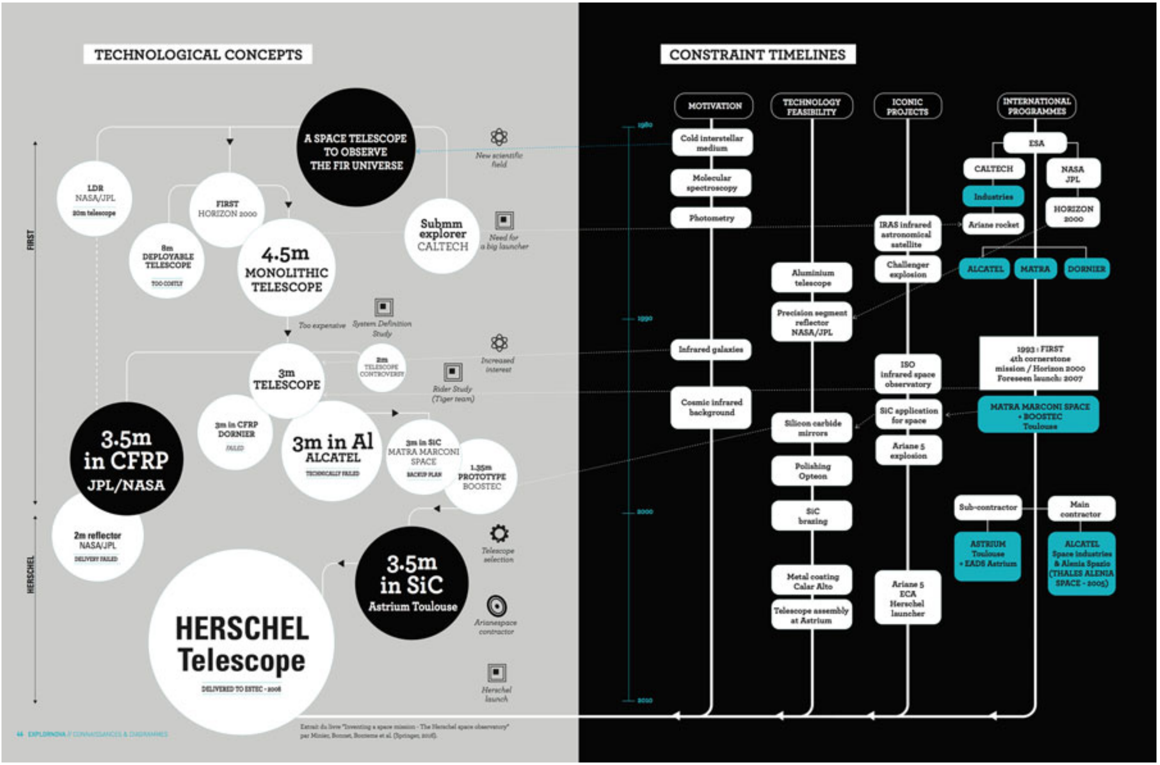


Fig. 6.8 This diagram represents the evolution of the different telescope concepts between the undecidable concept in 1982 and the *Herschel* telescope in 2009 at launch. The *left panel* includes all the concepts, while the *right panel* introduces the knowledge that constrains or triggers the concept space (see C-K design theory introduced in Chap. 5). In addition, knowledge is classified under four timelines: the scientific and technical motivation, the technology feasibility, the emblematic and comparable space projects impacting on the project in study and the international research programme providing the overall guidelines, rules and budget. In between, *blue arrows* connect together knowledge and concepts, as well as illustrating the reconfiguration of the knowledge space. For instance, the discovery of the cold interstellar medium (knowledge) motivates the need for a large FIR space telescope (concept). In a similar way, the CIB observation with ISO justified new science goals achievable with smaller telescopes than initially conceived. The research programme between industries MMS and Boostec gave birth to a novel SiC technology. © CEA (V. Minier, S. Leroy)

from the initial concept to the concretised SiC mirror, taking into account the most significant constraints that have influenced this process. The diagram is inspired by the C-K diagram (Chap. 5). The left part (technological concepts) shows the telescope concepts starting with the undecidable one. The right part (constraint timelines) illustrates the factors and events that constrained the telescope design. Timelines have been added to classify the main constraints or knowledge into four categories (see Sect. 6.1): scientific motivation, technology feasibility, emblematic missions and programmatic framework. 'Design projects aim to transform undecidable propositions into true propositions in K. Concepts define unusual sets of objects called C-sets, i.e. sets of partially unknown objects whose existence is not guaranteed in K. During the design process C and K are expanded jointly through the action of design operators' (Hatchuel and Weil 2009). The numbers of technical concepts and industries in competition were impressive. All the dominant technologies and previously well-established businesses in space mirror design were discarded in favour of a novel concept based on an advanced material technology. With this type of diagram, it is possible to visualise the innovative design and realise that it is not a linear process. It also includes crises as a normal part of the project life, often stimulating ingenuity and problem-solving.

Finally, the process mapped in Fig. 6.8 represents the evolution of coopetition, bearing in mind that cooperation and competition have played a major role in designing the most advanced telescopes. As described in Chap. 5, coopetition has three characteristics (complex relations, separation between the phases of cooperation and competition through a sharing of tasks and responsibilities and/or with an alternation of the phases, regulatory body). The complexity of relations between actors, which are potential collaborators and competitors, is emphasised by the frequent merging of space industries throughout the design process. The relative separation between the phases of cooperation and competition are illustrated by Thales Alenia Space (formerly Aerospatiale) becoming the main contractor with Astrium Toulouse (formerly MMS) and Astrium Friedrichshafen (formerly Dornier) as subcontractors for the telescope and cryostat, respectively. The regulatory body was the European Space Agency.

Innovation being characterised and mapped, is there any way of measuring it, i.e. quantifying it and representing it on a graph? As seen in Chap. 5, the benefits of an innovation are relative because there is no single criterion to measure the performance of a technical object, and the choice of any criterion depends on the functionality considered. The design of the *Herschel* telescope deals with a major passive functionality (see Chap. 5), because the choice of the material and the shape of the telescope structure were crucial to reduce the weight and give the correct resistance and stability to the mirror during its travel and its life within its final associated environment. It also deals with a median informational functionality, which is to transmit the expected information at a given resolution and expected sensitivity. The measure of innovation then results from the necessity to achieve specific scientific capabilities without losing certain characteristics of the ideal telescope design (cf. Altshuller 2007). It is a compromise between resolution and sensitivity (the informational functionalities) as scientific parameters on one hand

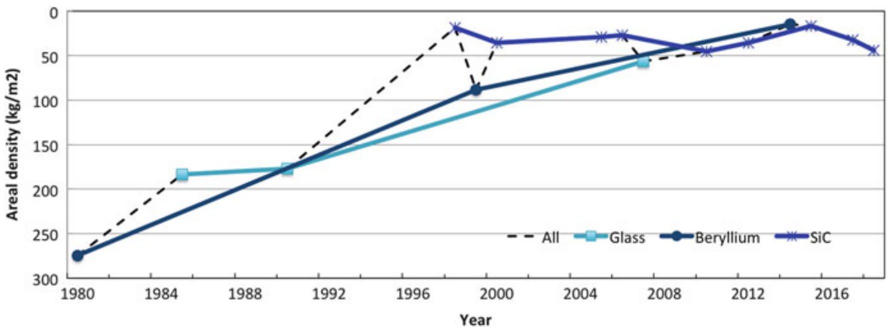


Fig. 6.9 Areal density evolution for different types of material used in space mirror industries: glass (ULE fused silica), beryllium and SiC. Note that the areal density of the mirrors depends not on the material choice but also, and mainly, on the support structure of the mirror and its machining. The areal density value reflects these two properties. © CEA (V. Minier)

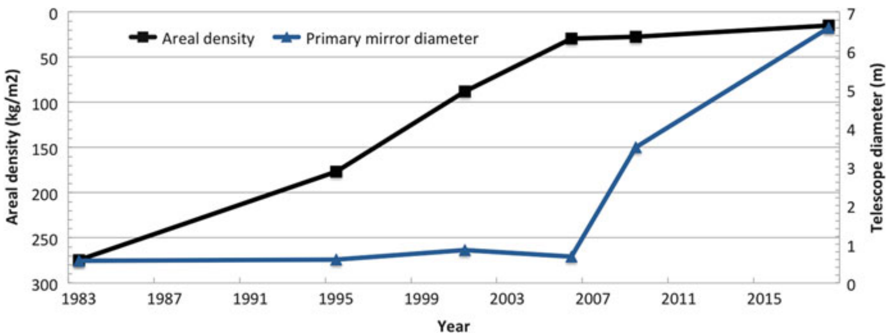


Fig. 6.10 Areal density evolution for infrared space telescopes since IRAS and evolution of their diameter. We note a strong decrease (improvement) in the areal density between 1983 and 2006, while telescope diameter remains under 1 m (limited by the need to cool space telescopes with liquid cryogens), and then a strong increase of the telescope diameter by a factor 10 while the areal density remains roughly constant. © CEA (V. Minier)

and structure and areal density (the passive functionality) as engineering parameters on the other hand, with the associated medium playing a key role. Radical innovation emerges, unsurprisingly, from a challenging trade-off between diameter and weight (or between science achievement and industrial cost).

The areal density—the ratio of the primary mirror mass (M) divided by its surface ($\pi D^2/4$, simplified to D^2)—would be an interesting figure of merit to take into account both scientific and industrial goals. Figures 6.9 and 6.10 show the evolution of the areal density from IRAS to *Euclid* and a comparison between this evolution and the size of infrared space telescopes. With a primary mirror of 3.5-m diameter and an areal density of 30 kg/m^2, *Herschel* was larger and less dense than the *Hubble* Space Telescope fused silica telescope (2.4 m and 180 kg/m^2). The *James Webb* Space Telescope (JWST) will go even further with a 6.5 m primary mirror consisting of deployable panels in beryllium. Boostec, now part of Mersen, a company with expertise in advanced materials, continues to develop SiC mirrors with technology programmes such as OTOS reaching the areal density of 16.5 kg/m^2 (Bougoin et al. 2014). The desired capability is around 1 kg/m^2 by around 2030 (Matson et al. 2008).

As seen on Fig. 6.9, the SiC technology performance is relatively constant over the last 20 years, while beryllium-based technologies have maintained a strong improving trend. The glass mirror technology remains below the levels reached by SiC and beryllium mirrors. A possible interpretation is that SiC technology emerges after a long period of development, beryllium technology has yet to reach its maturation point and glass technology is starting to saturate. Interestingly, as shown in Fig. 6.10, there is a delay between areal density increase and primary mirror diameter increase for the case of infrared space telescopes. A possible explanation is that 20 years are needed for a technology to mature sufficiently to be implemented on large telescopes such as *Herschel* and JWST. Industries then gain on their investment and do not generate major breakthroughs.

6.7 Conclusion and Future Perspective

In summary, the SiC telescope for *Herschel* is a true invention leading to radical innovation in space industries in the sense of Schumpeter (Chap. 5), i.e. a new combination of production means, which increases the value of production and, as a result, profit. This invention and the possibility of building large telescopes also led to science achievements thanks to major improvements in angular resolution and sensitivity. The industrial and scientific interests merge together in a common figure of merit, the areal density, the time evolution of which provides a historical perspective.

The future of FIR and submillimetre astronomy needs further major innovation to go to the next level: 10, 20 or 30-m diameter equivalent apertures in space. Indeed, the science cases (e.g. galaxy evolution, protoplanetary systems) will require 0.1–1″ angular resolution at 100 μm and hence telescope primary mirror having a diameter greater than 10 m. However, fairings put a strict limit to the size of a single dish aperture that can be launched in space. The JWST, with its deployable 6.5-m diameter main telescope mirror, just fits in the Ariane 5 fairing. The next step is either to decrease significantly the areal density of deployable telescopes or to combine a set of smaller telescopes in an interferometer network. As an example, TALC (Thinned Aperture Light Collector) is a 20-m diameter deployable concept that explores some unconventional optical solutions, between the single dish and the interferometer, to achieve a very large aperture. Its collecting area is 20 times larger than *Herschel*'s, giving access to very faint and distant sources (Sauvage et al. 2014). The deployable mirror structure exploits the concept of tensegrity, i.e. structural rigidity achieved through compression.

The TALC mirror is a segmented ring of 20-m diameter and 3-m width. For launch, the identical mirror segments are stored on top of each other and a deployable mast pulls a series of cables that deploy the stack of mirror into the required shape. Innovations such as this will be needed to provide the ambitious observational capabilities that far-infrared space astronomy will need in the future.

References

Altshuller G (2007) The innovation algorithm. Technical Innovation Center, Worcester

Bougoin M, Lavenac J, Coatantiec C, Costes V (2014) A new technological step for SiC mirrors preparing OTOS. In: International conference on space optics

Catanzaro B, Connell S, Mimovich M, Backovsky S et al (2001) Cryogenic (70 K) measurement of an all-composite 2-meter diameter mirror, Optomechanical design and engineering. In: Hatheway AE (ed) Proceedings of SPIE, vol 4444, pp 238–255

ESA (1982) A Far-Infrared and Submillimetre Space Telescope (FIRST), a mission proposal to the European Space Agency

ESA (1983) FIRST assessment study ("The Yellow Book"), SCI (83) 1

ESA (1993) Red book FIRST, SCI (93)6

Fazio GG (1983) Planned NASA space infrared astronomy experiments. Adv Space Res 2(4):97

Hatchuel A, Weil B (2009) C-K design theory: an advanced formulation. Res Eng Des 19:181–192

Lehman DH, Helou G (1990) Telescope technology for space-borne submillimeter astronomy, ESA, from ground-based to space-borne sub-mm astronomy, p 381

Matson LE, Chen MY, deBlonk B, Palusinski I (2008) Silicon carbide technologies for lightweighted aerospace mirrors. In: Advanced Maui optical and space surveillance technologies conference

Safa F, Levallois F, Bougoin M, Castel D (1997) Silicon carbide technology for submillimetre space based telescopes. In: 48th international astronautical congress

Sauvage M, Chanial P, Durand GA, Rodriguez LR et al (2014) The science case and data processing strategy for the Thinned Aperture Light Collector (TALC): a project for a 20m far-infrared space telescope. Proc SPIE 9143:14, id 91431B

Singer C (1995) The infrared space observatory. Space Sci Rev 74(1–2):67–72

Submillimeter Explorer (1986) JPL proposal to NASA, July 1986

Swanson PN, Kiya M (1983) LDR: an orbiting submillimeter-infrared telescope for the 1990s. SPIE Proc 0365:13

Takashi O, Salama (2009) AKARI: space infrared cooled telescope. Exp Astron 27:9

Toulemont Y, Passvogel T, Pilbratt G, de Chambure D, Pierot D, Castel D et al (2004) The 3.5 m all SiC telescope for Herschel. Proceedings of the 5th International Conference on Space Optics (ICSO 2004). In: Warmbein B (ed) ESA SP-554. ESA Publications Division, Noordwijk, pp 341–348, ISBN 92-9092-865-4

Chapter 7
Far-Infrared Bolometers: Technical Lineages

Abstract Bolometers are used as detectors in the *Herschel* SPIRE and PACS instruments. This chapter focuses on the history of bolometer development leading to the SPIRE and PACS detectors, operating in the submillimetre and far-infrared wavelength ranges. This historical and technological review describes the main technical lineages of bolometers and how the SPIRE and PACS instruments on *Herschel*, respectively, employed both an adapted and improved variant of a well-established technology and a novel bolometer architecture developed especially for the mission.

7.1 Bolometric Detectors

A bolometric detector measures electromagnetic power through a change in temperature produced when power is absorbed. Bolometers are the most suitable detectors for astronomy at wavelengths above ~200 μm (the long-wavelength limit for stressed Ge:Ga photoconductors) and below ~2 mm (where heterodyne radiometers start to become competitive; see Chap. 8). Depending on the details of the application, bolometric detectors are competitive over the 50 μm–3 mm range.

The principle dates back to 1800, when William Herschel used a prism to produce a spectrum of the Sun in order to investigate the heating effects of different colours of visible light. He placed the bulb of a mercury thermometer in the different colours of the spectrum and observed that the heating effect increased as he moved the thermometer from the blue to the red end of the spectrum. However, he carried on and found an even higher temperature increase when it was positioned beyond the red end. In so doing, he discovered infrared radiation.

The detailed theory of modern bolometric detectors was developed by Jones (1953) and further elaborated by others, especially Mather (1982). There is an extensive literature on the theory and practice of bolometric detectors and systems, which has been comprehensively reviewed by Richards (1994). The essential features are shown schematically in Fig. 7.1. Incident electromagnetic power heats up the absorber. Heat is allowed to flow from the absorber to a heat sink at a fixed temperature T_o via a thermal link. A thermometer, the resistance of which is strongly temperature-dependent, is attached to the absorber. A current is passed through the thermometer, via a bias voltage source and load resistor, to measure the

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_7

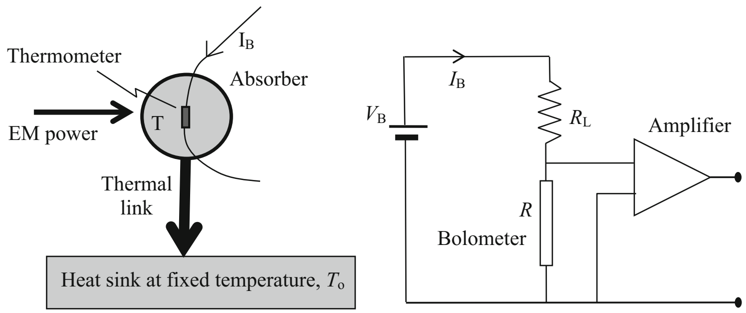


Fig. 7.1 *Left*, principle of operation of a bolometer; *right*, method of reading out the bolometer signal. © Cardiff University (Matt Griffin)

resistance, dissipating some electrical power, which heats the bolometer to a temperature, T_1, slightly higher than T_o. If no radiant power is incident, the absorber will be at this temperature. If some radiant power is absorbed, it will rise to a slightly higher temperature $T > T_1$. The value of T depends on the amount of power absorbed and the thermal conductance between the absorber and the heat sink. A readout amplifier measures the voltage across the bolometer, which depends on its temperature and thus on the incident radiant power.

The most important performance parameters, as with many detectors, are the noise-equivalent power (*NEP*), usually quoted in W $\mathrm{Hz}^{-1/2}$, and the time constant (τ). The *NEP* is a measure of sensitivity and characterises the smallest level of incident power than can be detected, so that a low value corresponds to a good detector. The time constant corresponds to the typical time needed for the detector to respond to a change in the amount of incident radiation. A low value of τ is also good, as it means that the detector has a fast response. Astronomers are interested in the overall sensitivity of an instrument rather than that of the detector alone and express this in terms of the noise equivalent flux density, or *NEFD*, which is the level of flux density required to be equivalent to the noise present in the system and depends on many additional characteristics of the instrument and the telescope (as well as the atmosphere for ground-based observatories).[1] Of course, one of the most important parameters affecting the *NEFD* is the detector *NEP*—the better the *NEP*, the better the *NEFD*.

Bolometer sensitivity depends strongly on the operating temperature of the device, with, in principle, the best performance achieved at the lowest possible temperature. There are two fundamental reasons for this. First, the relative temperature rise generated by a given incident power depends on the heat capacity of the bolometer—the lower the heat capacity the higher the temperature rises and so the

[1] *NEFD* is the overall *NEP* divided by the aperture area of the telescope, the bandwidth of the instrument, the atmospheric transmission if on earth, and various parameters characterising the optics transmission and coupling to incident radiation. In astronomy, it is often expressed in terms of Jansky (Jy) with 1 Jy = 10^{-26} W m^{-2} Hz^{-1}.

larger the signal. The heat capacity of materials decreases strongly with decreasing temperature, so that a bigger signal is produced at lower temperature. Second, measurement of the signal from the bolometer must be made in the presence of electrical and thermal noise from the bolometer itself, both of which decrease with operating temperature.

The need for extremely low-temperature operation and the difficulty and complications of achieving such low temperatures, especially in working astronomical instruments at observatories, meant that it was not until the 1960s that sensitive astronomical instruments using bolometers were developed.

7.2 Early Bolometer Instruments for Astronomy

A major advance was made in 1961 by Frank Low of the National Radio Astronomy Observatory at Greenbank, who developed a bolometer using a single crystal of gallium-doped germanium (acting as both the absorber and the thermometer), cooled to a temperature of around 2 K using a compact liquid helium cryostat and capable of being attached to an astronomical telescope (Low 1961). It had an *NEP* of 5×10^{-13} W Hz$^{-1/2}$ and a time constant of 0.4 ms, far superior to anything achieved previously.

This development was the starting point for the technical lineage leading to the SPIRE and PACS detectors. The essential features of a liquid-helium-based cryogenic system, the use of a doped semiconductor as the thermometer, and operation in the far-infrared part of the spectrum were all in place in Low's setup and were systematically and steadily enhanced in the following decades. Using his systems, Low and colleagues made many pioneering infrared to millimetre wavelength observations during the 1960s, with ground-based and airborne telescopes, including observations of the Sun, the Moon and the planets and of bright stars, galactic clouds and external galaxies.

Low's original bolometer used the germanium crystal to carry out both of the key functions of the bolometer—absorption of the incoming radiation and measurement of the corresponding temperature rise. But these two tasks actually placed different requirements on the device, so that it was impossible to optimise them both simultaneously. A significant step forward was made by Coron et al. (1972) and developed further by Clarke et al. (1974), Nishioka et al. (1978) and others, with the development of the composite bolometer, which involved separating the two functions. A carefully designed absorbing substrate was used to absorb radiation as efficiently as possible, and a separately optimised thermometer element was attached to it in order to measure the temperature. The best implementation of the absorber was in the form of a thin dielectric disk coated with a metal film of the correct impedance. This allowed incoming radiation to be absorbed with high efficiency while keeping the heat capacity as low as possible. The thermometer, optimised purely for thermometric performance, could then be simply attached to the absorber. The composite design subsequently became the standard architecture

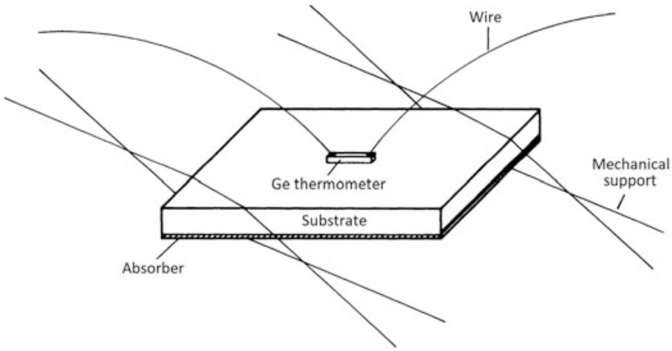


Fig. 7.2 An early composite bolometer. © Kreysa and Haller (1982)

for bolometers in astronomy. The essential features of a composite bolometer are shown in Fig. 7.2 (from Kreysa and Haller 1982).

7.3 ^{3}He-Cooled Bolometers

Further advances became possible with the advent of portable cryogenic systems capable of attaining lower temperatures in an operational astronomical instrument. The temperature of ^{4}He, the most abundant isotope of helium, is 4.2 K at atmospheric pressure and can be lowered by pulling a vacuum on the liquid surface, which leads to the preferential evaporation of higher-energy helium atoms. The lowest temperature that can be achieved in this way is around 1 K. ^{3}He does not suffer from this limitation, and so a pumped ^{3}He vessel can reach lower temperatures—down to around 0.25 K (see Chap. 9 for more details). The precise temperature dependence of the bolometer *NEP* depends on the details of the design and the materials used, but it typically goes as T^2. So a decrease in operating temperature from around 1.5 K to 0.3 K can enable about a 25-fold improvement in sensitivity. In the late 1970s, Ira Nolt and Jim Radostitz of the University of Oregon developed the first compact refrigerator based on ^{3}He rather than ^{4}He as the coolant for the final stage (Radostitz et al. 1978). The system was first used for astronomy in the submillimetre range in 1980 on the 3.8 m United Kingdom InfraRed Telescope (UKIRT), on Mauna Kea, Hawaii, using a composite germanium bolometer developed by Peter Ade and colleagues at Queen Mary College London and with an *NEP* of around 10^{-15} W Hz$^{-1/2}$ (Ade et al. 1984). Similar photometer systems were developed and used by other groups in the same period (e.g. Kreysa and Haller 1982; Roellig and Houck 1983). Two other cryogenic techniques, adiabatic demagnetisation and ^{3}He–^{4}He dilution refrigeration, have since been used to cool bolometers for astronomy and can achieve temperatures as low as a few tens of mK.

During the 1980s and 1990s, ground-based submillimetre observatories, such as the James Clerk Maxwell Telescope (JCMT), the Caltech Submillimeter

Observatory (CSO) and the Kuiper Airborne Observatory, operated routinely with state-of-the-art bolometer instruments. One design challenge with these instruments was coupling the radiation collected by the telescope efficiently to the detector. The optimum size for the bolometer is comparable to the wavelength being detected, or a little larger. Making it smaller compromises absorption efficiency and any bigger introduces unwanted additional heat capacity. So the bolometer was no more than a few mm in size. However, the size of the diffraction pattern (the image of a point source) produced by a telescope is approximately $2F\lambda$ where F is the focal ratio (the ratio of the telescope's focal length to its diameter) and λ is the wavelength. Infrared telescopes performed best with high focal ratios typically around 30, making the diffraction pattern several tens of wavelengths in diameter. The usual method for concentrating the radiation onto the bolometer was to use a condensing optic known as a Winston cone (Winston 1970; Harper et al. 1976). With an off-axis parabolic shape, this has the property that radiation coming into a large entrance aperture from within a clearly defined field of view is directed through a much smaller exit aperture. Radiation coming from outside that field of view is reflected back out again. The entrance aperture could be made comparable to the size of the diffraction spot and the field of view matched to the angle subtended by the telescope. The bolometer was placed in an integrating cavity fed by the exit aperture so that as much as possible of the radiation could be absorbed. To avoid having a very large and massive Winston cone, a lens placed at the telescope focus in front of the Winston cone was sometimes necessary to focus the beam onto a smaller Winston cone entrance aperture.

This scheme is exemplified by the design of UKT14 (Duncan et al. 1990), a submillimetre photometer installed in 1986 as a facility instrument for the UKIRT telescope, and subsequently operated on the nearby 15-m JCMT. The size of the UKIRT diffraction spot at the longest observing wavelength of 1.1 mm was 90 mm. For good performance, it was necessary to capture the radiation efficiently over a diameter of at least 70 mm. The UKT14 optical design is shown in Fig. 7.3. A lens placed at the telescope focus directed the radiation onto the entrance aperture of the Winston cone which in turn concentrated the radiation onto the bolometer. The filter wheel allowed different submillimetre bands to be selected. To minimise the background power on the detector from the instrument itself, the whole system was cooled, with the lens at 77 K, the filters at 4 K and the Winston cone and detector at 0.3 K. The UKT14 bolometer achieved an *NEP* of ~10^{-15} W Hz$^{-1/2}$.

Composite bolometers were also flown on the Cosmic Background Explorer (COBE) satellite, launched by NASA in 1989, which made groundbreaking measurements of the spectrum and anisotropies of the cosmic microwave background (CMB) radiation. COBE's FIRAS and DIRBE instruments both had 1.6-K bolometers using doped silicon thermometers with metalized diamond substrates (Serlemitsos 1988). These devices were developed especially for the mission to achieve low *NEP* ($<10^{-14}$ W Hz$^{-1/2}$) with a design robust enough to be launched and operated in space. FIRAS covered a broad wavelength range (0.1–10 mm) to measure the CMB spectrum, and DIRBE measured diffuse emission from the Universe at 140 and 240 μm.

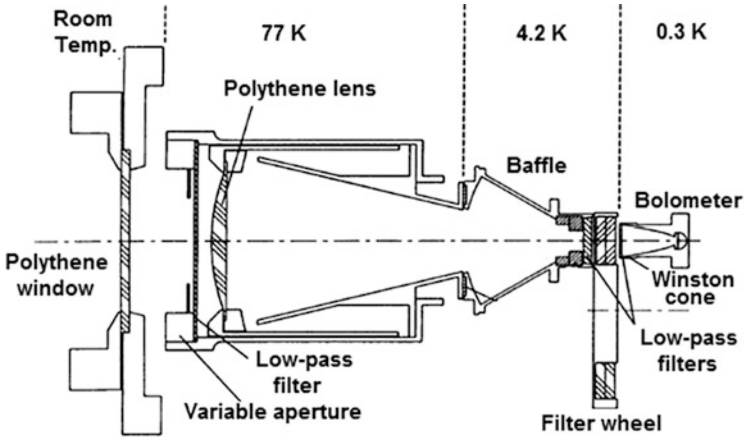


Fig. 7.3 Optical design of the UKT14 submillimetre photometer. © Duncan et al. (1990)

Ideally, the dominant noise contributions in a bolometric detector are so-called photon noise, which is a fundamental and unavoidable property of the radiation incident on the detector, and ‘phonon noise’, which is caused by thermodynamic fluctuations in the bolometer temperature and always present at some level dictated by the operating temperature. In infrared astronomy, most of the radiation falling on the detector comes not from the source being observed but from various background elements such as the telescope, the instrument, or the Earth’s atmosphere. Although the background can easily be subtracted to measure the true source signal, statistical fluctuations in the background radiation lead to photon noise, which then represents the ultimate limit to sensitivity. For this reason, great efforts are made to minimise background radiation, such as cooling the whole instrument as well as the detectors and cooling the telescope itself (only possible in space). In a well-optimised bolometer system, photon and phonon noise will be the main contributions, with photon noise being the main component. The bolometer is then as good as it needs to be and is said to achieve background-limited sensitivity.

Some semiconductors have excess noise over and above these components, degrading the detection sensitivity. The most common techniques for doping semiconductors are melt doping and ion implantation, both of which can have unwanted side effects of causing crystal imperfections and/or impurities which can adversely affect the electrical behaviour of the crystal, including its noise performance at low temperature. In developing semiconductor bolometers, it was important to ensure that the thermometer crystal behaved as closely as possible to the ideal in two key respects: (1) it should exhibit as little excess noise as possible and (2) its resistance should be a function only of its temperature, with no dependence on the electric field across it. Because of its widespread use in the electronics industry, silicon is the best understood semiconductor material, with well-established techniques for achieving desired properties for devices and manufacturing large arrays of devices with uniform properties. A monolithic silicon bolometer was developed by Downey et al. (1984) using integrated circuit fabrication techniques with ion implantation, achieving *NEP* as low as 4×10^{-16} W Hz$^{-1/2}$ at 0.35 K. Although many sensitive devices were made from silicon, germanium tended to be generally superior in terms of low-noise performance and negligible electric field effects. The most well-behaved material turned out to be neutron transmutation-doped (NTD) germanium (Haller et al. 1984). This is made by irradiating an extremely pure germanium crystal with neutrons. Some neutrons are captured by germanium atoms, which subsequently undergo β-decay to form atoms of gallium, which is a P-type dopant in germanium (an impurity atom that tends to acquire an electron from a neighbouring Ge atom, creating an unfilled electron state known as a 'hole', which allows electrical conduction to occur). The doping level is easily controlled as it depends on the total neutron dose and is uniform throughout the material because neutrons are so penetrating. After irradiation, the radioactivity associated with the transmutation process (half-life ~11 days) must be allowed to decline. NTD germanium bolometers provide a very close approximation to ideal bolometers: resistance shows negligible dependence on the applied electric field, and the material exhibits very low excess noise, even down to very low frequencies.

7.4 The Development of Ground-Based Bolometer Array Instruments

The advent of ^{3}He cooled composite bolometers enabled more sensitive observations than had previously been possible. For example, UKT14 had a detector *NEP* of around 10^{-15} W Hz$^{-1/2}$ with an 800-μm *NEFD* of 0.5 Jy Hz$^{-1/2}$. Comparable performance levels were achieved by other instruments at the time. But instruments were still limited by detector noise rather than photon noise and were not sensitive enough to detect distant galaxies. Furthermore, they still only had single detectors. With many astronomical applications requiring observations of extended sources such as star-forming clouds and galaxies, making an image was a time-consuming process with the instrument's single beam having to be raster-scanned back and forth across the source—a poor use of observing time on expensive telescopes. Bolometer cameras—instruments with detector arrays—were desperately needed to speed up such observations.

Bolometer cameras were built for the Institut de Radio Astronomie Millimetrique (IRAM) 30-m telescope, the Caltech Submillimeter Observatory (CSO) and the JCMT. A sequence of cameras (Kreysa et al. 1999), built by MPIfR, Bonn, for operation at 1.2 mm wavelength on IRAM between 1992 and 2001, was known as MAMBO (Max Planck *M*illimetre *Bo*lometer). The JCMT camera (Holland et al. 1999), built by the UKATC and Queen Mary College, London, was known as SCUBA (Submillimetre Common User Bolometer Array). It worked at 450 and 850 μm and became operational in 1996. Bolometer cameras called SHARC (Submillimetre High Angular Resolution Camera; Wang et al. 1996) and SHARC-II (Dowell et al. 2003) were built for the CSO by Caltech and NASA GSFC and operated at 350 or 450 μm. A 1.1 mm and 2 mm camera for the CSO, Bolocam (Glenn et al. 1998, 2003), was led by the University of Colorado. The IRAM and CSO instruments had bolometers cooled to around 0.3 K using ^{3}He refrigerators, but SCUBA used a ^{3}He–^{4}He dilution refrigerator to achieve an operating temperature of around 0.1 K, making the individual detectors much more sensitive.

Besides the need for focal plane arrays and bigger cryostats, building bolometer arrays posed some new challenges compared to single-detector systems. Firstly, the use of Winston cone optics for coupling the detector to the telescope was not easy given the difficulty and expense of manufacturing large numbers of Winston cones. SCUBA, MAMBO and Bolocam all implemented a simpler alternative using straight-walled conical feedhorns. A feedhorn collects less power than the 'multi-moded' Winston cone but is easier to manufacture in large numbers. Such a horn acts as a 'single-mode' antenna, coupling reasonable well (with aperture efficiency[2] around 70%) to the diffraction spot when sized to have an entrance aperture of $2F\lambda$. The circular feedhorn apertures were close-packed in the focal plane, fitting in as many as possible given constraints of mass, volume and cooling power. In the case of the MAMBO-2 camera, there were 117 bolometers operating at 1.2 mm. For SCUBA, there were two arrays observing simultaneously at 450 and 850 μm, with 37 and 91 bolometers, respectively. Figure 7.4 shows a schematic of the two SCUBA arrays and a photograph of the 450-μm array.

Secondly, bolometers had usually been made individually by hand, and more reliable standardised processes were needed for fabricating them. The SCUBA bolometers used bismuth-coated sapphire substrates and were built individually and integrated into individual feedhorns in a successful but very laborious process. For the MAMBO-2 IRAM camera, a hybrid approach was adopted. Substrates were formed by depositing a silicon nitride film on a silicon wafer and etching away the silicon to form an array of thin (<1 μm) free-standing silicon nitride membranes. The individual bolometers with their absorbers were then attached manually to the membranes. The bolometer array was then integrated in one step with an array of single-mode conical horns.

The SHARC cameras used silicon bolometers and also adopted a different approach to optical coupling. In contrast to the feedhorn-coupling method, the bolometers were 'absorber-coupled'. In this configuration, a 'CCD-style' array of bare bolometer pixels is exposed directly to the radiation in the focal plane. For SHARC-II, individual 32-element linear arrays of silicon bolometers were close-packed to form a 12×32 rectangular array, shown in Fig. 7.5. The bolometers were made using a technique developed at NASA GSFC, which allowed the legs for

[2]Aperture efficiency is defined as the fraction of the power from a point source that is coupled to the detector by the optical system.

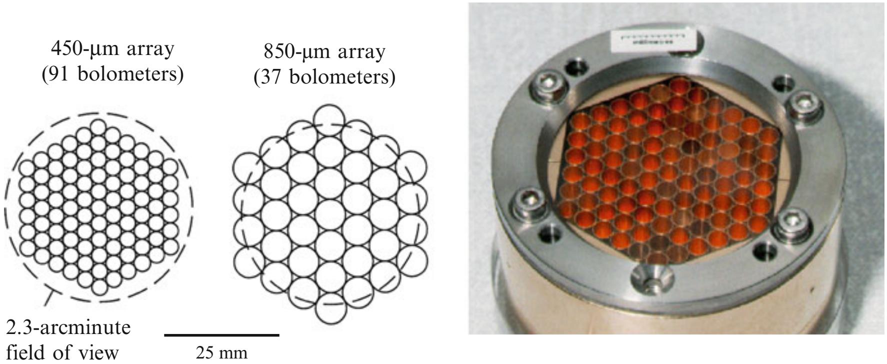


Fig. 7.4 *Left*, schematic diagram of the two SCUBA submillimetre arrays. © Holland et al. (1999); *right*, photograph of the 450 μm array. © James Clerk Maxwell Telescope

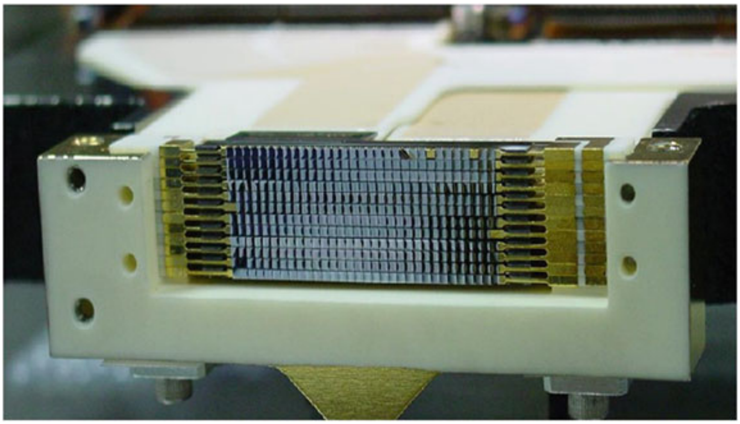


Fig. 7.5 Photograph of the SHARC II focal plane array. © Dowell et al. (2003). The bolometers cover an area of 32 × 12 mm. © Caltech Submillimeter Observatory

mechanical support and electrical connections to be folded back by 90 degrees allowing the lines of pixels to be butted up against each other (Moseley et al. 2000).

The two methods of coupling the bolometers to the telescope beam—feedhorn-coupling and absorber-coupling—have their own advantages and disadvantages, and deciding which approach is best for a given application is by no means an easy or uncontroversial task. This would prove to be a very important issue for the *Herschel*-SPIRE instrument, as described below.

In parallel with the revolution being brought about by array instruments on the ground, other researchers were developing balloon-borne telescopes to measure more accurately the CMB anisotropies that had been detected by COBE. The 2.7-K black body spectrum of the CMB peaks just beyond 1 mm wavelength, where bolometric detectors offer best performance, and accurate characterisation of the angular power spectrum of CMB anisotropies would constitute a powerful probe of the physics of the Big Bang. A collaboration led by Caltech/JPL developed the BOOMERanG balloon-borne instrument to map a large area of sky to search for these anisotropies. It was designed for a long-duration Antarctic balloon flight, in which it would be launched from the McMurdo base on the coast of Antarctica, flying at stratospheric altitudes (close to 40 km) while being carried in a roughly circular itinerary around the South Pole by the circular stratospheric vortex wind, coming back near the launch site after a flight of about 10 days. A particular problem faced by a bolometer instrument operating at altitude near the South Pole is the enhanced ionising radiation background due to charged particles from the solar wind that are funnelled down onto the poles by the Earth's magnetic field. When a bolometer is struck by an ionising particle, it causes a glitch in the output signal due to the pulse of energy absorbed. These events are rare on the ground but would be much more common and troublesome for BOOMERanG.

The best way of coping with ionising particles make them hit the bolometer as infrequently as possible, by minimising its cross section. But the absorber size needs to be a few wavelengths, presenting a significant cross section for impacts. The solution adopted by the BOOMERanG team was to make the absorber mostly empty space as far as the ionising radiation is concerned but to behave as though it were a continuous absorbing sheet as seen by the incident millimetre-wave radiation. This was done by fabricating the absorber as a fine mesh structure with gaps much smaller than the wavelength. This arrangement behaves electromagnetically in exactly the same way as a continuous metal sheet. Another advantage is that the absorber mass (and thus its heat capacity) is kept to a minimum. The 'spider-web' bolometer (Mauskopf et al. 1997), based on this principle, was developed for BOOMERanG (Fig. 7.6). The absorber was a metalised silicon nitride structure fashioned after a spider web (a configuration developed by nature to provide maximum mechanical robustness with minimum mass) and the thermometer was an NTD germanium crystal. BOOMERanG flew successfully in 1998, mapping a sky area of over 2000 sq. deg., with 16 feedhorn-coupled spider-web bolometers operating in four bands between 90 and 410 GHz with *NEP*s of typically 4×10^{-17} W $\text{Hz}^{-1/2}$ (Crill et al. 2003). It flew again in 2003 with an added capability to measure polarisation. Both flights produced very high-quality CMB power spectra, which were later improved upon only by the WMAP and *Planck* satellites.

The development of sensitive cryogenic bolometers for far-infrared and submillimetre detection has been paralleled by the use of similar devices for the detection of X-rays. The same basic principle of thermal detection used in a bolometer can also be used to detect single X-ray photons and measure their energy, a process known as calorimetry (e.g. Moseley et al. 1984; Porter 2013). The calorimeter is very similar to the bolometer, but the absorber is optimised for X-ray instead of far-infrared absorption. The deposition of energy when an X-ray photon is absorbed is quickly converted to heat, causing a rapid jump in the absorber temperature and a corresponding change in the thermometer resistance. Many of the requirements relating to device design and materials, cryogenic

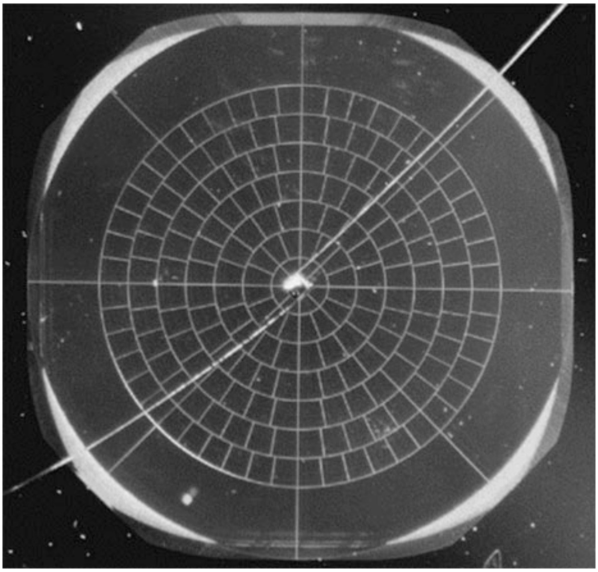


Fig. 7.6 Photograph of an early spider-web bolometer. © Mauskopf et al. (1997). The 5.6 mm diameter absorbing mesh was made of 1-μm-thick silicon nitride, with a metalised coating. The long radial legs were 1 mm long and 5 μm wide, and the legs in the mesh were 200 μm long and 4 μm wide. The NTD germanium thermometer was attached in the centre and read out with two long Nb-Ti lead wires. © JPL

cooling and signal readout and processing are common to far-infrared and X-ray applications, and the instrumentation communities in both domains have maintained close relationships and collaborations. X-ray calorimeters using silicon-based detectors have been successfully used in rocket instruments (e.g. Porter et al. 2000) and on the Japanese Hitomi satellite (launched in Feb. 2016, but unfortunately lost shortly afterwards due to a fault in its attitude control system) and are planned for the European ATHENA X-ray observatory (using superconducting transition edge thermometers similar to the ones described below).

7.5 The *Herschel* SPIRE Instrument

At the end of the 1990s, FIRST had been studied for a number of years. It was already clear that three instruments would be appropriate, with two of them based on direct detection and covering the entire wavelength range. One (which became PACS) would cover the shorter-wavelength part of the range (up to around 200 μm), and the other (which became SPIRE) would operate at longer wavelengths. Each would have a multi-band camera and a spectrometer. The third instrument (which became HIFI) would be a single-pixel high-resolution heterodyne spectrometer covering as much of the wavelength range as possible.

ESA issued the Announcement of Opportunity for the FIRST instruments in October 1997, with a response due in early 1998. Potential instrument consortia had already been forming and were working on instrument concepts. For SPIRE, some key requirements and design features were already clear. The camera would have three bands, nominally centred at 250, 350 and 500 μm, operating simultaneously for maximum data gathering efficiency. It would have the largest possible field of view, at least 4 × 4 arcminutes and more if possible. The spectrometer would be a Fourier transform spectrometer (FTS), covering at least 200–400 μm and preferably extending to longer wavelengths. Bolometric detectors were the only option, with the requirement to achieve photon noise-limited sensitivity under the background coming from the 80-K FIRST telescope. This required the bolometers to be operated at 0.3 K.

At that time, a feedhorn-coupled bolometer system, similar to those deployed on ground-based and balloon-borne telescopes, constituted the most mature and well-developed option, although many engineering challenges would need to be addressed to enable large arrays to be flown in space. But two alternative technologies, both based on absorber-coupled arrays, were emerging as potential alternatives. It was not at all clear which detector technology and focal plane architecture would prove to be the best option, and this posed a major debating point for the SPIRE team.

7.6 Superconducting Transition Edge Bolometers

An alternative to the semiconductor thermometer is the superconducting transition edge sensor (TES). A superconducting material exhibits a very sharp transition between the normal (resistive) and superconducting states at a well-defined transition temperature. So there is a very strong dependence of the resistance on temperature in the transition region between the two states, as shown in Fig. 7.7. This makes the TES an excellent thermometer. The TES thermometer is a thin superconducting film, set up to operate in the normal-superconducting transition region. Operation on the sharp transition means that there must be a fixed value of total power, equal to the power needed to warm the TES from the heat sink temperature to the transition temperature. The sum of the electrical bias power and radiant background power is thus always equal to this value. Changes in the absorbed radiant power due to the signal are compensated by opposite changes in the bias power (i.e. in the current for a fixed voltage). This change in current is read out using a sensitive current amplifier known as a SQUID (superconducting quantum interference device), which is inductively coupled to the TES, as shown in Fig. 7.7. At the time of the SPIRE proposal, this was a relatively new technology which had not yet been deployed in a working instrument. A team from NASA GSFC proposed to implement absorber-coupled arrays of bolometers, of the sort used in SHARC-II, but using the more sensitive TES sensors instead of silicon thermometers. A cold SQUID multiplexer could allow a number of detectors to be sampled using the same readout lines.

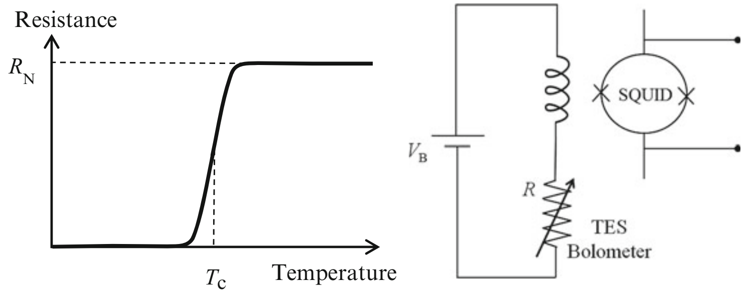


Fig. 7.7 *Left*: Principle of the TES thermometer. Resistance varies rapidly with temperature as metal goes through the transition from the normal (resistance equal to R_N) to superconducting (zero-resistance) states; *Right*: Essential features of a TES biasing and readout circuit incorporating inductive coupling to a SQUID. © Cardiff University

7.7 Micromachined Silicon Bolometers

Another novel absorber-coupled bolometer array technology was also under development as an option for SPIRE by the CEA/SAp group in Saclay in collaboration with LETI, Grenoble. This technology, illustrated in Fig. 7.8, used metalized silicon grids supported by silicon beams designed to provide the appropriate thermal conductance to the silicon substrate (Agnese et al. 2003). An ion-implanted thermometer was located in the centre of each grid and another on the substrate, with a readout that compared the impedances of the two. The readout and multiplexer used silicon CMOS transistors operating at 2 K. CMOS transistors are very convenient for such readout and multiplexing functions, but are relatively noisy. To overcome this, the detector impedance was made very high ($>10^9$ Ω) to give it a high responsivity, so boosting the photon noise above the readout noise level. Indium bump bonds made the electrical connections between the detector array chip and the readout chip. This technology had the key advantage that the entire array could be fabricated using micromachining techniques.

At the time when the SPIRE proposal was submitted, in early 1998, it was not possible to decide which option would be best, so the question was left open. The institutes developing the competing technologies were all willing to embark on a 2-year programme of modelling, experimental development and testing, with selection of the flight array technology to take place in 2000, and followed by the detailed development, space qualification and flight hardware provision of the chosen technology.

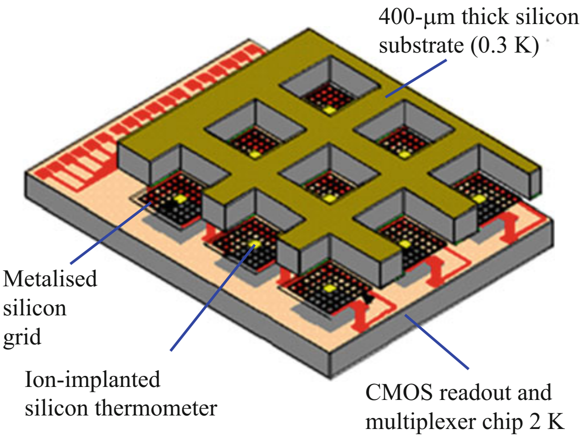


Fig. 7.8 Design of the CEA/LETI micromachined bolometers. © CEA

7.8 Absorber-Coupled vs Feedhorn-Coupled Arrays

One of the key issues that needed to be understood by the SPIRE team was the fundamental relative merits of feedhorn-coupled vs absorber-coupled arrays, something that was studied in detail during the array development programme.

As noted above, highest feedhorn efficiency is achieved for a horn diameter close to $2F\lambda$. Each feedhorn has a beam on the sky which is approximately λ/D in size, where D is the telescope diameter. Although the horns are close-packed in the focal plane, their beams on the sky are separated by $2\lambda/D$ and so do not overlap. The image is therefore not instantaneously fully sampled—either jiggling of the array (several separate telescope pointings) or scanning of the telescope is needed to create a fully sampled image.

The alternative absorber-coupled configuration dispenses with feedhorns, just filling the focal plane with an array of square bolometers of size $0.5F\lambda$, so the individual detector beams overlap on the sky giving instantaneous full sampling of the image. The individual pixels have low coupling efficiency to a point source, around 15%, because they are smaller than the diffraction limit. But they are close-packed and whatever is missed by one pixel is picked up by another. In principle, all of the power in the focal plane can be absorbed, and the total emission from a point source captured by adding up the signals in several pixels over which it is spread.

The main advantages of the feedhorn array lineage are as follows: (1) best efficiency for detection of a point source with known position because as much of the power as possible is concentrated onto one pixel, (2) good stray light rejection in which the bolometer field of view is restricted to the telescope and (3) minimum number of detectors for a given total field of view, so minimising the power dissipation, the number of wires to the cold focal plane and the data rate. The main disadvantages are that the observing modes are complicated (repointing or scanning are needed) and, more importantly, the efficiency for mapping is less than the ideal because the feedhorn acts as an antenna with some of the incident power rejected due to destructive interference.

The main advantages of the filled array lineage are as follows: (1) instantaneous full sampling of the image, with no need for jiggling or scanning, and (2) better overall sensitivity for mapping because no power is rejected, a theoretical factor of 3.5 for point source extraction if the detector noise is negligible, decreasing to 2 if the detector and photon noise *NEP*s are equal (Griffin et al. 2002). Significant disadvantages are as follows: (1) around four times lower background per detector due to the small pixels, giving a photon noise *NEP* which is lower by a factor of 2, and so more difficult to achieve, and (2) 16 times more detectors needed for a given field size than the $2F\lambda$ feedhorn design, putting greater strain on instrument resources such as power, wiring and data rate.

It may be surprising that increasing the number of detectors by a factor of 16 does not result in a comparable improvement in observing speed. The improvement is more modest because individual detectors collect much less power due to their small size and the co-addition of the signals (and noise) from a number of pixels to recover the total emission from a point source results in a higher overall noise level than one gets when all of the signal is intercepted by a single detector.

7.9 The SPIRE Detector Development Programme

At a meeting of the SPIRE consortium at Queen Mary College (QMC), London, in October 1997, the three options were reviewed and it was agreed that a two-year programme involving development and experimental evaluation of the three technologies would be implemented, with regular meetings and progress reviews. For the consortium's proposal to ESA, due in February 1998, it was decided that absorber-coupled planar arrays of bolometers would be the baseline technology, based on the potentially higher mapping speed. The question of which kind of bolometers—NTD germanium, TES or micromachined silicon—was to be left open in the proposal.

The three array options would be developed until late 1999. The three groups were to build demonstrator systems, which would all be evaluated in the laboratory at QMC in a standard test setup in preparation for the final array selection, planned for early 2000. It was agreed that the decision on which technology to fly would be based on the demonstrated performance of the arrays, the existence of a detector system design compatible with the SPIRE instrument, the ability to build enough of sufficient quality and confidence that they could be space qualified. In the event that none of the planar array technologies was sufficiently advanced for inclusion in the instrument, feedhorn-coupled spider-web bolometers with NTD germanium thermistors would be the fallback option. In the consortium's proposal, the different options were described in detail, as was the plan for evaluating them and selecting the best technology compatible with the instrument requirements and budgets and the mission schedule.

Over the next two years, the different array teams worked intensively to develop their array designs, to build and test prototypes and finally to demonstrate them at QMC in the standard setup. In preparation for the selection, detailed criteria and performance specifications were defined, and an array selection panel was appointed, including members of the SPIRE team, independent technical and scientific experts and ESA representatives. The array selection meeting was held at the Rutherford Appleton Laboratory in Oxfordshire, UK, in February 2000. A key point was the available improvement in mapping speed from the use of filled arrays. With a maximum theoretical factor of ~3.5, and closer to 2 in the case of SPIRE, because the detector noise would not be negligible, the advantage could easily be lost if the sensitivity of the filled array pixels were degraded even by a small amount (as observing time is proportional to the square of the *NEP*, an *NEP* degradation by a factor of 1.5 would cancel out the expected mapping speed gain). Given the maturity of the feedhorn-coupled option, it was thus seen as critical that the filled array options demonstrate convincingly that they could achieve the required sensitivity.

The TES option still had some problems with excess noise and was judged technically rather complex, less mature and more risky because TES sensors were vulnerable to becoming completely inoperable if the telescope background turned out to be higher than expected. The micromachined Si option faced problems in achieving the lower (more challenging) filled array *NEP* target and still exhibited excess noise. In addition, it had slower speed of response, which could degrade the photometer performance in scanning mode, and was not compatible with the speed of response requirements for the spectrometer. For the feedhorn-coupled spider-web NTD detectors developed for SPIRE by Caltech, JPL and the University of Colorado (Turner et al. 2001), it was clear the all key specifications could be met. While it was recognised that both the TES and micromachined silicon technologies had great potential, the panel decided that given the technical and schedule requirements for the mission, and the fact that the key science requirements could be met or surpassed by the feedhorn-coupled NTD bolometer option, it should be selected for SPIRE.

It is worth noting some of the personal and political aspects of this decision. The array selection was a difficult and painful process at the time, with everyone keenly aware of the enormous and sustained efforts made by the detector teams during the previous two years. The stakes were high for the SPIRE team as a whole, with the need to choose the best detectors, and for the detector teams with the prospect of building the detectors for a major space instrument, and the academic success and funding that this would bring to their laboratories and institutes. In addition, there was a natural preference in Europe for the selection of the micromachined silicon detectors as a European rather than a US option, given that this was a European-led mission. Given all of these factors, the selection decision could even have destabilised the SPIRE consortium, but the professionalism of the teams and their commitment to the success of the instrument and to the mission and its science, ensured that they remained on board.

Following the selection by the PACS and SPIRE teams of the detector technologies, many new challenges still had to be addressed. One of the most difficult was the need to connect the ultra-cold detectors to their housing in a way that provided extremely strong mechanical support (to withstand launch vibration) while at the same time having a very weak thermal link to ensure minimal thermal load on the ^{3}He refrigerator. The PACS and SPIRE teams developed a mechanical support system based on the use of Kevlar®, a composite material with very low thermal conductivity and which can be spun as a thread which is extremely strong in tension. Great efforts were also needed to implement the cold electronics needed to read out the low-level detector signals for relay to the warm electronics. In the case of SPIRE, a particular challenge was posed by the required operating temperature (about 120 K) of the readout transistors, which also had to be located as close as possible to the much colder bolometers. This problem was solved by putting the transistors in a special housing located just outside the main instrument enclosure, allowing them to operate optimally without interfering with the bolometers.

The intensive technology development did not go to waste for the unselected options. In a remarkable development shortly afterwards, it was realised that the micromachined silicon detectors, as developed for SPIRE by CEA, could prove to be well-optimised for the PACS camera (see below). TES technology was further developed by the GSFC and other groups and has since became the technology of choice for most ground-based submillimetre instruments today and also for the next-generation FIR space instrument, SPICA-SAFARI.

The design of the SPIRE camera, together with its detector and filter technology, formed the basis of a balloon-borne experiment which flew successfully as a scientific pathfinder for SPIRE: the Balloon-borne Large Aperture Submillimeter Telescope (BLAST; Pascale et al. 2008). Given the shorter development timescales for balloon projects, it was possible to build and fly BLAST a number of times before the launch of *Herschel*. With its smaller (2-m diameter) and warmer telescope, the effects of the residual atmosphere even at balloon altitudes, and the limited amount of observing time, BLAST was not intended to compete with SPIRE scientifically but to provide a valuable, and at the time unique, foretaste of the science that SPIRE would enable.

For the latest generation of ground-based instruments operating in the submillimetre, absorber-coupled arrays, usually of superconducting detectors, have become the norm in the case of systems optimised for general astrophysics, such as SCUBA-2 (Holland et al. 2013) and NIKA (Monfardini et al. 2013). For instruments designed for CMB studies, in which very good control of the telescope beam, with extreme suppression of sidelobes, is important, feedhorn- or other antenna-coupled systems are normally preferred—e.g. for the South Pole Telescope (Chang et al. 2012) and the BICEP-2 and Keck experiments (Ade et al. 2015).

7.10 CEA Silicon Bolometers: From SPIRE to PACS

During the selection of SPIRE bolometers, the PACS team was experiencing a crisis because the photoconductive detector technology that they had planned to use for their camera had a major technical problem with the readout electronics, meaning that it could not meet the instrument requirements. Although the silicon bolometers had struggled to meet the SPIRE sensitivity specifications, the PACS specifications were not so stringent. PACS would operate at shorter wavelengths where the telescope background would be higher, producing a higher photon noise-limited *NEP*, and so easier to achieve, and the SPIRE evaluation programme had shown that the micromachined silicon bolometers could actually meet the PACS specifications.

The CEA team leading the silicon bolometer development wanted to secure their investment by proposing their SPIRE-like prototype to become part of an instrument on SOFIA, a joint German-US airborne observatory. Unfortunately, SOFIA instruments were already decided and complete. However, Albrecht Poglitsch at MPE, Garching, principal investigator of PACS and engaged in SOFIA, realised that the CEA bolometer array could be suitable for the PACS camera: ‘We have two choices: we have a guaranteed failure or at least we have a certain chance for success’. The filled array bolometer camera came as a last-minute innovation, saving the PACS camera. The adaptation to PACS requirements was not at all straightforward. Major technical adjustments were needed, and consortium rules about the time allocation for each consortium partner had to be reevaluated (see Chap. 10).

The PACS team and the CEA Saclay group carried out a quick study to determine if such bolometers could be used in PACS. Although many technical issues would have to be addressed (not least the incorporation of a ^{3}He cooler, similar to the one being developed for SPIRE, into the PACS focal plane unit), it indeed proved to be feasible and attractive, and these detectors were eventually flown in PACS with great success. In addition, they have since been used both in a balloon-borne polarimeter experiment, PILOT (Misawa et al. 2014) and in ARTEMIS, a large-format submillimetre camera for the APEX telescope in Chile (Revéret et al. 2014; André et al. 2016).

Figure 7.9 shows photographs of the SPIRE and PACS flight array designs. The SPIRE and PACS bolometer arrays performed superbly in flight, fully justifying the great care and efforts that had gone into their design and development.

7.11 Bolometer Technical Lineages

To quantify and illustrate the advances in the capability of the various bolometer types and architectures, we can define several figures of merit for a bolometer system. At the level of the observatory, the overall performance depends on the

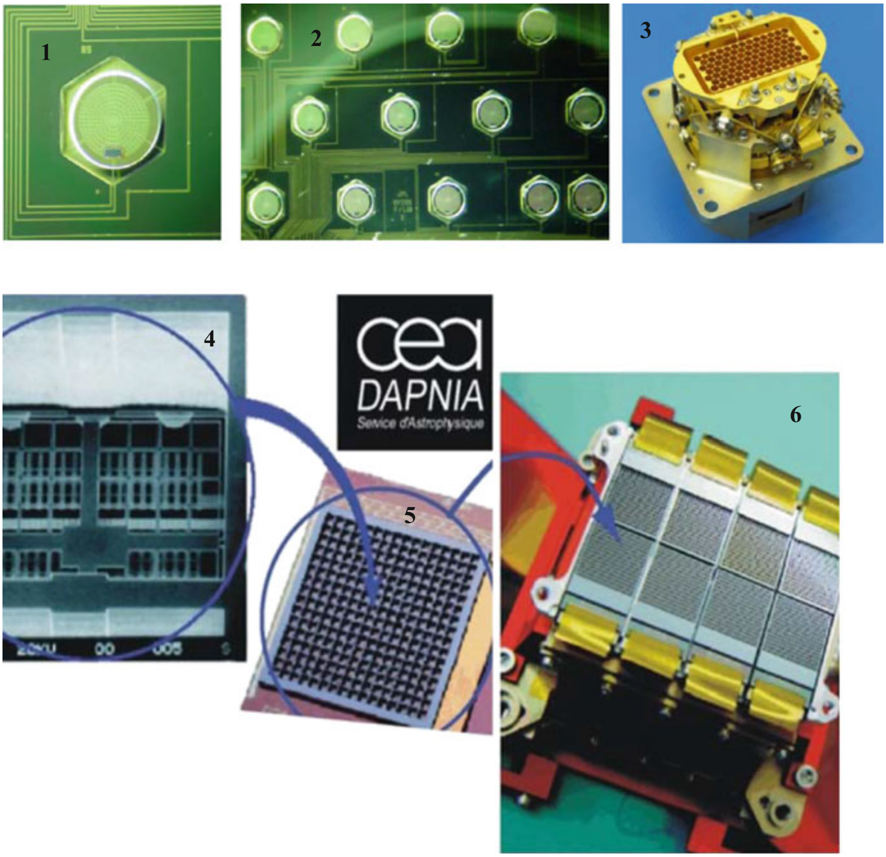


Fig. 7.9. Clockwise from *top left*: (1) An individual SPIRE spider-web bolometer. (2) Part of a SPIRE array of such bolometers. (3) A SPIRE 350-μm array module showing the array of 88 feedhorn apertures and the Kevlar threads used to support the bolometers and feedhorns at 0.3 K from the 2 K stage. (4) Close-up of an individual PACS bolometer. (5) A 16 × 16 sub-array. (6) Eight such sub-arrays butted together to form a 16 × 64 (1024-element) array.

characteristics of the Earth's atmosphere (if it is Earth-based), the telescope, the instrument, the detectors, the readout electronics and the data processing software. The mapping time (the time to map a given sky area to a given sensitivity), is a common figure of merit. It includes telescope and instrument characteristics such as angular resolution, field of view, optical efficiency and observing modes such as the scanning strategy. Considering the detector system alone, its contributions to the overall sensitivity depend on the *NEP*, the number of pixels and the aperture efficiency per pixel. We can therefore define a figure of merit for the detectors given by the product of 1/*NEP*, pixel count and aperture efficiency. Finally, at the bolometer level, the *NEP* as a measure of its intrinsic sensitivity appears as a good

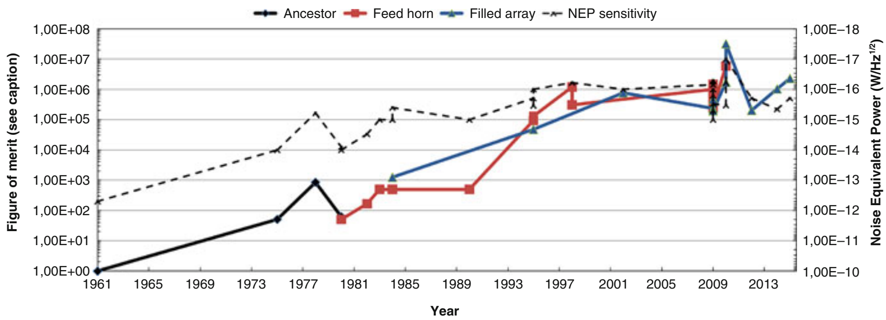


Fig. 7.10 Bolometer technical lineages. The *left y*-axis represents a figure of merit defined as the product of 1/NEP, pixel count and aperture efficiency. The *black curve* is for the ancestor, *red curve* for the feedhorn bolometer arrays and the *blue curve* for the filled arrays. The *right-hand y*-axis is the NEP (the *lower* the value, the better the sensitivity). The *dashed plot* shows the NEP for all bolometers. © Cardiff University (M. Griffin) & CEA (V. Minier)

figure of merit. Such figures of merit are plotted in Fig. 7.10 for various systems, starting with the first cryogenic bolometer of Low (1961).

It is important to note that these figures of merit refer to different levels: the observatory (telescope to detectors), the bolometer camera (number of detectors taken into account) and the bolometer itself. Important factors which are not taken into account are the optical efficiency and bandwidth of the instrument, the dependence of photon noise on the wavelength of operation and the presence of excess noise, caused by atmospheric turbulence, over and above the atmospheric photon noise level that would be expected for a quiescent atmosphere.

In Fig. 7.10, the bolometer camera figure of merit is represented as three lineages: the ‘ancestors’, the ‘feedhorn’ and the ‘filled array’. The feedhorn and filled array lineages have progressed in parallel. Interestingly, the feedhorn lineage increased in performance from 1980 to 1997 and then maintained a nearly constant level. This could be interpreted as invention and performance progress until the saturation point of this technology. The filled array concept seems to have followed a comparable path with a delay of few years. However, it is not clear whether they saturate or still progress. The bolometer *NEP* shows a comparable evolution, meaning that innovation of bolometer instruments as detectors is mainly triggered by innovation within bolometer detectors themselves and driven by the need to achieve photon noise-limited sensitivity at lower and lower backgrounds.

The best choice of technology depends on the particular characteristics of the experiment, and it is expected that both types will continue to be used. Future developments, beyond the scope of this work, will lead both technologies to evolve. For instance, the number of pixels which a filled array can have will be increased with the adoption of a new kind of detector, the kinetic inductance detector (KID), as an alternative to the bolometers described here. With KIDs it is possible to read out a much larger number of detectors with a single readout line. Likewise, the inconveniently high mass of feedhorns can be reduced by replacing them with lighter planar or lenslet-coupled antennas. The detector *NEP* will also continue to show a downward evolution as it becomes possible to cool space telescopes to lower temperatures, reducing the photon noise limit.

7.12 Conclusions

The semiconductor bolometers flown on *Herschel* rested on many years of innovation and practical experience with Earth-based instruments. Three types of innovation are identified for describing bolometer evolution since 1961. According to the typology of Chap. 5, these are: *synergy innovation* (coupling of different pre-existent technical lineages, such as bolometer and cryogenic technologies used for ground-based systems, to perform a common task in a space environment), *incremental innovation* (gradual progress of a technical lineage such as the evolution of the bolometer with a Winston cone to feedhorn bolometer arrays) and *disruptive innovation* (major progress of a technical lineage, either realised by a global reorganisation of the object or by substitution of a new lineage for another to perform the same task, such as Low's bolometer as the common ancestor and then filled bolometer arrays as an alternative to feedhorn bolometer arrays).

The question of whether an absorber-coupled array or an antenna-coupled array is best is one that does not have a single answer—it depends critically on the application and the technical constraints and scientific requirements. The choice of detectors for a space instrument is a complex exercise in which many technical, scientific, schedule and financial factors have to be considered. The SPIRE detector development and selection programme was an intensive and high-pressure activity which produced an optimised choice for the instrument. The technical lineage for the SPIRE bolometers ended up as an extension of the traditional and well-tried solution of using feedhorn-coupled NTD germanium detector arrays. More novel technical lineages involving filled arrays of micromachined silicon bolometers or TES sensors were also evaluated, and, in what was an unexpected surprise, this produced a technical solution for the PACS camera too. The selection process also paved the way for future innovation as represented by the TES technology selected for the forthcoming SPICA mission. But semiconductor detectors of the sort used on *Herschel* are becoming superseded by superconducting sensors, which do offer superior sensitivity.

In summary, the SPIRE bolometer programme triggered two types of successful innovation in our typology (Chap. 5): an incremental innovation (the SPIRE bolometers) and a breakthrough/disruptive innovation (the PACS bolometers). The feedhorn-coupled arrays turned out to be better for one application and the absorber-coupled configuration better for another. The same has proved to hold for subsequent state-of-the-art instruments, with both architectures in widespread use today.

References

Ade PAR, Griffin MJ, Cunningham CT, Radostitz JV, Predko S, Nolt I (1984) The Queen Mary College/University of Oregon photometer for submillimetre continuum observations. Infrared Phys 24:403–415

Agnese P, Cigna C, Pornin J-L, Accomo R, Bonnin C, Colombel C et al (2003) Filled bolometer arrays for Herschel/PACS. Proc SPIE 4855:108–114

André P, Revéret V, Könyves V, Arzoumanian D, Tigé J, Gallais P et al (2016) Characterizing filaments in regions of high-mass star formation: high-resolution submilimeter imaging of the massive star-forming complex NGC 6334 with ArTéMiS. Astron Astrophys 592:A54

BICEP2 Collaboration, Keck Array Collaboration, SPIDER Collaboration, Ade PAR, Aitkin RW, Amiri M, Barkats D, Benton SJ, Bischoff CA et al (2015) Antenna-coupled TES bolometers used in BICEP2, keck array, and SPIDER. Astrophys J 812:176

Chang CL, Ade P, Aird K, Austermannc J, Beall J, Becker D et al (2012) Detectors for the south pole telescope. Phys Procedia 37:1381

Clarke G, Hoffer I, Richards PL (1974) Superconducting tunnel junction bolometers. Rev Phys Appl 9:69

Coron, N., Dambier, G., and Leblanc, J. (1972) A new type of helium-cooled bolometer. In: Manno V, Ring J (eds) Proceedings of the infrared detector techniques for space research. Reidel, Dordrecht, p 121

Crill BP, Ade PAR, Artusa DR, Bhatia RS, Bock JJ, Boscaleri A, Cardoni P et al (2003) BOOMERANG: a balloon-borne millimeter-wave telescope and total power receiver for mapping anisotropy in the cosmic microwave background. Ap J Suppl 148:527

Dowell CD, Allen CA, Babub S, Freund MM, Gardner MB, Groseth J et al (2003) SHARC II: a Caltech Submillimeter Observatory facility camera with 384 pixels. Proc SPIE 4855:73

Downey PM, Jeffries AD, Meyer SS, Weiss R, Bachner FJ, Donnelly JP, Lindley WT, Mountain RW, Silversmith DJ (1984) Monolithic silicon bolometers. Appl Opt 23:910

Duncan WD, Robson EI, Ade PAR, Griffin MJ, Sandell G (1990) A millimetre/submillimeter common user photometer for the James Clerk Maxwell Telescope. Mon Not R Astron Soc 243:126

Glenn J, Bock JJ, Chattopadhyay G, Edgington SF, Lange AE, Zmuidzinas J, Mauskopf PD, Rownd B, Yuen L (1998) Bolocam: a millimetre-wave bolometric camera. Proc SPIE 3357:326

Glenn J, PAR A, Amariec M, Bock JJ, Edgington SF, Goldind A, Golwala S, Haig D, Lange AE, Laurent G, Mauskopf PD, Nguyen H (2003) Current status of Bolocam: a large-format millimeter-wave bolometer camera. Proc SPIE 4855:30

Griffin MJ, Bock JJ, Gear WK (2002) Relative performance of filled and feedhorn-coupled focal-plane architectures. Appl Opt 41:6543

Haller EE, Palaio NP, Rodder M (1984) NTD germanium: a novel material for low-temperature bolometers. In: Proceedings of the 4th international neutron transmutation doping conference, Gaithersburg, Maryland, 21

Harper DA, Hildebrant RH, Winston R, Stiening R (1976) Heat trap—an optimized far infrared field optics system. Appl Opt 15:53

Holland WS, Robson EI, Gear WK, Cunningham CR, Lightfoot JF, Jenness T et al (1999) SCUBA: a common-user submillimetre camera operating on the James Clerk Maxwell telescope. Mon Not R Astron Soc 303:659

Holland WS, Bintley D, Chapin EL, Chrysostomou A, Davis GR, Dempsey JT et al (2013) SCUBA-2: the 10000 pixel bolometer camera on the James Clerk Maxwell Telescope. Mon Not R Astron Soc 430:2513

Jones RC (1953) The general theory of bolometer performance. J Opt Sot Am 43:1

Kreysa E, Haller EE (1982) He-3 bolometers for millimeter and submillimeter photometry. In: Kjaer K, Moorwood AFM (eds) Proceedings of the 2nd ESO infrared workshop, Garching, April 20–23. European Southern Observatory (ESO), Garching, p 197

Kreysa E, Gemünd H-P, Gromke J, Haslam CGT, Reichertz L, Haller EE et al (1999) Bolometer array development at the Max-Planck-Institut für Radioastronomie. Infrared Phys Technol 40:191

Low FJ (1961) Low temperature germanium bolometer. J Opt Soc Am 51:1300

Mather J (1982) Bolometer noise: non-equilibrium theory. Appl Opt 21:1125

Mauskopf PD, Bock JJ, Del Castillo H, Holzapfel WH, Lange AE (1997) Composite infrared bolometers with Si_3N_4 micromesh absorbers. Appl Opt 36:765

Misawa R, Bernard J-P, Ade P, André Y, de Bernardis P, Bouzit M et al (2014) PILOT: a balloon-borne experiment to measure the polarized FIR emission of dust grains in the interstellar medium. Proc SPIE 9153:91531H

Monfardini A, Adam R, Adane A, Ade P, André P, Beelen A et al (2013) Latest NIKA results and the NIKA-2 project. J Low Temp Phys 176:787

Moseley SH, Mather JC, McCammon D (1984) Thermal detectors as x-ray spectrometers. J Appl Phys 56:1257

Moseley SH, Dowell CD, Allen C, Phillips TG (2000) Semiconducting pop-up bolometers for far-infrared and submillimeter astronomy. In: Mangum JG, Radford SJE (eds) Proceedings of the imaging at radio through submillimeter wavelengths. ASP conference proceedings, vol 217, p 140

Nishioka NS, Richards PL, Woody DP (1978) Composite bolometers for submillimeter wavelengths. Appl Opt 17:1562

Pascale E, Ade PAR, Bock JJ, Chapin EL, Chung J, Devlin MJ, Dicker S, Griffin M, Gundersen JO, Halpern M, Hargrave PC, Hughes DH, Klein J, MacTavish CJ, Marsden G, Martin PG, Martin TG, Mauskop P, Netterfield CB, Olmi L, Patanchon G, Rex M, Scott D, Semisch C, Thomas N, Truch MDP, Tucker C, Tucker GS, Viero MP, Wiebe DV (2008) The balloon-borne large aperture submillimeter telescope: BLAST. Ap J 681:400

Porter FS (2013) X-ray calorimeters. Observing photons in space: a guide to experimental space astronomy, ISSI Scientific Report Series, vol 9, chap 28

Porter FS, Almy R, Apodaca E, Figueroa-Feliciano E, Galeazzi M, Kelley R, McCammon D, Stahle CK, Szymkowiak AE, Sanders WT (2000) The XQC microcalorimeter sounding rocket: a stable LTD platform 30 seconds after rocket motor burnout. Nucl Inst Methods Phys Res A 440:220

Radostitz JV, Nolt IG, Kittel P, Donnelly RJ (1978) Portable ^{3}He detector cryostat for the far infrared. Rev Sci Instrum 49:86

Revéret V, André P, Le Pennec J, Talvard M, Agnèse P, Arnaud A et al (2014) The ArTéMiS wide-field sub-millimeter camera: preliminary on-sky performance at 350 microns. Proc SPIE 9153:915305

Richards PL (1994) Bolometers for infrared and millimeter waves. J Appl Phys 76:1

Roellig T, Houck JR (1983) A helium-3 cooled bolometer system for one millimeter continuum observations. Int J Infrared Millimeter Waves 4:299

Serlemitsos A (1988) Flight worthy infrared bolometers with high throughput and low NEP. Proc SPIE 0973:314

Turner AD, Bock JJ, Beeman JW, Glenn J, Hargrave PC, Hristov VT, Nguyen HT, Rahman F, Sethuraman S, Woodcraft AL (2001) Silicon nitride micromesh bolometer array for submillimeter astrophysics. Appl Opt 40:4921

Wang N, Hunter TR, Benford DJ, Serabyn E, Lis DC, Phillips TG et al (1996) Characterization of a submillimeter high-angular resolution camera with a monolithic silicon bolometer array for the Caltech Submillimeter Observatory. Appl Opt 35:6629

Winston R (1970) Light collection within the framework of geometrical optics. J Opt Soc Am 60:245

Chapter 8
Heterodyne Technology in Submillimetre Astronomy: Towards Implementation in *Herschel*

Abstract In this chapter, we introduce the heterodyne technique and its early history. The *Herschel*-HIFI instrument involved two important technical lineages that were just emerging at the time of the initial mission proposal to ESA in 1982: superconducting mixers and local oscillator (LO) devices. We review the 15 years of progress and development carried out to advance these technologies before the HIFI instrument project began in 1997. The status and performance of the HIFI instrument are reviewed, and HIFI's huge improvements in sensitivity and broad frequency coverage, securing both the robustness and reliability, which are mandatory for a successful space mission, are discussed. Like the other two *Herschel* instruments, HIFI represents a powerful example of how enabling technology and programmatic opportunities can fruitfully interact in a space project.

The Heterodyne Instrument for the Far-Infrared (HIFI) is probably the most complex instrument flown so far on a scientific satellite. Its development covered several technical lineages that were just emerging at the time of the mission proposal (1982) and were fully developed by the time of launch. The two key technology development areas were the superconducting mixers (SIS[1] and HEB[2]) and the planar integrated circuit local oscillator (LO) devices. HIFI represents a typical example of how, in a space project, enabling technology and programmatic opportunity interact over long-term periods to major developments in the key heterodyne components that ultimately provided huge improvements in sensitivity, broader frequency coverage and, last but not the least, the robustness and reliability which are mandatory for a successful space mission. This chapter first introduces the heterodyne technique followed by a review of the innovation paths for mixers and local oscillator. Then the chapter reviews the technology status used for the FIRST assessment study in 1983 and the development in the following 15-year period in preparation for FIRST. This is completed by the status in 1997 when the model of the heterodyne instrument (HET) was defined and used for the announcement of opportunity. This is followed by a description of the ultimate efforts to develop the mixers and local oscillator.

[1]Superconductor-insulator-superconductor

[2]Hot electron bolometer

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_8

8.1 Principle of the Heterodyne Technique

Radiation from the interstellar gas manifests itself in the form of spectral features at well-defined frequencies (or wavelengths), which are characteristic of the ambient chemical and physical conditions. To reveal further information about the gas chemical composition, temperature and density and its state of motion such as infall, outflow or rotation, high spectral resolution receivers are required to separate spectral lines and retrieve their detailed shape (see Chap. 4). However, obtaining high spectral resolution in the far-infrared (FIR) range of the electromagnetic spectrum is not easy for several reasons. The wavelength of the radiation is much longer than in the visible, making it difficult to use the kind of spectrometers that are built for optical astronomy, because the physical size of such instruments scales to first order with the wavelength. The energy of the FIR photons is also very low making the signals extremely faint and thus requiring cooling of the detection system. In addition, FIR signal frequencies are very high compared to the frequencies involved in conventional radiocommunication and radio astronomy, posing major technical challenges. For example, direct electrical amplification of these high-frequency signals is essentially impossible. To overcome these problems, heterodyne mixing technique can be used.

Heterodyne detection was invented in the beginning of the twentieth century and was applied in the radio and optical ranges of the electromagnetic spectrum. The Canadian engineer Reginald Fessenden was the first to demonstrate the principle in the radio-frequency domain in 1901 (Fessenden 1901), while a decade later, Edwin Howard Armstrong applied the heterodyne technique to build radio receivers during World War I (Armstrong 1918). Since 1930, almost all radio receivers include a heterodyne stage. On the other hand, Ted Forrester initiated the study of optical heterodyne mixing, more than half a century after the radio heterodyne technique was established. The first experimental demonstration was made by Forrester et al. (1955), and this experiment was of crucial importance for furthering research in the coherence properties of light and its applications. Forrester was also the first to point out the potential of the heterodyne technique for spectroscopy. Soon after H.G. van Bueren in the Netherlands and C.H. Townes in the USA started developments for applications in astronomy. The development of heterodyne technology for the FIR/submm/mm wavelength range used much from the developments in the two adjacent wavelength ranges.

In the radio domain, the electromagnetic field (EM) of the source is captured in an antenna and subsequently conveyed to the mixer (a non-linear signal processing detection element, Fig. 8.1) where the incoming radio frequencies (RF) of interest are combined with the frequency of a monochromatic, internal-generated oscillator, the so-called local oscillator (LO). In the non-linear mixer device, difference frequencies are generated, the so-called intermediate frequency (IF) band of which one IF frequency is simply the difference between the RF and the LO frequency. The IF signal preserves both the amplitude and phase of the original RF signal but at a much lower-frequency range where amplification (IF amp) and

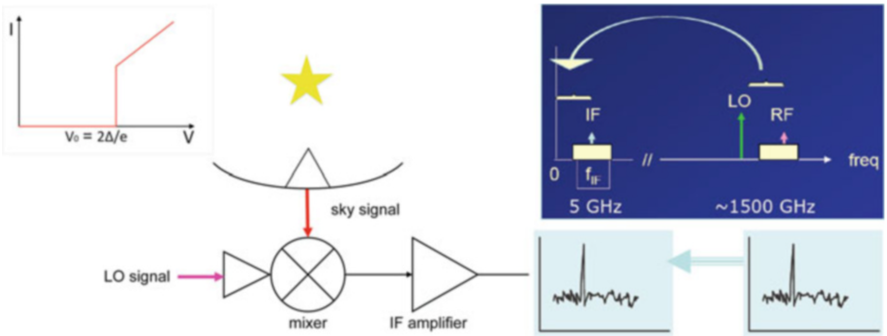


Fig. 8.1 The heterodyne principle in action. In the *left panel*, the sky signal is combined with the LO signal in a non-linear mixing element (with *I-V* curve given in the *top left*). The output signal, at the difference frequency (IF signal), can be amplified electronically. In the *right panel*, it is shown that the down converted signal has the original information content, but now at a much lower frequency, which can be analysed more easily. © HIFI Consortium

electronic analysis is easier for producing a spectrum. The non-linearity of the mixer, arising from a sharp bend in the current (I)-voltage (V) curve,[3] is crucial for the efficiency in generating and converting the difference frequencies. The amplitude of the IF signal depends on how efficient the conversion is (also called conversion gain).

The key subsystems of a heterodyne receiver and their main characteristics, which together determine the science capabilities, are therefore:

– The mixer: it fixes the RF input frequency coverage (RF bandwidth), the IF central frequency and the IF bandwidth. The mixer noise[4,5] temperature and its efficiency in the down conversion (conversion gain) determine the input level for the IF amplifier.
– The local oscillator: it provides the RF frequency tuning range, the frequency and amplitude stability and phase noise. The output power level needs to be compatible with the pumping power needed for the mixer.
– The IF preamplifier: it needs to be compatible with the IF central frequency and bandwidth and ensures the IF amplifier noise temperature and amplification gain.

The sensitivity of the heterodyne receiver is expressed in receiver noise temperature.

[3]*I-V* curve describes the relationship between current and voltage in an electronic device, also called current-voltage characteristic.

[4]Noise is a random fluctuation in an electrical signal.

[5]Noise temperature is the temperature difference between two black bodies that can be detected by a heterodyne mixer per unit time (1 s) and bandwidth (Hz). The number decreases with observing time (t) as $1/\sqrt{t}$ and increased bandwidth (ΔIF) as $1/\sqrt{\Delta \mathrm{IF}}$.

Detailed references for all of these subsystems can be found in Baryshev (2005), Jackson (2005) and Kooi (2008). It is important to add that a heterodyne receiver is a complex instrument in the sense that it requires well-adjusted interplay between its main subsystems, the mixer and LO. For example, in order to have the mixer properly functioning, it needs to receive sufficient radiation power from the LO to utilise the sharp bend in the mixer *I*-*V* curve. A second feature to realise is that the very low energy of the FIR photons requires mixers that can respond to these signal frequencies, while the signals are also extremely faint as is often the case in astronomical observations. The mixers and their materials that fulfil these requirements have been one of the main targets in the described developments.

8.2 Technical Lineages, Risk and Innovation

The Heterodyne Instrument on *Herschel* (HIFI) is a highly complex system of receivers consisting of seven spectral bands ranging in frequency from 480 to 1910 GHz (or 0.48 to 1.91 THz). It is the result of three decades of developments and innovations, initiated in the 1970s and ending in early 2000s. The technology used in HIFI originates for a large part from the heterodyne technology developed for ground-based observatories. Subsequently, the know-how and the developments that led to HIFI benefited to ground-based facilities. For example, the necessity for remote operation and control drove the local oscillator chain and mixer designs as now used in ALMA,[6] thus allowing to operate 66 antennas in a reliable way. Another example is the expertise of several SIS mixer groups acquired during the HIFI development that is now used for the mixers in ALMA (band 9) and SOFIA[7] GREAT.[8] Also software for calculating non-Gaussian optical systems and matching to antennas, for propagation of electromagnetic waves in material, developed for HIFI are now being applied in other domains (e.g. Jellema 2015). These examples show that the synergy between space and ground technology development has been beneficial to both sides.

Flying a heterodyne instrument in space with this level of ambition was a premier. Among the key subsystems presented in Sect. 8.1, potential technologies for mixers and local oscillators had each their own path of development and innovation of different types as identified in Chap. 5. Figure 8.2 displays the two-time evolution tracks of the advances in covered frequency range. This represented the biggest challenge in the programme. Various technical devices can be singled out as the main functioning schemes for local oscillators and mixers, thus allowing the identification of potential technical lineages. The design phase of

[6]ALMA, the Atacama Large Millimeter Array, is an interferometer based in Chile and consists of 66 radio telescopes operating at millimetre and submillimetre wavelengths.

[7]SOFIA: Stratospheric Observatory For Infrared Astronomy

[8]GREAT: German Receiver for Astronomy at Terahertz Frequencies

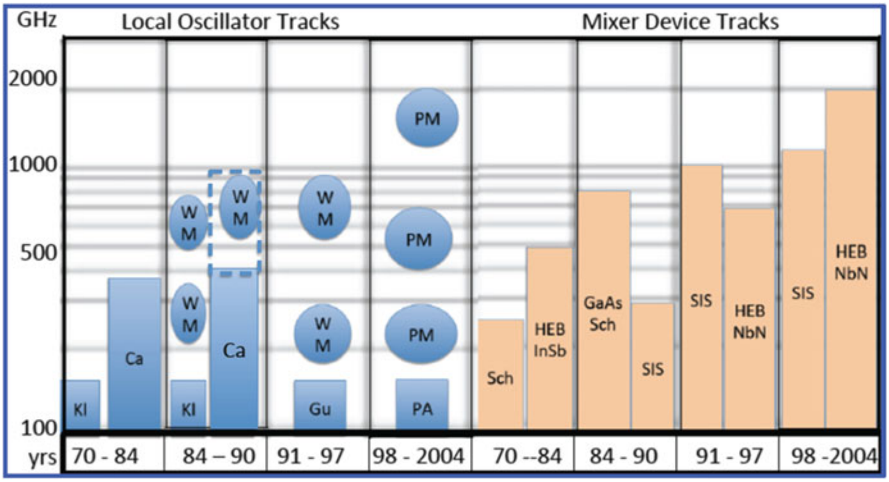


Fig. 8.2 The development tracks for the local oscillators and mixers as has taken place over the years and as was described before. Kl for klystron, Ca for carcinotron (the *dotted line* in the 84–90 slots represents the final experimental results), WM stands for Whisker-contacted diode multipliers, Gu for Gunn oscillator, PM for planar multipliers, PA for power amplifier, Sch for Schottky diode mixer, HEB for hot electron bolometer, SIS for superconductor-insulator-superconductor. © Th. de Graauw

FIRST/*Herschel* has been divided into two periods (1984–1990 and 1991–1997) to show the progress made. The 1998–2004 period corresponds to the time when the final choices were decided. The picture is however a simplified illustration of the many more detailed challenges and achievement that were made. The seven HIFI bands required optimisation for each band. Two types of mixer (SIS and HEB) were selected where the combination of superconducting and wiring material was different for each band, with two types of mixer mounts (waveguide and quasi-optical) and two different IF bands (4–8 and 2.4–4.8 GHz).

Following the typology of Chap. 5, key innovations paved the way of HIFI from early days (1970–1984) to HIFI in-orbit. Some innovations disappeared in the process as they were dismissed or replaced by more promising technical devices. ‘Synergies between innovations’, coupling two different preexistent technical lineages to perform a common task, took place with the cooling of the GaAs Schottky diodes. Operation of these solid-state devices in cryo-conditions reduced the noise temperature by significant factors. In the 1984–1990 period, the introduction of SIS devices as mixers can be characterised as a ‘disruptive innovation’, i.e. a major progress of the mixer technical lineage realised by substitution of a new lineage (SIS) to another (Schottky) to perform the same task. Coming from a development for the superconductor computer requiring fast switching superconducting devices, the proposed application for submm mixers was a real stellar moment. It was in the back of the mind of several persons who were looking for such a new supersensitive type of diode. In the same period, mid-1970s, the introduction of the quasi-optical mixer mounts (bi-conical and cube-corner mixer mounts) was also a disruptive innovation, in particular later on with the inclusion of planar antennas mounted on the back of a lens as also used in HIFI. This innovation was one of the triggers for a complete new field in optical physics: quasi-optics. However, the quasi-optical mount application in receivers was limited due to the incremental innovation of the waveguide technology. New precision fabrication machines with numerically controlled operation and improved electroforming allowed making smaller and more precise mandrills, waveguides and horns. For HIFI, the frequency breakpoint between the two was around 1000 GHz or 1 THz, but it is nowadays still shifting to higher frequencies. The reintroduction in the 1990s of the HEB mixer, using superconducting material, can also be characterised as disruptive innovation. Its impact is enormous, and till to date, it is the only mixer type that functions well in the frequency range beyond 1400 GHz.

For the local oscillator, the innovation concerns two stages of the local oscillator chain: the 100 GHz stage and from there onwards the multiplication stage. For the first stage, we have seen a long struggle with the use (or not) of the vacuum tube oscillators (klystron and carcinotron). Instead, the use of the already existing semiconductor Gunn diode oscillator was not a perfect solution either due to limitations in output power and tuning range. The idea to use monolithic microwave integrated circuit (MMIC) power amplifiers at such high frequencies (100 GHz) was a disruptive innovation enabling HIFI to feed all its 14 multiplier chains at a proper input level and thus covering the required RF range. The development of the multipliers was partly an incremental innovation but had also a substantial disruptive innovation component. Initially the whisker pointed diode multipliers went through a long period of improvement due to a better understanding of the RF circuits with modelling and experiments in trial and error (incremental). The resulting modelling computer programmes were used in the next phase to design the planar devices. The implementation of the high-frequency planar devices was only made possible by the use of the large advances in gallium arsenide (GaAs) IC technology that occurred in the 1990s.

Finally, the design and construction of the mixers and LO chains for the HIFI instrument, with all the crucial ingredients in place, required a large effort with incremental innovation to deliver mixers and LOs for all the frequency bands that had never been built before.

After HIFI was selected and accepted as one of the three *Herschel* instruments, developments for the key technologies continued but became very focused. Within 2 years, there was already a down selection for two items: the high-frequency LO and the superconductor material for the HEB mixer. The first one, the laser mixing LO (Verghese et al. 1997), was dropped because of too little progress in its development, being unattractive for space qualification and also due to the success of the competing technology of the power amplifiers and planar multipliers. Aluminium material was explored for HEB mixers because earlier measurements showed a large IF bandwidth (8 GHz) and also good sensitivity but with the disadvantage of requiring sub-kelvin cooling. These two developments were eclipsed so early in the project that they are not shown in the HIFI development and verification plan.

Table 8.1 Mission critical components and the associated risk factors considered by the HIFI consortium in their development and verification plan

	HIFI mission criticality		
Risk factor	Low	Medium	High
A: Feasible in theory		LOA 7	
B: Working lab model		LOA 1–6	SIS mixers
		HEB mixers (6,7)	Bands 1–5
C: Based on existing flight design		HRS-F	IF amps
			LOU
D: Extrapolated from existing flight design		HRS-S	FPU, WBS
			ICU, FCU, LCU
E: Proven design			

In the same plan, the team addressed the main risks arising from the various components under development. From this, a matrix emerged where mission criticality was displayed versus risk level (factor) arising from development status of each unit at the time of the start. This matrix is shown in Table 8.1. There was no item in the upper right corner (high criticality/feasible in theory) box. The local oscillator assembly (LOA) with the 2.7 THz multipliers was in the next lower level together with the SIS mixers used in laboratories and at ground-based observatories. The SIS devices, very critical for the HIFI mission, were not demonstrated in the field for frequencies beyond 500 GHz. The matrix describes the maturity of a technology in only five levels in the vertical scale. These levels can roughly be compared with the technology readiness levels (TRL, see Chap. 10) 1 to 9 as defined by ESTEC. Level 1 corresponds to basic technology research ('basic principles observed and reported') and level 9 to matured technology ('Actual system flight proven through successful mission operations'). From Table 8.1, we can conclude that HIFI was only for a limited part of its key technologies at a level that would have been acceptable for ESA today (levels 5 to 6[9] in the Cosmic Vision programme 2015–2025). This divergence from TRL probably explains the major innovations achieved during the development of HIFI.

8.3 Developments on Mixers and LOs (1970–1997)

By the early 1970s, optical and radio heterodyne principles were well understood. However, applications in the infrared (IR) to millimetre (mm) astronomy were limited. In the IR domain, there were only laboratory experiments and a very few spot frequency astronomical observations. In the millimetre spectral domain, several observatories operated with millimetre receivers, mainly around frequencies at

[9]Component and/or breadboard critical function verification in relevant environment (level 5) to model demonstrating the critical functions of the element in a relevant environment (level 6)

hundred gigahertz (GHz or 1 billion hertz). With the detection of interstellar molecules, like carbon monoxide (CO) in 1970, an interstellar molecule hunt was started that drove developments of components for receivers to higher frequencies and to lower noise. Considering the status of the technology at that time, research in all key subsystems (mixers, LOs, preamplifiers, back-end spectrometers) was essential to further astrochemistry and to allow the unveiling of the molecular universe. During the course of the following 25 years, heterodyne technology progressed step by step, ultimately resulting in near-quantum noise limited (theoretical limit) receivers and space qualified instrumentation.

In the early 1970s when these developments were started, the available LOs that could be used in astronomical receivers were klystrons, carcinotrons and Gunn diodes. Klystrons, invented in 1937, are devices where a linear electron beam in a small vacuum tube interacts with a resonance cavity, thus generating radio waves. Frequency selection is achieved by tuning the voltage and the resonant cavity. Klystrons can work up to about 200 GHz, but in astronomy, they are mainly used at 100 GHz. Most of the klystron development took place during World War II. Another type of radio-wave generator that can work at much higher frequencies is the backward wave oscillator (BWO). Here the electron beam, also in a vacuum tube, interacts with a periodic resonant metallic structure. It sustains the oscillations by propagating a travelling wave backwards against the electron beam direction. The radio power output (several 10 s of milliwatt) is located next to the electron gun. Its tuning is electrical, and the range can be much larger (15%) than in a klystron (5%). It was invented in 1951, and use for heterodyne receivers started in 1975 (de Graauw et al. 1978). Both types of vacuum tube oscillators are electrically fragile, their electron source (cathode) has a limited lifetime and they operate at high voltages. The voltage range for the BWO is up to 7 kV, and it needs a strong bulky magnet for keeping its electron beam well focused. Carcinotrons[10] have been shown to work up to 600 GHz. The only semiconductor suitable as LO at that time is the Gunn oscillator. The Gunn diode has a negative resistance in its *I-V* characteristic curve where an increased voltage causes a decrease of current. Its operation is based upon the Gunn effect detected in 1962, where a constant voltage applied to a semiconductor above a certain threshold causes oscillation. The output powers are moderate (10 mW at 100 GHz).

As described in Sect. 8.1, the mixer is the other key device of a heterodyne receiver. Initially almost all mixers used at mm wave observatories were made with semiconductor Schottky diodes,[11] operating at room temperature. From 1975 to 1983, the noise temperatures of these components dropped an order of magnitude by improving the solid state GaAs materials used, optimising the device layout, the

---

[10]Carcinotron is the brand name for the BWOs produced by Thomson-CSF in France. Russia is the other main supplier of BWOs with similar characteristics as Carcinotrons except that the oscillation starts at lower voltages, but using them requires more skills. Their highest oscillation frequency reported to date is 1.5 THz.

[11]Schottky diode (named after German physicist Walter H. Schottky) is a semiconductor diode formed by the junction of a semiconductor with a metal.

RF and IF embedding in the mixer mounts and the optical configurations. In addition, cooling of the Schottky mixers and their preamplifier was started. For 100 GH, the noise temperature went from 1800 K down to 120 K and for 230 GHz from 6000 K down to 800 K. An additional advantage associated with the cooling of the mixers was the reduction of LO power needed to pump the mixer.

An alternative for the Schottky mixer became the so-called HEB mixer using the InSb semiconductor. Kinch and Rollin (1963) were the first to recognise its detection capability from absorption of photons by the electron gas in the semiconductor. The absorption results in a temperature variation of the electron resistance. Putley (1965) reported a InSb submm detector for physics laboratory experiments using this mechanism, while Arams et al. (1966) were the first to carry out mixing experiments using these devices. This mixer technology provided a low-noise mixer but with very small bandwidth (1 MHz), and for spectral observations, the LO had to be swept across the spectral region of interest. On the other hand, LO power levels required were low (<0.1 mW), therefore more easily available and enabling access to a much wider high-frequency range. In 1973, a first successful demonstration in astronomy of the InSb HEB mixer (Phillips and Jefferts 1973) took place resulting in the first detections of important submm lines (e.g. the CO (2–1) line). For more than a decade, the receiver was successfully operated between 200 and 600 GHz, and important first detections were made with the 92-cm telescope on-board of the Kuiper Airborne Observatory (KAO) and with other optical/infrared telescopes. There were not many followers however, probably because of the drawback of the limited IF bandwidth. Glenn White and co-workers at Queen Mary College (see, e.g. Richardson et al. 1985) and Van Vliet et al. (1982) at the European Space Agency Technology Centre (ESTEC) and the University of Utrecht made regular observations from UKIRT[12] and other optical telescopes at Mauna Kea, using InSb HEB mixers.

Another solution emerged from the photon-assisted tunnelling junctions discovered by Dayem and Martin (1962) at Bell Labs. In such an electronic device, made of two superconductors separated by an ultra-thin film of insulating material (SIS), the current can pass through the barrier in a process of quantum tunnelling. There are two components to the tunnelling current. The first is from the tunnelling of Cooper pairs, consisting of two electrons that are weakly bound, the so-called Josephson tunnelling. This supercurrent is described by the ac and dc Josephson relations. The second is the quasi-particle current, consisting of ordinary electrons. At zero bias and low temperature, this current is virtually zero. It arises when the energy from the bias voltage and photon energy exceeds twice the value of the superconducting gap Δ, the value where Cooper pairs are breaking. This photon-assisted tunnelling changes the current-voltage curve, creating a non-linearity that can be used for mixer operation. When one talks about SIS junction, one usually addresses the quasi-particle current operation. The Cooper pair current operation is usually referred to the Josephson mixer that also shows a non-linear behaviour. Somehow the discovery of the quasi-particle tunnelling got eclipsed by the

[12]UKIRT: United Kingdom Infrared Telescope

discovery of the Josephson junction that provided more fascinating physics. The much later (1975+) developments at Bell and at IBM Labs on superconducting electrodes for a superconducting computer resulted in fast superconducting tunnel junctions. The possible application for mm/submm mixers was recognised by Tom Phillips at Bell Labs and later also by Paul Richards at UC Berkeley. Around 1976–1977, Tom Phillips and collaborators started to carry out SIS mixer experiments. They used lead (Pb) alloys that become superconducting at liquid He temperatures. By 1977, Richards, who had studied Josephson mixers, switched to SIS experiments. The first mixing results by these two competing groups were published, back to back, in 1979 in Applied Physics Letters. Personal accounts of the development of SIS mixers by several of the main players can be found in the ASP conference series no. 417, 'Submillimeter Astrophysics and Technology', a symposium honouring Thomas G. Phillips (Lis et al. 2009).

Following the FIRST assessment study carried out in 1983, the Schottky diode mixers were kept as a prime solution for the higher frequencies (>1 THz). However, Schottky mixers require high LO power levels, especially at higher frequencies. This is illustrated in Fig. 8.3 from Erik Kollberg (Chalmers University of Technology, Göteborg, Sweden), a THz pioneer who has played an important role in the development of heterodyne technology for FIRST. During his talk on submm receivers at the ESO-IRAM-Onsala[13] workshop on submm/mm astronomy in June 1985, he addressed the expected improvement in noise temperature from the development of SIS and Schottky mixers together with the local oscillator power needs for various types of mixers. As a possible alternative to overcome the LO power needs at the higher frequencies, subharmonically pumped Schottky mixers were developed where the LO frequencies are at half the signal frequency. They work well up to about 300 GHz, be it with somewhat higher conversion losses and higher noise than the directly pumped Schottky mixer. Therefore, this mixer scheme was not widely applied in receivers for astronomy but was acceptable for submm/mm atmospheric spectroscopy where sensitivity is less crucial.

By 1985, FIRST had become one of the cornerstones of the ESA long-term plan Horizon 2000 (cf. Sect. 3.2). As such, FIRST became a fixed element of the ESA Science Programme thus guaranteeing a level of funding for the essential preparatory activities of the main critical aspects of the mission (cf. Sect. 2.2.5), a.o. heterodyne technology. National space agencies and other funding sources were more willing to support heterodyne developments with increased focus on the main needs for the mission. In addition, ESA's Technology Research Programme[14]

---

[13]Onsala rymdobservatorium or Onsala space observatory, located at 45 km south of Göteborg, is the Swedish national facility for radio astronomy and depends on Chalmers University of Technology.

[14]The main participants in the TRP programme in the 1990s were Chalmers University of Technology, Sweden; DEMIRM/Observatoire de Paris, IRAM-Grenoble, LETI-CEA Grenoble, France; Farran Technology Ltd, Ireland; KOSMA/University of Cologne, Radiometer Physics GmbH (RPG), Germany; SRON/University of Groningen and University of Delft, the Netherlands; Technical University of Denmark; University of Cambridge, UK, with contributions from IREE and MSPU, Moscow, Russia.

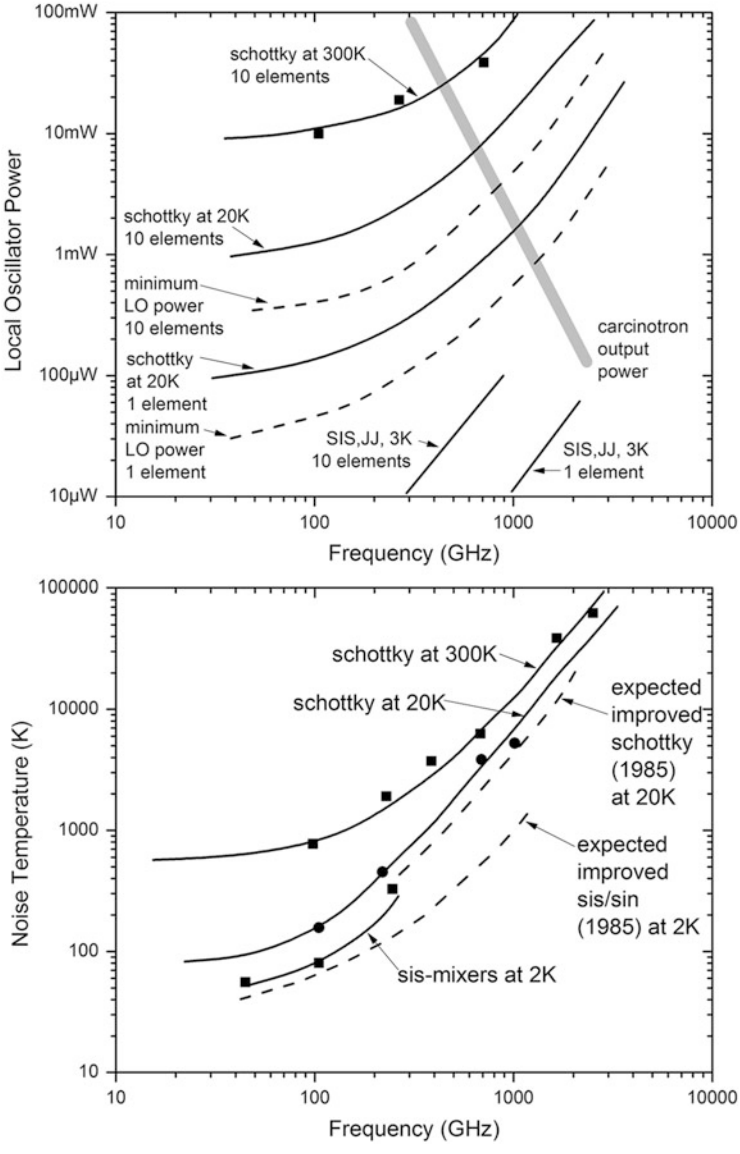


Fig. 8.3 Noise temperatures for Schottky and SIS mixers with present and expected performances as foreseen in 1985. The *upper figure* shows the needed LO power levels for Schottky and SIS mixers at various operating temperatures. © Erik Kollberg, Chalmers University of Technology

(TRP) started to devote considerable funds to this area and played a very important catalysing role in the preparation of the FIRST/*Herschel* mission (Chap. 10).

Insert: Focus on ESA TRP

In the period 1975–1999, ESA's Technology Research Programme was aimed at establishing a strong European technology basis for a lead space mission. In the early years (1975–1984), the submm instrumentation programme was focused on developing LO sources with sufficient output power to pump Schottky barrier diode mixers. The Carcinotron R&D effort was concentrated on improving almost all aspects of the unit (Kantorowicz and Palluel 1979) and included space qualification by addressing housing, shielding and last but not least the use of switched high-voltage power supplies. The mixer effort was mostly dedicated to the GaAs Schottky diode. ESA wanted to secure a European diode provider together with other European funding agencies and started to support the European diode laboratory in Cork, Ireland. At the end, the main source for high-frequency Schottky diodes was Bob Mattauch from the University of Virginia in Charlottesville who became the main worldwide supplier of Schottky diodes for mixers, including European receiver developers. When FIRST was accepted as one of the cornerstones of ESA's Horizon 2000 programme, the TRP needed to be re-evaluated, and more funding would be needed. A development plan was defined that focused on Nb and NbN SIS mixers, cooled Schottky receivers and GaAs multipliers. During 1990–1994, it was decided that the ultimate operating frequencies for SIS mixers should reach 1 THz, to be demonstrated in steps with 500 GHz in 1992, 750 GHz in 1993 and 1000 GHz in 1994. The programme suffered from some delays, but by 1997, the technology was demonstrated and did meet the requirements for the *Herschel* heterodyne instrument (HET). In 1995, it was realised that the TRP programme should be complemented with a HEB mixer development in order to have an alternative for the Schottky diode mixer at frequencies above 1.2 THz. The success of this HEB mixer development enabled the community to propose that the operating frequency range for HIFI should be extended above 1.2 THz. Indeed, no Schottky mixers were going to be flown, and no Carcinotrons would be needed.

As described above, when moving to submm wavelengths, the alternative to Schottky mixers is SIS or HEB mixers. The initial SIS devices available were based on lead (Pb) alloys: Pb, PbBi and PbBiAu with a native oxide as the insulating layer, usually formed on thermally evaporated alloys. These junctions were not chemically stable at room temperature and had a limited lifetime. A much more stable superconductor, niobium (Nb), has the disadvantage that its oxides are either semiconducting or metallic, making a poor tunnel junction. Recognising this problem, Gurvitch et al. (1983), at Bell Labs, involved in the development of superconducting structures for computer applications, developed Josephson tunnel junctions based on Nb/Al-oxide/Nb and Nb/Al-oxide-Al/Nb structures. They put on fresh Nb a thin layer of aluminium (Al), which was oxidised. This turned out to make a high-quality tunnel barrier. The so-called proximity effect took care that the Nb superconducting qualities were induced in the Al. The junctions were fabricated using standard photolithography, but the deposition and etching processes required precisely tuned apparatus and operation. Duplicating the recipes to get a very sharp diode *I-V* kink appeared to be extremely difficult. Nevertheless, this technique fulfilled the needs for SIS receivers in astronomy. The first high-quality Nb devices for astronomy were made by Inatani (1985) at Nobeyama Radio Observatory (1985).

Mixer Mount: Quasi-optical or Submicron Precision Engineering?
Although understanding the mixing process was key to progress in the detection process, the practical implementation of micron-sized structures with optimum RF and IF characteristics was equally crucial. Figure 8.4 gives examples of such structures, located inside a so-called mixer mount. The dimensions of the structures scale with wavelength and are thus smaller for the shorter wavelength/higher frequencies. The shorter wavelengths and the associated smaller dimensions stimulated the development of quasi-optical (QO) techniques for mixer mounts, in which the RF and LO signals were both delivered to the mixer through free-space optics rather than via waveguides. In 1982, it was believed that the QO technique would become superior above 300 GHz. However, in subsequent years, the waveguide manufacturing technologies became more and more sophisticated. By 2000, when HIFI selected its technologies for the various mixer bands, the transition from waveguide to QO was at around 1 THz.

It has since been extended up to around 2 THz. Although SIS devices showed a large promise, at that time (in 1984), it was still questionable whether they still would be suitable for operation in space and in particular at frequencies above the superconducting gap frequency (~700 GHz for Nb).

When moving to higher-frequency characterising photons with energy above twice the band gap of niobium, it becomes impossible to create sensitive tunnel junctions. Another mixer technology was needed. It was in the late 1980s that a hot electron effect, as it was presented earlier within the InSb semiconductor, was observed in superconductors. The first report was by E.M. Gershenzon and his Moscow group (Gershenzon et al. 1990a, b; 1991) including Gregory Gol'tsman (Gol'tsman et al. 1991). Since that time, Gol'tsman and his group have played a crucial role in the HEB mixer development by providing superior layers to all HEB mixer groups. HEB mixers have become so attractive because their response to radiation is truly bolometric, therefore frequency independent. They can be used from radio to infrared wavelengths and do not require much LO power. The

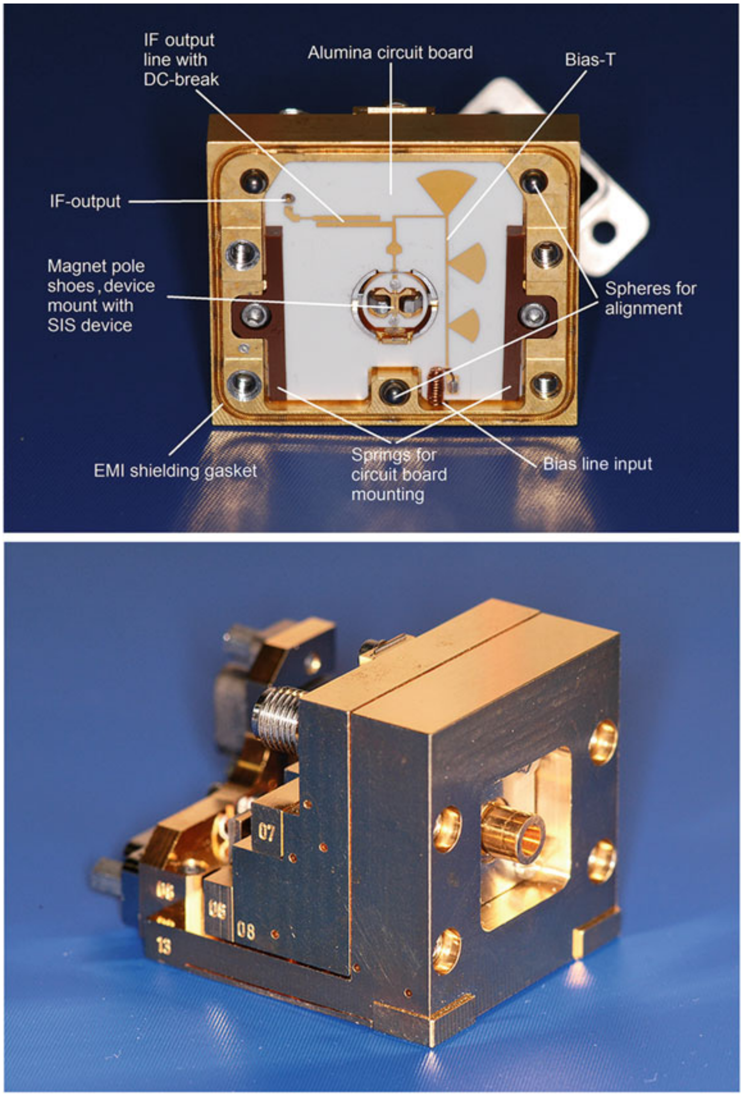


Fig. 8.4 A HIFI band 3.4 mixer. The *upper panel* shows the SIS mixer substrate. The *lower panel* shows the mixer mount with a horn looking to the right. © SRON

superconducting HEB mixers can achieve three orders of magnitude larger bandwidth than the bulk InSb crystals although this is not quite enough to meet the demands of today's receivers, and therefore efforts to widen the bandwidth beyond 5 GHz are still continuing. The acquired performance of the HEBs by 1995 made them workhorse mixer for frequencies above 1.2 THz.

In 1993, when FIRST was selected as Cornerstone 4, since little progress in high-frequency Schottky mixers had been achieved, and because the superconducting SIS and HEB mixers started to perform well, there was a general confidence that no Schottky mixers were needed on FIRST. This led to a huge relaxation of the LO power output levels and implied that cryo-cooling should be provided by the FIRST satellite for the heterodyne instrument with SIS and HEB mixers. In 1984, the highest-frequency local oscillator sources consisted of frequency-multiplied output of klystrons and carcinotrons. The multipliers consisted of GaAs diodes mounted in the wall of a waveguide where the radiation was intercepted by a hair-thin metallic wire (whisker) conducting the radiation onto the diode assisted by two mechanical back-short plungers. Despite several attempts, it appeared that the combination of the vacuum tube klystron and multipliers could not provide the necessary power levels to pump a 1 THz Schottky mixer. However, the output would be adequate for SIS and HEB mixers at that frequency. The only non-desirable alternative left was to fly the Carcinotron with its heavy magnet and high-voltage supply. Somewhat later, improved Gunn solid-state oscillators became available for the 100 GHz range, with somewhat lower output power levels than the klystrons but with the major advantages of much lower operating voltages (a few volts rather than keV), higher stability and much longer lifetime. An alternative scheme for generating LO power was realised with FIR pumped gas lasers providing spot frequency continuous wave (CW) output powers, but the bulky equipment with high-voltage supply for the pumping CO_2 laser, the need for hands-on tuning by an expert and the limited number of frequencies coinciding with interesting astronomical lines made this LO type one that never could and would be adapted for use in spacecraft (Buhl et al. 1981). However, it allowed characterisation of Schottky mixers at 890 GHz and above and determined the LO pumping power (Buhl et al. 1981). Whiskered contacts are also difficult to space qualify. Triggered by the realisation of planar diodes for lower frequencies ($<$300 GHz), the goal for an instrument like HIFI was to operate with an all-solid-state, completely electronically tuneable LO chain with a wide tuning range and sufficient LO power to pump SIS and HEB mixers. With planar diodes in an integrated structure, it would be easier to make the micron-sized structures to optimise the design for efficient operation of the doublers and triplers over larger bandwidths ($>$10%).

The solution to have high-frequency planar diodes emerged in the 1990s and came from several directions. A number of complex GaAs processes had been developed for the processing of planar devices. To carry out these processes successfully in a direct interactive way, a microelectronic development facility would be needed. JPL had opened a new laboratory, the Micro Devices Lab (MDL), which became a centre of excellence in space microelectronics. It got into full operation by 1990. A second positive evolution was the development of high-frequency modelling with electromagnetic simulation tools. These tools, enabled by high power state of the art computers, were capable of predicting the frequency response of microwave circuits with a very high degree of fidelity. The sophistication of the programmes grew with the years, and by the time planar multipliers needed to be designed, there was sufficient know-how to design and model new high-frequency circuits. An example is SuperMix, a flexible software library for high-frequency circuit simulation, including SIS mixers and superconducting components (Ward et al. 1999). However, this was not sufficient. For multiplication to higher frequencies (> 600 GHz), the delivered output power from planar diode devices at 100 GHz, like a Gunn oscillator, was far from sufficient. Moreover, Gunn oscillators had a very limited tuning range. When new power amplifiers were developed for 100 GHz, using monolithic microwave integrated circuit (MMIC) techniques, Todd Gaier (JPL) immediately recognised the opportunities for THz heterodyne local oscillators. These amplifiers worked very well with high output powers, while the low-frequency input could come from a set of frequency multipliers sourced by an X-band[15] oscillator. The 100 GHz amplifiers were sufficiently robust to cope with the required bias dissipation, while the microwave output was sufficient to drive a three-step multiplier chain, up to 2 THz.

During this period, several other key technology developments were undertaken. Preamplifier developments motivated by the needs of telecommunications and military applications allowed the mm/submm astronomy community to gain access to high-electron mobility transistors (HEMTs), which could be used for high-frequency, low-noise, cryogenic preamplifiers. Several excellent batches were developed for radio astronomy, some of which were used for millimetre-wave receivers. Back-end spectrometer developments were gradual in this area. Acousto-optical spectrometers (AOS) were developed for ground-based observatories. HIFI relied on the AOS expertise from the Submillimeter Wave Astronomy Satellite (SWAS) and *Odin* satellites.

8.4 From the Heterodyne Model Payload Instrument (HET) to HIFI

In January 1997, the *Herschel* Payload Working Group (PWG) made an assessment of the technology readiness. The heterodyne model payload instrument, known as HET, was to operate till 1.2 THz. A critical item list was made to identify which components needed more development for HET. The PWG concluded that no completely new technology is needed to be developed. However, further development was needed on tuning and control for the LO and for the IF amplifiers and on reduction of power dissipation in the transistors. They assumed that SIS mixers would work well up to 1.5 THz. In 1997, the demonstrated performance of an SIS mixer at 1.1 THz was considered to be good enough (T_{noise}= 1000 K), and they envisaged that it could be extended to 1.5 THz. They also commented that a 1THz LO had already been made with a 100 GHz Gunn oscillator followed by a cascade of whisker point contact frequency multiplier and had provided sufficient output power to pump an SIS mixer. Such a LO chain would be extendable to 1.5 THz.

[15]X-band: radio frequency band between 8 and 12 GHz

Their conclusions were based on a report to ESA TRP, made in 1999 by the company RPG that commercialised the developments by Peter Zimmermann. He reported a 2 μW of LO power at 1.39 THz that was demonstrated with an LO chain consisting of a solid-state Gunn oscillator (77.3 GHz) followed by three stages of multiplication (doubler-tripler-tripler) with an output at 1.39 THz. These multipliers however used whisker-contacted diodes and sliding back-short tuners in the (small) waveguides to optimise efficiency and extend the tuning range. The PWG ignored the limited tuning range of the available Gunn oscillators. With the Gunn multiplier approach, the multiplication scheme of HIFI would have had a total of 48 sliding back-short tuners. The unreliable whisker contacts for space applications had to be replaced by all means. Such development became one of the top priorities for HIFI, and RPG, as subcontractor, played an important role together with JPL where a rigorous and ambitious development plan was started in 1998 in collaboration with the group at the Observatoire de Paris[16] that designed several THz frequency multipliers. The aim was to have multiplier chains up to 2.7 THz based on the new technologies. Although the extra effort by the HIFI consortium to prepare the instrument and its components for space flight was mainly due to its higher-frequency ranges as compared to the HET specifications, the claim pronounced by the FIRST PWG in 1997 that 'no new technology had to be developed and that therefore the development risk was low' was still quite an understatement.

By the time the proposal had to be submitted to ESA in 1998 and based on technology developments under the TRP programme and in the USA, it was clear that the HIFI instrument could be built in principle. However, much work was still needed to fulfil the science objectives and to meet all interface requirements with the spacecraft to obtain full space qualification. The multitude of development items spread over, so many institutes required strong management to direct and control activities in a transparent way. For that purpose, the 'HIFI Instrument Development and Verification Plan' (IDVP) was written. It defined in full the intended technical implementation, including engineering, manufacturing, assembly integration and testing and engineering aspects of product assurance and verification, and included model philosophy, timeline, to realise the specified technical product. The IDVP, the *HIFI Project Management Plan*, the *HIFI Science Implementation Plan* and the *HIFI Product Assurance Plan* were the top four documents defining the baseline for the HIFI Project. With so many parts and components to be used in space for the first time and so many developments to be done, a strict focused approach was needed. From 1998 till flight instrument delivery, the remaining technology gap was overcome, and the HIFI instrument reached almost completely its promised capabilities.

The HIFI consortium was unanimous including higher-frequency bands in their instrument proposal. These science requirements, as given in the *HIFI Science User Requirement Document*, defined the frequency coverage from 480 to 1250 GHz, 1410 to 1910 GHz and 2400 to 2700 GHz, frequency resolutions of 1 MHz and

[16]Initially with DEMIRM that became LERMA

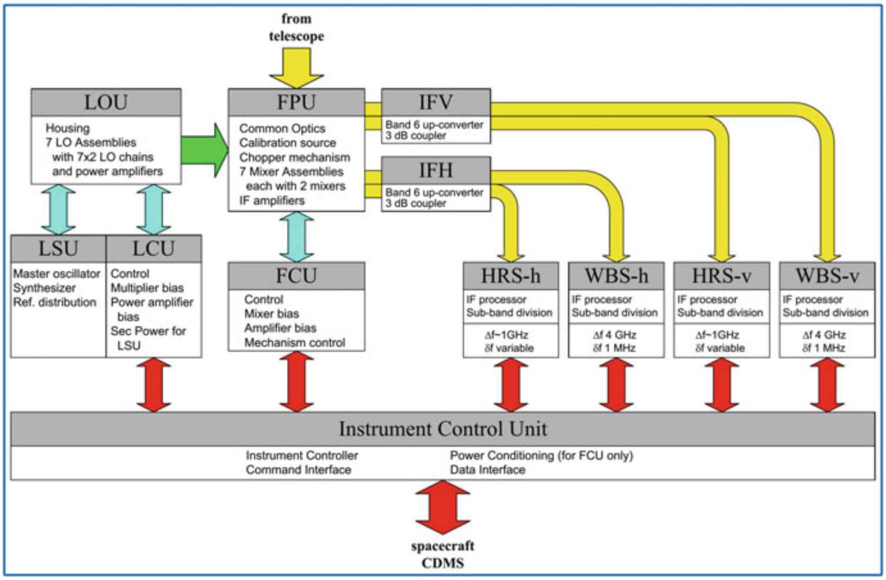


Fig. 8.5 HIFI block diagram with all subsystems and units. Acronyms are spelled out in the text. Basically the block diagram represents a multitude of the one shown in Fig. 8.1. © SRON

125 kHz and instantaneous IF bandwidth of 4 GHz. Below 1250 GHz, both polarisations should be measured. Sensitivities three to four times the quantum noise limit should be reached in all bands. The choice for mixer type was SIS, for frequencies up to 1250 GHz, and HEB, above this frequency. As a consequence, special attention would need to be given for the mixer sensitivity, dynamic range, gain stability and radiometric calibration and for the case of the LO, to frequency resolution and accuracy. In addition, the overall observing efficiency would need attention. These requirements, derived from the main science goals, plus the space qualification, drove the heterodyne hardware design and development.

The HIFI instrument and its calibration as flown are described in de Graauw et al. (2010) and Roelfsema et al. (2012). Here only the most important aspects are reproduced. A block diagram of the system is shown in Fig. 8.5. The HIFI focal plane unit (FPU), located on the *Herschel* optical table in the cryostat, contained seven mixer assemblies covering seven frequency bands. Each mixer assembly contained two mixers, for horizontal and vertical polarisations, whose beams coincided. Each mixer assembly was looking at a slightly different part of the sky. By repointing the satellite, the beam from the source in the sky could be directed to one of the seven frequency bands. Only one mixer assembly (frequency band) was operating at a time with one beam per band. The beams went first through some common optics that included insertion of beams from the calibration units and a beam switcher/chopper. A local oscillator unit (LOU), with seven LO assemblies, each with two LO chains, was located outside the cryostat and sent seven adjacent LO beams through cryostat windows into the respective frequency mixer assemblies in the FPU. Each of the LO assemblies, with dedicated chains of power amplifiers and frequency multipliers, received corresponding monochromatic signals from a local oscillator source unit (LSU). The FPU was controlled by a dedicated processor with control circuits, the so-called focal plane control unit (2) (FCU), while the warm local oscillator control unit (LCU) controlled and powered the local oscillator source unit (LSU) and the local oscillator unit (LOU). All together they covered the frequency ranges of interest, lying between 480 and 1250 GHz and 1410 and 1910 GHz. Note that band 7 in the proposal covered 2400–2700 GHz. After 3 years in the programme, this band was given up due to limited resources. Band 6, covering 1410–1910 GHz, was split into two sub-bands (6a and 6b, later renamed band 6 and band 7). All mixer IF signals were amplified by cryogenic HEMT[17] amplifiers. Afterwards, the IF signals were split into two, one going to the wideband spectrometer (WBS) which employed an AOS technique and the other one going to a digital autocorrelation spectrometer, the high-resolution spectrometer (HRS). The digital outputs of the two spectrometers were sent to the Instrument Control Unit (4) (ICU), which served as the overall instrument main controller and as the interface with the satellite for power supply, instrument control and data archive.

Besides dedicated developments, *Herschel*-HIFI benefitted from two previous satellite projects that used heterodyne receivers: SWAS (Melnick et al. 2000) and *Odin* (Hjalmarson et al. 2004). These small missions, not being cryogenic, had to use state-of-the-art ambient temperature Schottky diode mixers and Gunn oscillators for their receivers and AOS, similar to the one later used in HIFI. Although the technology heritage from these missions applicable in HIFI was limited, the experience gained by the SWAS and *Odin* teams in operating the AOS and its calibration was valuable.

However, although the complete instrument was presented in the HET/HIFI proposal, many crucial subsystems were not fully developed. In the next two sections, we restrict the description of HIFI mixers and LO developments as these were the most challenging and the most important to achieve the scientific capabilities of the instrument.

8.5 Ultimate Technological Developments for HIFI

8.5.1 Mixer Developments

At the time of the HIFI proposal to ESA at the end of 1998, most ground-based observatories had moved to SIS mixers for their heterodyne instruments, mostly at frequencies around 115, 230 and 345 GHz (corresponding to CO lines accessible in atmospheric windows). The mixers to be developed for HIFI and their space

[17]HEMT is a high-electron mobility transistor used for amplifying current.

readiness would of course rely on earlier research, among others in the TRP programme, but would still require a huge effort to be realised. In particular, above 1 THz, there was very little experience in manufacturing and testing devices. We summarise first the requirements for the HIFI mixers as written in the HIFI Instrument Development and Verification Plan (IDVP): for frequency bands 1–4 (see Table 8.2), the baseline mixer designs were a waveguide structure with a corrugated horn. For bands 5–7, quasi-optical mixers would be used. For these, the baseline antenna was a double slot on a hyper-hemispherical lens. It was required to use mixers without mechanical tuners. This was very challenging in view of the large instantaneous bandwidth of the individual mixers (for all bands 1–4, 160 GHz), especially for the lower-frequency bands. All SIS mixer units would have de-flux heaters and superconducting electromagnets to allow a fast and reproducible removal of trapped flux and noise due to Josephson effects. Three to four times the quantum noise limit was the baseline sensitivity for bands 1–4 and a goal for bands 5–7. For the IF, bandwidths as wide as 4–8 GHz were required. Most (laboratory) receivers at that time operated with a 1–1.5 GHz bandwidth centred at 1.5 GHz; a 4–8 GHz bandwidth had only once been demonstrated. These requirements demanded further development of SIS junctions themselves and their embedding circuitry technology to increase RF and IF bandwidth, reduce RF losses, increase conversion gain and improve the overall receiver sensitivity. With the SIS mixer bands covering such a wide frequency range for each band, specialised design concepts were needed for each of the bands. The technologies and challenges are summarised in Table 8.2 together with the names of the responsible institutes for each band.

The situation for the HEB mixers is a similar one. HEBs are in principle frequency independent, and in 1998, there were two flavours: phonon-cooled and diffusion-cooled devices, where cooling concerned the photon-heated electron gas. Both were thoroughly investigated, and after careful testing, the phonon-cooled HEBs were found to give the best and most reliable results, albeit with a somewhat reduced IF bandwidth (2.5 GHz). While SIS mixers were used in many ground-based telescopes, this was not the case for HEBs with very restricted experience for operation above 1 THz. During the *Herschel* development phase and environmental testing for space qualification, it turned out that the HEBs with submicron dimensions were not so rugged and stable as expected. Accelerated lifetime tests were needed to demonstrate that the invented remedial measures worked. In addition, for the operational settings, there was some lack of experience. More laboratory testing after *Herschel*'s launch showed also that the sensitivity could have been improved by 30% applying a bias feedback to the signal output, while Allan variance times could have improved by orders of magnitude when stabilising the LO in a feedback loop with the mixer bias.

Regular meetings involving all mixer teams showed an open exchange of experiences and ideas. Reaching the baseline sensitivity turned out to be a strong motivator for the mixer teams and a stimulus as effective as direct competition. At the end, the baseline sensitivity was reached or approached in almost all mixer bands (Roelfsema et al. 2012) and in band 1 also the goal values. Meanwhile all

Table 8.2 Overview of HIFI mixer bands (B), their RF and IF frequency ranges, the responsible institutes and the technologies used for mixer mounts, mixer devices including the various wiring layer technologies and associated technical challenges

B	Institute	RF freq. (GHz)	IF freq. (GHz)	Technical challenge	Mixer mount	SIS/HEB device; wiring layer
1	LERMA Obs. Paris	480–640	4–8	RF and IF bandwidth	Waveguide with horn	Nb/Al2O3/Nb Nb/Nb
2	KOSMA Uni. Köln	640–800	4–8	RF and IF bandwidth Noise temp.	Waveguide with horn	NbTiN/Al2O3/Nb NbTiN/Nb
3	SRON	800–960	4–8	Noise temp. Layer Tech.	Waveguide with horn	NbTiN/Al2O3/Nb NbTiN/Al
4	SRON	960–1120	4–8	Noise temp. Layer Tech.	Waveguide with horn	NbTiN/Al2O3/Nb NbTiN/Al
5	JPL/Caltech	1120–1250	4–8	Noise temp. Layer Tech. Device Fab.	Si lens, planar twin slot	NbTiN/AlN/NbTiN NbTiN/Al
6	Chalmers with JPL	1410–1703	2.5–6	Noise temp. Device Fab. IF bandwidth	Si lens, planar twin slot	HEB NbN phonon cooled
7	Chalmers with JPL	1703–1910	2.5–6	Noise temp. Device Fab. IF bandwidth	Si lens, Planar twin slot	HEB NbN phonon cooled

space qualification tests had been carried out, and all mixer tuning and optimum settings could be performed electronically, including the biasing and removal of flux trapped in the devices.

The results of the receiver development during the HIFI design and construction phase are illustrated by Fig. 8.6. The dark green line presents the state of the art at the time of the proposal (SOAP). The lighter green line gives the baseline values that are expected to be reached for the flight model. In almost all cases for SIS, the baseline values were reached, except for band 5 due to a miscalculated polarising effect of the substrate. The HEB performances were slightly less sensitive than the baseline values.

8.5.2 *Local Oscillator Developments*

By 1998, the main ingredients to develop the HIFI LO had just become available and were in place. However, there were no LO chains that could demonstrably operate at the extended HIFI frequency range. The LOs for the high-frequency bands 6 and 7 presented a huge challenge. The baseline adopted in the HIFI

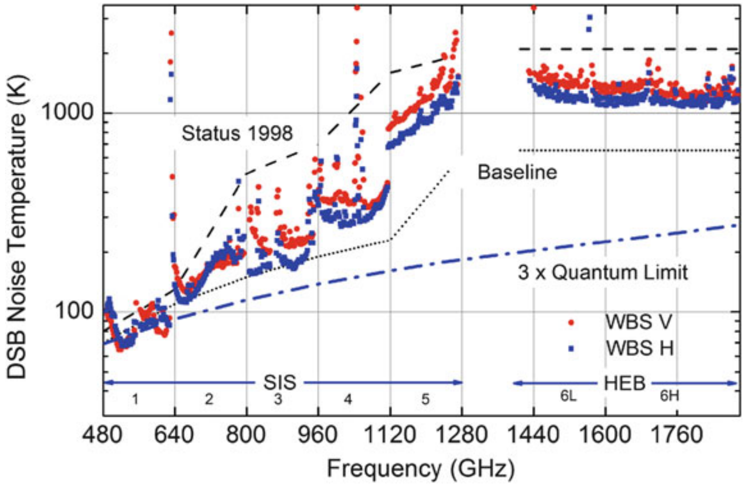


Fig. 8.6 The measured HIFI receiver noise temperatures. Shown are the 1998 noise values and the expected baseline. The *straight line* represents three times the quantum limit. © SRON: G. de Lange

proposal for these bands was to use a laser photo mixer system. This was a very new development which had been demonstrated in the IR. A description and the status in 1997 were presented by Verghese et al. (1997). While sufficient tenability and stability with a narrow line width had been demonstrated, one of the main goals of the development here was to increase the output power level. However, by the year 2000, the laser mixing solution had to be abandoned. The parallel development for the 100 GHz frequency source, the mentioned MMIC[18] power amplifiers (Huang et al. 1997), was very successful. In addition, the use of GaAs MMIC technology adopted for planar multiplier circuits was likewise successful. The devices provided not only wide bandwidths but also good reproducibility, ease of mounting and high reliability. This way, each LO chain became a chain starting with a 10 GHz oscillator, followed by multipliers, a power amplifier around 100 GHz and a series of higher-frequency multipliers (Bruston et al. 1997).

The HIFI development and verification plan, finalised in 2000, described comprehensively all foreseen steps, including a timeline, to build all the HIFI LO units (LOA, LCU, LSU). The development of planar technology for the power amplifiers and frequency doublers/triplers was fully described and rigorously pursued. The outcome had to be a set of fully electronically tuneable LO chains operating over the entire HIFI frequency range with sufficient power to pump the SIS and HEB mixers and with high spectral purity. Apart from power and bandwidth requirements, there were a few other main requirements that drove the designs, notably, (a) allow easy, accurate and reproducible assembly, (b) robustness for flight

[18]MMIC stands for monolithic microwave integrated circuit

Table 8.3 Multiplier scheme and resulting LO frequency bands (in bold) as used in the HIFI LO assemblies

Power amplifier/ multiplier scheme	71–79 GHz	80–92 GHz	88–99 GHz	92–106 GHz
×2	142–158	160–184	176–198	184–212
×2 ×3	284–316	320–368	352–396	368–424
×2 ×3	426–474	**1a: 488–546**	528–594	**1b: 560–633**
×2 ×2 ×2	568–632	**2a: 647–710**	**2b: 724–793**	**3a: 807–848**
×2 ×2 ×3	**3b: 862–953**	**4a: 967–1042**	**4b: 1056–1113**	**5a/b: 1127–1242**
×2 ×2 ×2 ×2			**6a: 1418–1596**	**6b2: 1472–1696**
×2 ×3 ×3			**6b1: 1584–1782**	**7a/b: 1704–1906**

The input to the power amplifiers came from a synthesiser followed by triplers and power amplifiers. Note that only four types of power amplifiers were needed for all 14 LO chains

qualification and (c) low-temperature operation (<100 K) where the output levels increased considerably. A detailed scheme was designed to develop all the power amplifiers and the multipliers for the 14 LO chains. Table 8.3 shows the realised multiplier chains.

Although Table 8.3 suggests that several multipliers could be used for different chains, each multiplier element in a chain was actually an individual unit. The performance depended strongly on its matching to preceding and succeeding components. If needed, impedance matching devices (isolators) were added between these. This was essential to achieve the required bandwidth. As a rule of thumb for conversion efficiencies of each multiplier component, the first stage needed to be at 25%, the second at 15%, the third at 5% and the fourth at 1–2%. These values were seen at frequencies below 1 THz. For higher frequencies, the efficiencies could be a factor 2–4 lower. It must also be noted that in 1996, the experience in cascading more than two multipliers was very scarce.

Figure 8.7 shows the achieved power output as a function of frequency (J. Pearson, private communication). In all cases, the output levels are well above the requirements. Actually, the requirements for the SIS mixers were too conservative (high), and in several bands, special measures had to be taken to attenuate the LO output power levels.

8.6 HIFI as Built

More than 20 years elapsed between the time of the proposal for the FIRST mission (1982) until the delivery of HIFI (2006) for integration into *Herschel*. In this period, the performance of the heterodyne components underwent numerous developments as described in the previous section. Sensitivity improved by two orders of magnitude through the adoption of superconducting devices, tuning became fully electronic and so on. These improvements enabled *Herschel* to address in an effective

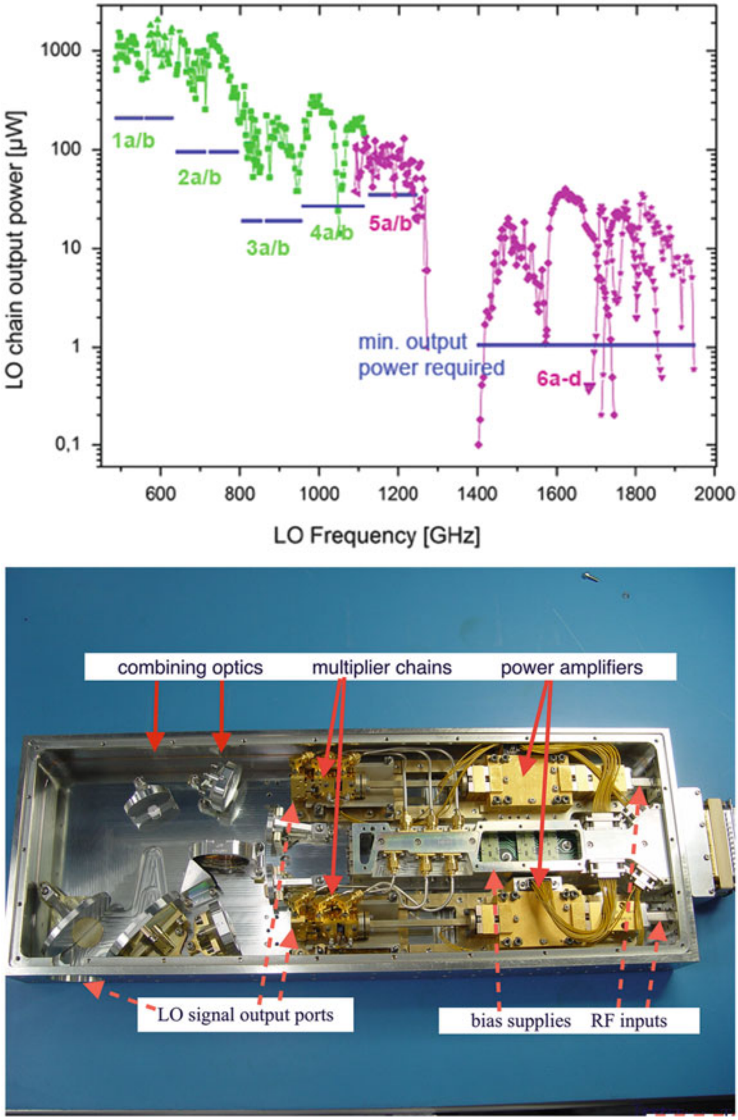
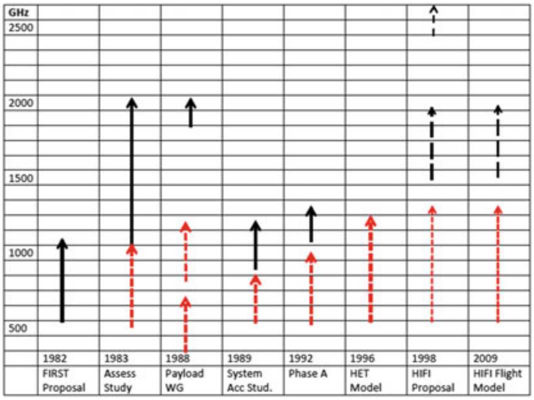


Fig. 8.7 *Top*: the generated output power levels in the LO chains of the flight model. The *horizontal lines* give the requirements for the HIFI mixer assemblies as given in the development plan. Note that band 6 a–d was later split into bands 6a,b and 7a,b for naming conventions © J. Pearson. *Bottom–top view* of the local oscillator assembly (LOA). © Th. Klein (MPIfR), private communication

Table 8.4 Frequency coverage for the heterodyne receivers in the FIRST/*Herschel* mission through the various stages of development, shown on the horizontal axis. The vertical axis gives the frequency range in GHz. The drawn lines represent Schottky mixers. The red dotted lines are for SIS mixers (<1200 GHz) and for striped for HEB mixers (>1200 GHz)



way the important scientific questions and helped keep its science programme unique and competitive (see Chap. 4). On the other hand, the requirements for the frequency coverage remained almost the same since the assessment study. Table 8.4 illustrates the frequency coverage in the various phases of the project, from the proposal and assessment study, via the various working group results in the late 1980s and early 1990s, the model payload definition, to the actual *Herschel* instrument proposal and the finally launched flight model.

It is remarkable that the envisaged frequency coverage at the time of the assessment study is very similar to what was implemented many years later in HIFI. For the assessment study, a team of receiver experts with a good understanding of the field projected a highest frequency up to 2 THz. This range was somewhat reduced with some continuous coverage dropped, but the overall plan was maintained by the payload working group in 1988. ESA cut the frequency extent in the subsequent model payloads of 1989, 1992 and 1996, adopting a maximum frequency of 1.2 THz. However, rapid developments in the mixer and multiplier areas eventually allowed the HIFI team to cover frequencies up to 2 THz as was originally foreseen.

HIFI was constructed, tested and delivered on a timescale compatible with the overall *Herschel* schedule. Its ten boxes and two smaller units are depicted in Fig. 8.8, following the block diagram of Fig. 8.5. The in-orbit performance was close to 'as designed'. The sensitivities were at the design baseline values or better. It provided unprecedented accuracy in frequency calibration; it reached its highest spectral resolution (demonstrated in Caselli et al. 2012) with a wide IF band (4 GHz in bands 1–5, 2.4 GHz in bands 6–7) and operated over all of the advertised RF coverage (480–1250 GHz and 1401–1810 GHz). The photometric calibration accuracy was everywhere within 10% and in many cases close to the goal of 3%.

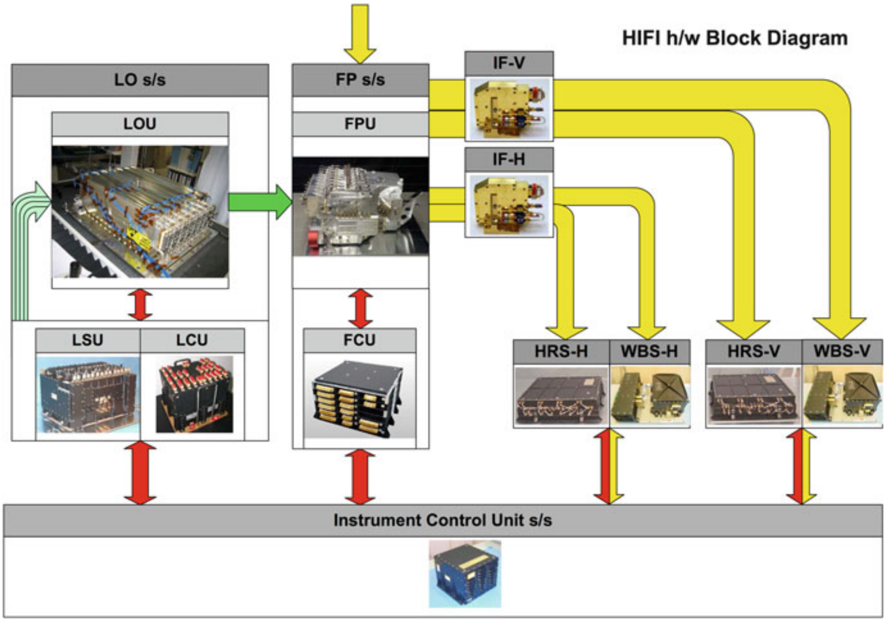


Fig. 8.8 Photographs of all the (sub-)units of the HIFI flight hardware are shown following the HIFI instrument block diagram as given in Fig. 8.5. © SRON

The details about the in-orbit performance are described in de Graauw et al. (2010) and Roelfsema et al. (2012) (Fig. 8.9).

However, there were also some issues. There were many spurious signals created in the LO chains. These could be readily identified. Also impure (i.e. non-monochromatic) LO outputs did exist in specific frequency regions. These could be cured by fine-tuning of the bias settings of the components in the LO chain. Electrical standing waves, due to impedance mismatches, caused quite some ripple, particularly in the bands 6 and 7. These could be eliminated in post-operation data processing. Optical standing waves were found to be unpredictable and had to be addressed empirically and to be cleaned in the data processing phase. Finally, for the first time ever, variations in sideband ratio could be determined with high accuracy, down to the one percent level, while imbalances as high as 10% were detected, sometimes over small RF ranges. All these effects could be accurately measured and compensated thanks to the stability and sensitivity of the HIFI instrument and the absence of an unstable atmosphere, which normally veils or smears these effects. Detailed results on the optical design and performance verification can be found in the PhD thesis by Willem Jellema (2015).

During operations, a most dramatic event for HIFI occurred after 214 days in orbit. On Monday 3 August 2009, the Mission Operation Centre (MOC) reported that *Herschel*-HIFI was found in an undocumented state. The currents of the HEB mixer of band 7 changed within one periodic House Keeping (HK) reading, from pumped to un-pumped, showing that there had been a sudden drop of LO power.

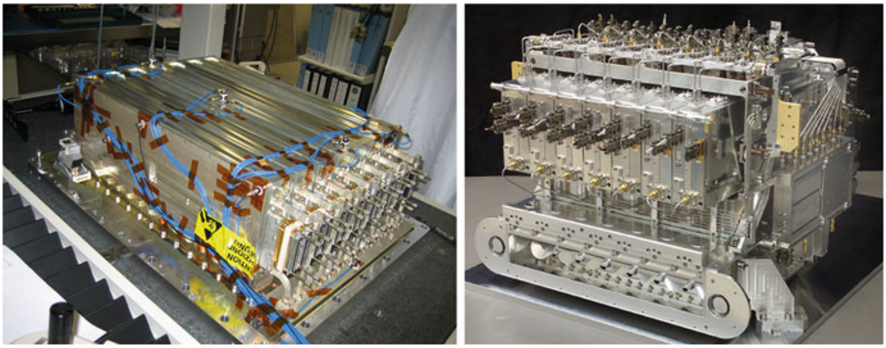


Fig. 8.9 At the *left*, a photograph of the Local Oscillator Unit (LOU) with the seven LO assemblies, each containing two multiplier chains. At the *right*, a photo of the Focal Plane Unit (FPU) with its seven mixer assemblies. At the lower end of the FPU, there are seven input openings for the seven beams from the LOU with an alignment flat at each side. The LOU is located at the outside of the *Herschel* cryostat, and its beams are entering the cryostat through seven optimised windows located opposite the FPU LO optics as shown. © MPIfR and SRON

HIFI also suddenly lost its communication with the local oscillator control unit (LCU). At the time of the incident, the instrument was operating in nominal mode in band 7, making a map at a frequency of 1893 GHz. It was necessary to switch off the HIFI instrument and start an investigation. But in order to understand better the status of the LCU and to find out what was the root cause of the failure, it was decided to restart HIFI and to read and check again the settings of all subsystems. At the same time, an urgent effort was started with the spare units in the laboratory to simulate the phenomena that had been observed. After several months of testing and analysis by dedicated experts from HIFI and ESA, the cause of the problem was understood, and changes in the software and operation procedures were implemented to restart HIFI and put its redundant electronic equipment in a full but safe operation mode. The investigations indicated that more than one issue had occurred that led to the mishap (see Orleanski et al. 2012; Jellema et al. 2010).

8.7 Conclusion

Throughout the HIFI project, it was possible to align the interests and coordinate the work of a large number of institutes (23) in a large number of countries (11). Everyone involved was working towards one common goal. Close to 500 people were involved in the project with contributions covering definition, design, development, software, calibrations, administration and science utilisation. The continuous presence of an unchanged scientific motivation of the team leaders ensured alignment of the consortium teams during the long implementation phase and turned out to be a key factor in the success in this complex project. HIFI was a science-driven project, while for the engineers, the incentive came from the challenge to build a unique high-tech instrument. All these factors did offset the inconveniences of a large consortium.

The long-term funding prospects provided by the various funding agencies in each country and by ESA played a crucial role to have the enabling technologies well developed and available in time for the mission, despite its low 'TRL' level at the start of the project. This preparatory phase also played an important role in community building, leading to an efficient and smooth consortium formation. Cooperation and competition were thus bundled into one project. The outcome was a world-class instrument for an ESA Cornerstone mission that produced a wealth of scientific results. It also generated dozens of trained people at the highest levels in engineering, instrument science and project management.

In all of this, communication and open knowledge dissemination played a crucial role. Joint activities such as the TRP programme, the yearly International Symposia on Space Terahertz Technology (ISSTT), science preparatory workshops and Key Programme symposia, all fostered collaboration among the HIFI aficionados. The few problems that occasionally arose between persons or groups generally showed lack of communication rather than conflicting or diverting opinions.

A final and important conclusion is that the TRL qualification is not a definitive discriminator for the go-ahead (or not) of a mission concept in particular when dealing with ambitious projects that require long lead time for preparation. HIFI has shown that in a relatively short time, elements with low TRL could be made ready for use in a complex space mission.

References

Arams CA, Peyton B, Sard E (1966) Millimeter mixing and detection in bulk InSb. Proc IEEE 54:308

Armstrong EH (1918) Notebook from Harold Miller Lewis. http://www.columbia.edu/cu/lweb/eresources/exhibitions/treasures/html/171.html

Baryshev AM (2005) Superconductor-insulator-superconductor THz mixer integrated with a superconducting flux-flow oscillator. Thesis University of Delft

Bruston J, Smith RP, Martin SC et al (1997) Fixed-tuned submillimeter waveguide multipliers using MMIC technology. In: Blundell R, Tong E (eds) Eighth international symposium on Space Terahertz Technology. Proceedings of a conference held in Cambridge, MA, 25–27 March 1997, p 198

Buhl D, Koepf GA, Chin G et al (1981) 433 micron laser heterodyne observations of galactic CO from Mauna Kea. Proc SPIE 280:108

Dayem AH, Martin RJ (1962) Quantum interaction of microwave radiation with tunneling between superconductors. Phys Rev Lett 8:246

de Graauw Th, Anderegg M, Fitton B, Bonnefoy R (1978) In: Gustincic JJ (ed) Third International conference submillimeter waves. Guilford University of Surrey

de Graauw Th, Helmich FP, Phillips TG et al (2010) The Herschel-Heterodyne Instrument for the Far-Infrared (HIFI). A&A 518:6

Fessenden RW (1901) Apparatus for signaling by electromagnetic waves. US patent US 706747. https://www.google.com/patents/US706747

Forrester TF, Gudmundsen RA, Johnson PO (1955) Photoelectric mixing of incoherent light. Phys Rev 99:1691

Gershenzon ME, Gershenzon GN, Gol'tsman GN et al (1990a) Electron-phonon interaction in ultrathin Nb films. Sov Phys JETP 70:505

Gershenzon EM, Gol'tsman GN, Gogidze IG, Gusev YP, Elantev AI, Karasik BS, Semenov AD (1990b) Millimeter and submillimeter range mixer based on electronic heating of superconducting films in the resistive state. Sov Phys Superconduct 3:1582

Gershenzon EM, Gol'tsman GN, Gousev YP, Elant'ev AI, Semenov AD (1991) Electromagnetic radiation mixer based on electron heating in resistive state of superconductive Nb and YBaCuO films. IEEE Trans Magnetics 27:1317–1320

Gol'tsman GN, Semenov AD, Gusev YP, Zorin MA, Gogidze IG, Gershenzon EM, Lang PT, Knott WJ, Renk KF (1991) Sensitive picosecond NbN detector for radiation from millimeter wavelengths to visible light. Superconduct Sci Technol 4:453

Gurvitch M, Washington MA, Huggins HA (1983) High quality refractory Josephson tunnel junctions utilizing thin aluminum layers. Appl Phys Lett 42:472

Hjalmarson A et al (2004) Adv Space Res 34(3):504–510

Huang PP, Huang T-W, Wang H et al (1997) IEEE transactions on microwave theory and techniques. 45(12):2418–2421

Inatani J (1985) US-Japan workshop on Josephson junction electronics. Hawaii, USA

Jellema W (2015) Optical design and performance verification of Herschel-HIFI. Thesis, University of Groningen

Jellema W, Jacobs H, Van Leeuwen B-J et al (2010) In: Proceedings of the 21st ISSTT, Oxford, UK, p 33

Kantorowicz G, Palluel P (1979) Backward wave oscillators. In: Button K (ed) Infrared and millimeter waves, vol 1, Chap. 4. Academic Press

Kinch MA, Rollin BV (1963) Detection of millimetre and sub-millimetre wave radiation by free carrier absorption in a semiconductor. Br J Appl Phys 14:672

Kooi JW (2008) Advanced receivers for submillimeter and far infrared astronomy. Thesis, University of Groningen

Lis DC, Goldsmith PF, Bergin EA et al (2009) Hydrides in space: past, present and future. In: Lis DC, Vaillancourt JE, Goldsmith PF, Bell TA, Scoville NZ, Zmuidzinas J (eds) Submillimeter astrophysics and technology: a symposium honoring Thomas G. Phillips. ASP conference series 417

Melnick GJ, Stauffer JR, Ashby MLN, Bergin EA, Chin G et al (2000) The submillimeter wave astronomy satellite: science objectives and instrument description. Astrophys J 539(2):L77–L85

Orleanski P et al (2012) In: 19th International conference on microwave radar and wireless communication (MIKON), vol 2, pp 426–432

Phillips TG (2009) In: A symposium honoring Thomas G. Phillips. Submillimeter astrophysics and technology. ASP conference series 417, p 160

Phillips TG, Jefferts KB (1973) A low temperature heterodyne bolometer for millimter astronomy. Rev Sci Instrum 44:1009

Putley EH (1965) Indium antimonide submillimetre photoconductive detectors. Appl Opt 4:649

Richardson KJ, White GJ, Avery LW et al (1985) Submillimetre CO studies of high velocity flow sources. Astrophys J 290:637

Roelfsema PR, Helmich FP, Teyssier D et al (2012) In-orbit performance of Herschel-HIFI. A&A 537:17

Van Vliet AHF, De Graauw T, Lidholm S et al (1982) A low-noise heterodyne receiver for astronomical observations operating around 0.63 mm wavelength. Int J Infrared Millim Waves 3:817

Verghese S, McIntosh KA, Brown ER (1997) Optical and terahertz power limits in the low-temperature-grown GaAs photomixers. Appl Phys Lett 71(19):10

Ward J, Rice F, Zmuidzinas J (1999) Supermix: a flexible software library for high-frequency circuit simulation including SIS mixers and superconducting components, Tenth international symposium on space Terahertz technology. University of Virginia, Charlottesville

Chapter 9
Superfluid Helium Cryostat Customisation

Abstract Here we describe the various elements of the *Herschel* cryogenic concept and its evolution since the first proposal submitted to ESA in 1982, following developments from previous cryogenic missions, and continuous adaptation to the evolving needs of the focal plane instruments, ending in a substantially different concept. A temperature below 2 K was necessary for the *Herschel* instruments to achieve the required sensitivity. *Herschel* therefore carried a liquid He cryostat of 2300 l capacity, with the liquid in the superfluid phase at a temperature of 1.6 K. This temperature was optimum for the photoconductive detectors of the PACS spectrometer and for the SIS mixers of the heterodyne HIFI instrument. The detectors of the PACS camera and of the SPIRE camera and spectrometer were bolometers, which required an even lower operating temperature. To achieve this, PACS and SPIRE had their own ^{3}He sorption coolers to reach a temperature of 0.3 K. The telescope was launched warm and, once in orbit, cooled passively down to a temperature of ~85 K.

Herschel's instruments could only provide the desired sensitivity if they were cooled down to cryogenic temperatures. To do this, *Herschel* was launched with a cryostat containing 2370 l of liquid helium. When exposed to the vacuum of space, the liquid attained a temperature of 1.6 K and was in a superfluid state. The liquid helium slowly boiled off, with the gas routed through pipes that cooled the instrument enclosures to around 5 K. Some internal parts of the instruments were cooled to about 1.7 K by direct straps (high-conductance metallic connections) to the liquid. The 1.7 K provided by the cryostat was optimum for the photoconductors of PACS and the SIS mixers of HIFI, but the PACS and SPIRE bolometers needed to have their own internal ^{3}He sorption coolers to achieve their lower operating temperature of around 0.3 K. The telescope was launched warm, and once in orbit it was cooled passively down to a temperature of ~85 K. This cryogenic system was quite different to what had been originally proposed in 1982. We describe here its evolution with respect to previous cryogenic missions and its adaptation to the instrument and scientific needs as a customisation of the ISO[1] cryostat.

[1]Infrared Space Observatory

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_9

9.1 Principles of Cryogenic Devices

One of the simplest ways to achieve cold temperatures in space is to use the available natural environment: the almost perfect vacuum and the ability to radiate thermal energy into cold space. The absence of any atmosphere limits heat exchange by gas conduction, leaving only conduction by solid elements and radiation, which can be influenced by careful design. A warm object will tend to cool down by radiation if it is in a colder environment. The cosmic microwave background in space has an equivalent temperature of about 2.7 K, so that a surface on a satellite could cool down to temperatures of a few K if it were isolated from any other heat input such as solar or Earth radiation or parasitic conduction from the satellite itself. The efficiency of this radiative cooling is proportional to T^4 and so decreases rapidly with temperature. With well-designed systems using optimised passive radiators, temperatures of around 30–40 K can be achieved. To go lower than that requires active cooling (Fig. 9.1).

One solution is to use a cryogenic liquid such as helium, liquefied on Earth and stored in a reservoir called a cryostat. Cryostats are devices, similar in principle to a vacuum flask, that maintain the liquid inside, and any objects in contact with the liquid, at low temperatures. The thermal load on the system from any heat input (conducted, radiation or dissipation) is carried away by the gas which slowly boils off the liquid, maintaining the liquid at constant temperature and removing heat from the system. The gas is usually vented to the vacuum of space, leading to a decrease of the internal pressure and thus reaching even lower temperatures (this phenomenon is similar to water boiling at a lower temperature at high altitude, where the atmospheric pressure is lower). Solid cryogens are also sometimes used and are easier to control in space than liquids (phase separation and microgravity effects are less problematic). However the achievable cryogenic temperature is

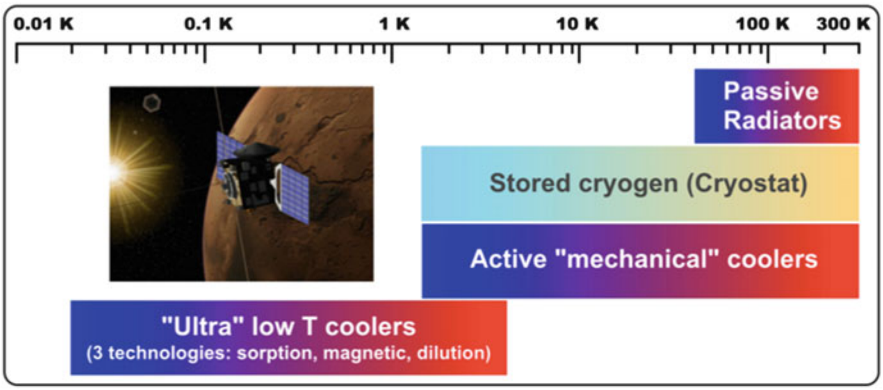


Fig. 9.1 Cryogenic system and their performance in terms of reached temperature, from radiative cooling to cooling machine. © Duband (2015)

limited to about 10 K using solid hydrogen (Duband 2015). Other cryogens are liquid neon (30 K), solid nitrogen (50 K) and liquid nitrogen (77 K).

Refilling a cryostat with liquid helium in space would be a very difficult and costly exercise. So the design of a space-borne cryostat must minimise the heat load on the liquid to ensure that it boils off as slowly as possible and so lasts as long as possible. For this reason, the outer vessel of the cryostat needs to be kept as cold as possible. Parts of it that are exposed to the rest of the spacecraft are covered in highly reflective material to reduce heat absorption, and the parts that are exposed to cold space are made black so as to radiate heat with maximum efficiency. Inside, the cryostat consists of multiple layers, like a Russian doll, to isolate radiatively the interior from the surroundings. A well-designed cryogenic system uses the cold gas generated as the liquid evaporates to cool the surrounding structures and further reduce the heat load on the cryogen.

Although the use of helium tanks is a simple technique that offers many advantages, it has major drawbacks: high mass, large volume and limited lifetime, as explained above. Therefore there is a strong push in the field of space cryogenics to replace cryostats with mechanical cryocoolers (Duband 2015). In a mechanical cooler, the work of an engine is utilised to provide refrigerating power, much as in a domestic refrigerator. They operate by alternately compressing and expanding a working fluid, such as helium, in a closed cycle, with cooling occurring as the fluid expands. These cryocoolers use no consumables and are sufficiently reliable to run almost indefinitely as long as electric power is available from the satellite's solar panels. A typical active cryocooler is the Stirling cycle refrigerator, which can achieve temperatures of around 70 K. Lower temperatures (around 4 K) can be achieved with Joule-Thomson coolers.

Active Coolers

Active coolers can be categorised into regenerative cycles (Stirling, pulse tube, Gifford coolers) and recuperative cycles (Joule-Thomson or Brayton coolers). The necessary power source can be obtained by using photovoltaic panels, available for almost all the space missions. Rando (2010) defined these devices as follows: 'Regenerative coolers are based on a pressure wave generated by a compressor (usually mechanical), and a cold finger, using a mobile regenerator. The heat is extracted at the cold end when the gas expands, and rejected at the warm end when the gas is compressed. The recuperative cycles use the enthalpy difference between high- and low-pressure gas'. In 1998, NASA successfully tested another type of mechanical cooler on a Shuttle mission: a reverse Brayton-cycle cryocooler that can cool detectors to temperatures as low as 60–70 K. This test showed that operation under weightless conditions was not a problem for this type of cooler, but long-term reliability was still an open question, though the same coolers had a good record in the laboratory and could run for many months to

(continued)

a few years. Reliable closed-cycle cryocoolers are certain to affect critically the design and life spans of future infrared astronomical missions in space. The reverse Brayton-cycle cryocooler has been retrofitted to the NICMOS instrument on *Hubble* Space Telescope to extend its useful lifetime beyond the almost two years achieved with the initial charge of solid nitrogen (Jedrich et al. 2002).

Instruments using bolometers as detectors require even lower temperatures, in the range 0.1–0.3 K. To reach these ultralow temperatures, three main technologies have been developed: ^{3}He evaporative cooling using a sorption pump (sorption cooler), dilution refrigeration and adiabatic demagnetisation (Duband 2015). Evaporative cooling relies on the same basic physical mechanism involved in cooling by perspiration: the transition of water from the liquid state to the vapour state requires energy. The vapour carries away the most energetic molecules and so takes heat away from the liquid. In the case of sweating, this mechanism helps to regulate body temperature. This property extends to almost all fluids. In the case of *Herschel*, evaporative cooling of the liquid helium in the cryostat resulted in the liquid cooling to around 1.6 K, at which temperature the liquid is in a quantum mechanical state known as a superfluid. Lower temperatures are difficult to achieve because superfluids do not evaporate easily. ^{3}He, a different isotope of helium in which the atomic nucleus has two protons and only one neutron instead of the usual two, does not become a superfluid and can be used to attain temperatures as low as around 0.25 K. However, an ultralow temperature refrigerator will not be able to operate from room temperature as the starting point.[2] In the case of the ^{3}He coolers used on *Herschel*, the starting temperature was 1.7 K, provided by the liquid helium tank.

^{3}He coolers, including those used on *Herschel*, operate via a closed cycle in which the ^{3}He gas that evaporates is collected and recycled, never escaping the system. The energy needed to cool the gas to convert it to liquid form is provided electrically. Once the liquid has formed, the chamber in which it resides must be pumped, creating a near vacuum into which the liquid can evaporate efficiently. This is achieved without any moving parts using the adsorption properties of a highly microporous material such as activated charcoal acting as a 'cryopump'. Such materials have very large internal surface area (over 1000 m^2/g). When the material is cold (a few K), ^{3}He gas molecules are adsorbed and so trapped, and when it is warmer (a few tens of K), the molecules are ejected. In a ^{3}He refrigerator, once all the liquid has boiled off and all the ^{3}He is trapped in the cryopump, it must then be recycled by warming up the pump, ejecting the gas, reliquefying it and starting the cycle once again. The adsorption and desorption temperatures depend on the nature of the gas. In the case of helium, by cycling between thermally

[2]It is important to note that at low temperatures the gains should not only be assessed in terms of difference but rather also in temperature ratio. In this case cooling from 2 K to 300 mK (ratio of 7) is almost as difficult as cooling from room temperature (293 K) to −230 °C (43 K).

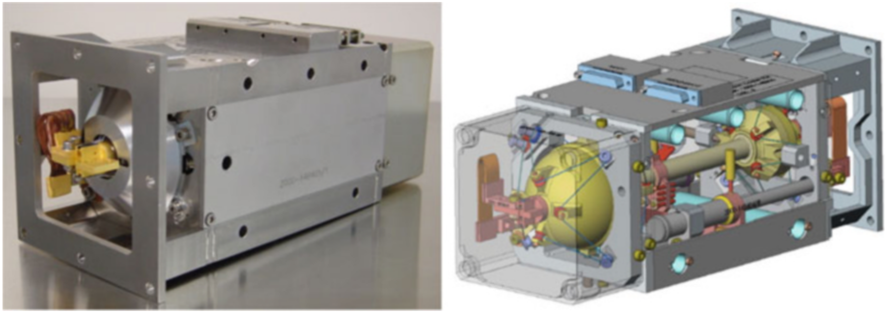


Fig. 9.2 ^{3}He cryocooler design used for PACS and SPIRE. © Lionel Duband (CEA)

activated charcoal at −230 °C (45 K) and −270 °C (3 K), one can generate the pressures required for the condensation of the liquid and then for its pumping to lower the pressure and thus lower the temperature. Figure 9.2 shows a photograph and a cut-away diagram of a ^{3}He sorption cooler similar to the ones used on *Herschel*.

9.2 Initial Concepts: From Cryocoolers to Superfluid Helium Cryostat

The original proposed payload for FIRST consisted of three instruments with the photometers and medium-resolution spectrometers using bolometers and photoconductors in a direct detection mode and the heterodyne receivers operating with Schottky diode mixers (see Chap. 3). As the direct detector instruments required similar temperatures to those in ISO but with the addition of a ^{3}He stage, it was envisaged in 1982 that a cryogenic system, similar to the one being considered at the time for ISO,[3] would be used. The heterodyne receivers with Schottky diode mixers would benefit from cooling as well, but at 70 K most of the advantage of cooling would have been obtained. This temperature level was going to be provided by Stirling cycle coolers for which very promising developments were already taking place in the USA and Europe. So at that time, the cryogenic system of the FIRST proposal relied on a combination of a cryostat and mechanical coolers.

However, by 1983 the ISO cryogenic concept was not so clear cut. The phase-A concept for the cryostat was not based on superfluid liquid helium (He-II) only but on a dual cryogen system consisting of 750 l of liquid hydrogen and 750 l of liquid helium, to achieve the required 18 months minimum lifetime within the constraints of the Ariane 2 launcher. A serious worry was that ground operation with liquid hydrogen was an unexplored territory and considered very risky. In 1985, when it

[3]ISO was launched in 1995 with a totally different cryogenic system.

became clear that by the time of the ISO launch a more powerful rocket, Ariane 4, would be available, the dual hydrogen-helium cryogen concept was abandoned. The switch to an Ariane 4 launcher enabled a helium-only system to be adopted.

Around the same time, successful demonstration of a space He-II cryostat by the IRAS mission was very encouraging. One of the key design features was the use of a porous plug as 'phase separator'. This controlled the flow of the liquid as it boiled away, allowing the liquid-phase helium to evaporate into the gas phase in a controlled manner. The successful demonstration of the IRAS cryogenic system validated the concept for the ISO cryostat. However, the cryogenic design for FIRST was still quite open as there were several uncertainties to deal with. First of all, the space qualification of mechanical coolers, and specifically the Stirling refrigerator, was not yet demonstrated. Furthermore the positive developments of SIS mixers (see Chap. 8) with higher sensitivity, and capable of operating at higher frequencies, changed the cooling requirements for the heterodyne instrument—it now needed to have mixers operating at a few K. In the mid-1980s, several cryogenic studies were carried out with mechanical coolers only and with a He-II cryostat only as the extreme options and several hybrid solutions in between with a combination of a He-II cryostat and 4 K mechanical coolers. These studies were pursued in an attempt to put FIRST in the context of the Columbus space station programme.

By 1987–1988, the Science Advisory Group (SAG) and the Telescope and Payload Working Groups (TWG, PWG) were in place with the task of trading off the scientific objectives against the technical complexity and cost constraints supported by a System Definition Study (SDS). This study was carried out by Dornier Systems and aimed at investigating various solutions concerning the use of a liquid helium cryostat or cryocoolers. It was completed in December 1990. On the basis of technical risk, total programme cost and scientific priorities, the SAG recommended that the baseline mission include the following elements: a non-deployable 4.5 m passively cooled telescope, a single He-II cryostat for the payload with both heterodyne and direct detection instruments and a launch by Ariane 5.

Later in the SDS, it was recommended to have a cryostat containing 3200 l of superfluid helium to allow a lifetime of three years, the mission requirement. However this did not fit in the cost budget and the concept had to be abandoned. A 'rider' study took place in 1992–1993 (see Chap. 3). Several options were explored, from decreasing telescope diameter to removing the cryostat completely and having a heterodyne-only mission. The SAG rejected the heterodyne-only option. In June 1993, the mission concept, as presented in the phase-A report (the Red Book), contained a 3-m single-panel telescope, a cooling system of closed-cycle Stirling refrigerators providing a 4-K environment for the science payload, a suite of instruments consisting of multi-frequency heterodyne receivers and a direct detection imaging/spectroscopy instrument, a 24-h highly elliptical Earth orbit and a lifetime of at least two years. However, the Stirling cycle cooler technology was not yet fully developed. As the instruments' detectors were to operate in the 0.1–4 K temperature range, further cooling would need to be provided by the instruments themselves. ESA set up and implemented a technology development programme that included the development and test of the Stirling cooler cryogenic subsystems including the qualification of the 20-K and 4-K coolers.

In parallel to the research and development of cryogenic technologies for FIRST, the development of the ISO He-II cryostat and its qualification tests were very successful. It was decided to carry out a study to assess the implications of using the liquid helium cryostat technology developed for ISO as an alternative to the Stirling cooler system. The Payload Working Group re-optimised the scientific payload for the cryostat alternative, which provided the input for the cryostat study. Daimler-Benz Aerospace carried out the study confirming that an ISO-like cryostat was a viable alternative to the Stirling cooler system. This conclusion led to trade-off studies between the cryocooler and the cryostat options, which took place in 1995–1996. The SAG and the associated working groups provided the science inputs to both studies. Both options were based on a 3-m telescope. Matra Marconi Space (MMS) studied the cryocooler option. The cryostat study was carried out by Aerospatiale and Daimler Benz-Aerospace (AS/DASA). The trade-off study report was made available in October 1996. In their analysis of the study results, the cryostat option emerged as the priority option preferred by the SAG.

Following the extensive scientific and technical evaluation by ESA of the trade-off study results, including risk, choice of orbit and cost, and taking account of the SAG recommendation, the decision was made to implement the FIRST mission with a cryostat spacecraft and to conduct operations from an orbit around L2. This concept was used for the Announcement of Opportunity (AO) for the FIRST scientific instruments issued in 1997 (Fig. 9.3).

At launch in 2009, the *Herschel* cryostat was described (Pilbratt et al. 2010) as:

> The Payload Module is dominated by the cryostat vacuum vessel from which the superfluid helium tank is suspended, surrounded by three vapour-cooled shields to minimise parasitic heat loads. The optical bench with the three instrument focal plane units (FPUs) is supported on top of the tank, which has a nominal capacity of 2367 l. A phase separator allows a continuous evaporation of the liquid into cold gas. The FPUs and their detectors are kept at their required temperatures by thermal connections to the liquid cryogen in the tank and to pick-off points at different temperatures of the piping that carry the helium gas from the tank, which is routed around the instruments for this purpose, and to the vapour-cooled shields for eventual release into free space.

Having described the main options for the *Herschel* cryogenic system, we devote the next sections to analysing the technical lineages of helium cryostats and their level of innovation (as introduced in Chap. 5).

9.3 Liquid Helium Cryostat Lineage

The liquid helium cryostat was of course not a new concept or a radical innovation. It was known since the first liquefaction of helium by Kamerlingh Onnes in 1908 and afterwards by Collins in 1947. But until 1960 the cryostats were all made from

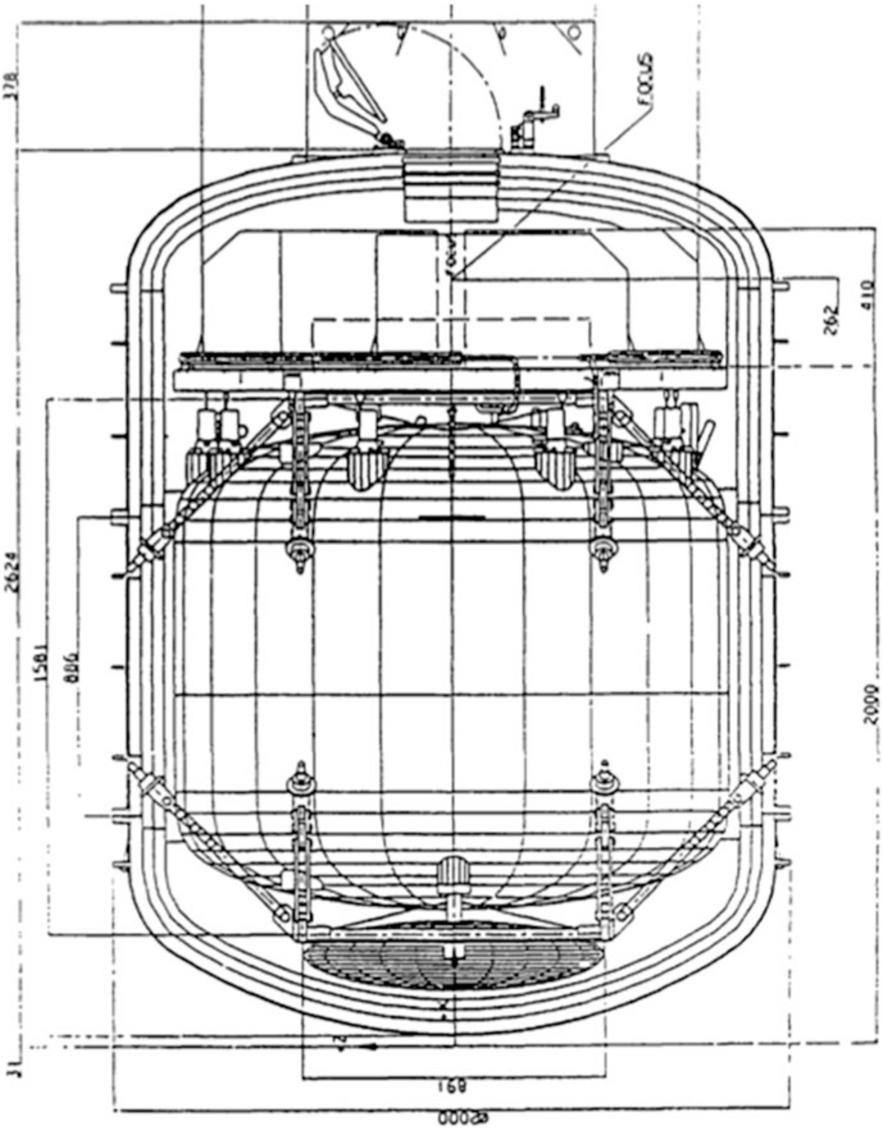


Fig. 9.3 *Herschel* cryostat design in 1997. © Juillet et al. (1997)

glass. An important breakthrough occurred when Frank Low (see Chap. 2) introduced an all-metal liquid He cryostat with radiation shields. He used these in various infrared instruments that he commercialised (via Infrared Laboratories, Inc.) together with the germanium bolometers that he had also developed. For decades Infrared Labs provided the cryostats and detectors to the worldwide astronomical community for ground-based and balloon-borne observatories.

This concept was applied in the IRAS satellite, which had a 600 l superfluid helium cryostat. Superfluid helium was known since the work of Soviet physicist Pyotr Leonidovich Kapitsa in 1938. It presented two advantages for a space mission: its greater cooling capacity, achieving temperatures down to around 2 K, and the fountain effect, useful for fluid management under zero gravity (Castles and Di Pirro 1990). This required also the development of a vapour-liquid phase separation using a porous plug to control the flow of the He gas and thus the temperature. The porous plug was developed in the USA as well as in Europe. This cryostat concept was used in numerous space projects (IRAS, IRTS, WIRE, COBE, etc.) as attested by several NASA patents, and back in 1990 it was foreseen to be used in SIRTF (later known as *Spitzer*). IRAS demonstrated the capability of passively controlling the superfluid helium temperature in a space-borne cryostat (Petrac and Mason 1984).

Example of NASA Patent: Cryostat System for Temperatures on the Order of 2 K or less
US 3983714 A—1975–1976

A cryostat system for cooling a device to a temperature on the order of 2 K or less includes a dewar, in which helium, in other than the superfluid state, is stored. Helium flows from the dewar through a heat exchanger tube and a restrictor tube, which controls the helium flow rate, into the cavity of a heat exchanger, to whose outer wall the device to be cooled is attached. A pressure regulator valve controls the pressure in the cavity to be very low, e.g. on the order of 30 torr. As the helium exits the restrictor tube into the cavity, due to low pressure cavity, it becomes an aerosol mixture of helium gas and superfluid helium droplets at the desired temperature. The latter form a thin layer or film of superfluid helium on the inner side of the heat exchanger wall and thereby cool the device, which is attached to the wall to the desired temperature. The helium gas, formed during the exit of the helium into the cavity, and the helium gas, formed from the superfluid helium, which is evaporated by absorbing heat from the device being cooled, are evacuated from the cavity. As they flow around the heat exchanger tube, through which helium flows from the dewar, heat is absorbed by the helium gas from the helium in the tube, so that the helium entering the cavity is at a lower temperature than the helium entering the tube from the dewar. The evacuated helium gas may be used for one or more purposes, including reducing the amount of radiated heat reaching the dewar, as well as serve as the propellant for spacecraft attitude control. The cryostat may be used to cool different devices to different temperatures on the order of 2 K or less during an entire mission or during selected independent periods for each device.

http://www.google.com/patents/US3983714

The launch of IRAS in 1983 coincided with the selection of ISO as an ESA mission. But, as described above, the ISO cryostat cannot be considered as a direct scaled-up version of the IRAS cryostat. Actually, He-II cryostat designs were already developed in Europe well before IRAS was launched. In 1977, Dornier conducted studies on a cryogenic system using superfluid helium to cool the telescope structure, optics and detectors of a space-borne IR measurement system (Becker et al. 1977). For the company that developed the ISO cryostat (MBB in Ottobrunn near Munich), the cryostat and the porous plug development started with GIRL (German IR Lab), a space experiment intended to be flown on Spacelab. Developments of a 300-l He-II cryostat and the instruments for that project were quite advanced. The scientific programme included studies related to infrared astronomy, the determination of characteristics of the Earth's atmosphere and He-II physics. The scientific instruments included a camera, a photo-polarimeter, a grating spectrometer and an FTS spectrometer. Dietrich Lemke was developing the photometers, and Siegfried Drapatz from MPE was developing the spectrometer and in particular the drive mechanism. GIRL would be mounted to the Spacelab Instrument Pointing System (Lemke et al. 1981). The design of GIRL, as presented in 1979, had a toroidal liquid He tank like IRAS. Whether there was significant information exchange on cryostat designs between the IRAS and GIRL teams is not clear, but it is very unlikely. With the development of ISO confirmed, the German space agency (DLR/DARA) decided to abort the GIRL project.

In summary, the *Herschel* cryostat fits within the technical lineages of liquid helium cryostat following the path of ISO, the GIRL project and IRAS using a porous plug to vapour-liquid phase separation.

9.4 Customisation and Innovation

We may then ask whether there was any innovation that could be identified in the cryostat design and development. The level of innovation was defined in Chap. 5 as ranging from customisation to radical innovation, i.e. invention. The *Herschel* telescope was described as a radical innovation (Chap. 6), while the bolometers in PACS and SPIRE were breakthrough and incremental innovations (Chap. 7).

As the cryostat is a potential single-point failure for a cryogenic space mission, a cautious approach to the design is essential. Besides cooling the scientific instruments, space-borne cryostats need to meet several important technical requirements. They must be strong enough to survive launch; parasitic heat inputs from mechanical supports, radiation, wires, pipes, etc., must be minimised; and one has to optimise the use of the cooling powers available such as the liquid, the cold boil-off gas at various temperatures and passive cooling by radiating into deep space. There must be a compromise between mechanical robustness, requiring bulky supports for the inner cold stages, and low parasitic thermal conduction, which is best achieved with thin supports. Then radical innovation might become risky. With the intention of having a secure and robust approach, the *Herschel* cryostat design

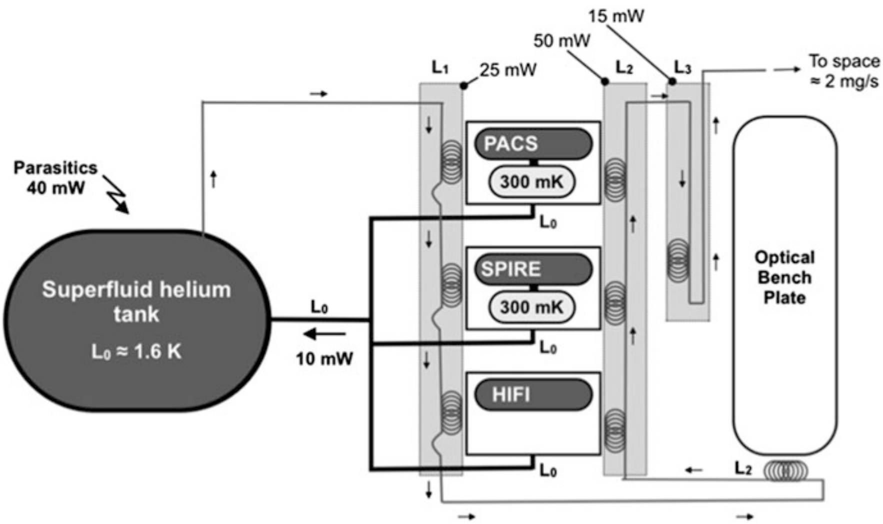


Fig. 9.4 Thermal architecture used for the *Herschel* mission, which featured a 2370-l superfluid helium tank. The exhaust vapours were used to cool several radiation shields and the SPIRE, PACS and HIFI instruments to temperatures ranging from 4 K to 15 K, in order to minimise the overall heat load on the liquid. © Duband (2015)

was based firmly on the ISO cryostat. The thermal architecture of the *Herschel* system is shown on Fig. 9.4. To reach the 0.3 K temperature for the PACS and SPIRE bolometers, the additional ^{3}He cryocoolers were installed within the instrument's focal plane units, and close to the detectors. The key features of the system were already well established.

Air Liquide,[4] the company that built the cryostat tank, claimed that innovative techniques and materials were used in that the project was very demanding in terms of leak tightness (around one million times better than the fuel tanks on the Ariane rocket), with leak rates at the limit of detection. To achieve this, specialised materials had to be used, and new welding methods had to be developed and qualified. Ultrahigh cleanness was another challenge: the tank was assembled in a specially built class-100 clean room. The cold temperature was brought to the instruments using ultrahigh thermal conductivity thermal links. Some links at 1.6 K were protrusions from the main tank enabling the superfluid liquid to be brought as close as possible to the instruments through the use of the superconducting fountain effect (Balcet et al. 2005).

Although the essential features of the cryostat system were well defined, the instrument teams raised some technical issues about the potential electromagnetic interaction between the solar panels and the cryostat and consequently with the instruments inside. Albrecht Poglitsch, the PACS PI recounts a conflict concerning

[4] http://www.airliquideadvancedtechnologies.com/en/our-offer/space/programs/un-cryostat-dhelium-liquide-pour-Herschel.html

the routing of cables on the outer wall of the vacuum vessel, which was important to avoid electrical interference with the instrument:

> Astrium decided to route the cable harness in a certain way, distributing the cables all around. There were cables running on the outer wall of the vacuum vessel, and thus under the blanket of super-insulation, or MLI (multi-layer insulation), which was used to minimise the radiative thermal load from the environment on that vessel, such that it stays as cool a possible and, thus, minimizes the thermal load on the helium tank inside. But it provided no magnetic shielding although the solar panels and their wiring were close. Cutting MLI open and patching it up introduced leak paths between the layers and thus deteriorated the insulation. In addition, from an electromagnetic point of view, it was a sort of a deathly scene because it was creating loops, and then when there are variable magnetic fields it induces voltages and perturbations in the system. This issue was discussed with the Project Manager who claimed that industry had studied this and that there was no problem. Then there was an ongoing debate for a long time. I remember our electric system engineer was incredibly concerned when the Astrium people were about to put on the cables and warned them again. But they did not change their plan. Thomas Passvogel did not want to change it either. So, they put it the intended way, which was wrong. It cost us an enormous effort to prove it to Thomas, theoretically but also experimentally, including a real test on the real satellite. But when he saw that, the same day he got the people from whoever was in charge of the super-insulation to cut it open, after it was already closed, and re-route the cables, and then patch the insulation back on.

The *Herschel* cryostat (Fig. 9.5) was thus the result of customisation of the ISO cryostat rather than a breakthrough innovation. As defined in Chap. 5, customisation is the adaptation of a preexisting technology (ISO cryostat) to a new environment (L2) and requirements—unlike ISO, the instruments in the *Herschel* cryostat were not cooled just by the liquid but by an elaborate combination of liquid and gas boil-off. An over-adaptation could induce a loss of genericity, which means that the instrument will not work as well or not at all in another environment or if a strong variation of the environment occurs. Indeed the ISO and *Herschel* cryostats were specially adapted to their own orbital environments and to the different requirements of their scientific instruments.

The cryostat of the *Spitzer* satellite was also a liquid helium cryostat, based on an IRAS-like cryostat. However, it benefitted from an innovation in managing passive cooling which had improved significantly with the cryogenic concept of the EDISON project, proposed in 1990, involving the V-groove concept (Hawarden et al. 1992). V-groove shields are based on a small number (typically three) of angled and highly reflective solid plates, with the open end of the V directed towards cold space and emitting radiation after a number of reflections between the shields (Rando 2010). Frank Low (again) immediately saw the merits of this design and incorporated it into the SIRTF/*Spitzer* cryogenic system, enabling a dramatic cost reduction for the mission. The same approach was used by ESA for the *Planck*[5] satellite but not for *Herschel*. Like GIRL, the EDISON concept thus had a major technological impact even though the mission was never flown.

[5] http://sci.esa.int/planck/45498-cooling-system/?fbodylongid=2123

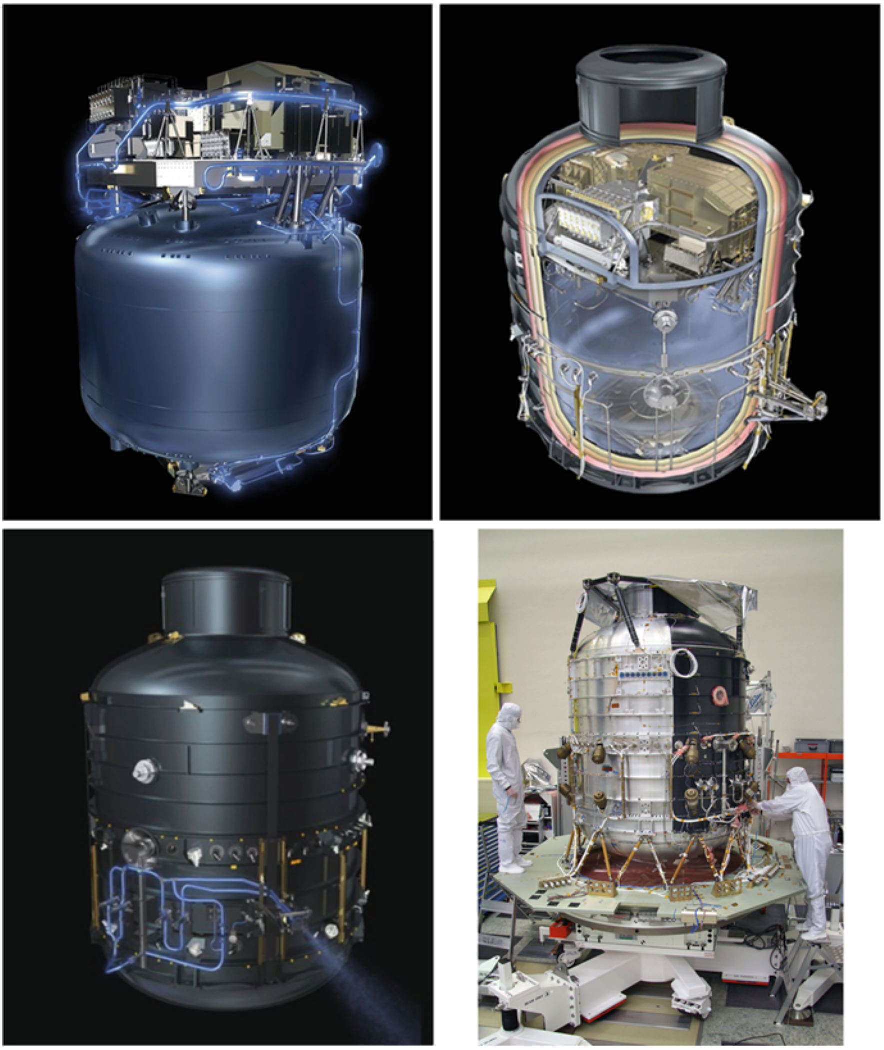


Fig. 9.5 *Top left*: *Herschel's* superfluid helium II tank, with the optical bench and the instrument focal plane units on top and piping (shown in *blue*) to route the cold system showing additional boil-off gas around the instrument enclosures. *Top right*: A cutaway view of the complete system now additional radiation shields around the helium tank and the outer cryostat vessel. *Bottom left*: A view showing the cryostat as seen from the cold space side. This side of the cryostat was made black to allow it to radiate effectively. The boil-off gas, now at the ambient temperature of about 74 K, is shown schematically emerging from the piping system. *Bottom right*: A photograph of the cryostat being assembled. © ESA/*Herschel*

9.5 Mission Lifetime and Cryostat Volume

In summary, several constraints are associated with the space environment, and when it comes to cryogenic cooling, the main ones are to survive the launch; to minimise the mass, volume and power required; to be compatible with the operation in microgravity; and to guarantee reliability and lifetime over several years. The latter leads to designs involving moving parts without friction and ideally no moving parts.

A system with a fixed amount of stored cryogen has a limited lifetime: unless replenished (an approach which has not been adopted for any astronomical space mission), the cryogen is eventually fully depleted and the system warms up. In the case of *Herschel*, none of the scientific instruments could work unless helium-cooled, so the mission then came to an end. As written by Duband (2015), 'although the liquid cryostat is a simple and straightforward technique, it leads to heavy reservoirs and structures and by essence limits the mission duration'. During the design of *Herschel* and later during its operation, every effort was made to maximise the mission lifetime. Then, once *Herschel* was operating, extensive efforts by engineers optimised the helium liquid consumption and extended the mission lifetime of a few months. Comparable optimisation for ISO expanded its mission lifetime from the required 18 months to 22 months.

Based on Table 9.1, Fig. 9.6 shows the evolution of the ratio between the mission lifetime (in days) and the cryostat liquid helium volume (in litres). The value remains constant at 0.4–0.5 from IRAS to *Herschel* with the exception of the *Spitzer*, which exhibits an improvement by a factor of 10. This is partly explained by the optimisation of the radiative cooling leading to a lifetime exceeding 2.5 years for only 360 l of liquid helium on board. Additional compromises between lifetime and science capabilities of the instruments also explain the *Spitzer* value—the *Spitzer* instruments were very stringently restricted in the amount of electrical power that they could dissipate, which meant that their scientific capabilities were correspondingly constrained. While customisation seems to prevail in the sequence from IRAS to *Herschel*, *Spitzer* included a specific innovation, the V-groove that was described above in Sect. 9.4.

Table 9.1 Volume and lifetime of helium (He II) cryostats flown (from Duband 2015)

Mission	Year launched	Cryogen volume (litres)	Lifetime (months)
IRAS	1983	600	10
COBE	1989	600	10
ISO	1995	2250	22
SFU–IRTS	1995	100	1
Spitzer	2003	360	48
Gravity Probe B	2004	2240	17
Suzaku	2005	30	0.5
Herschel	2009	2370	46

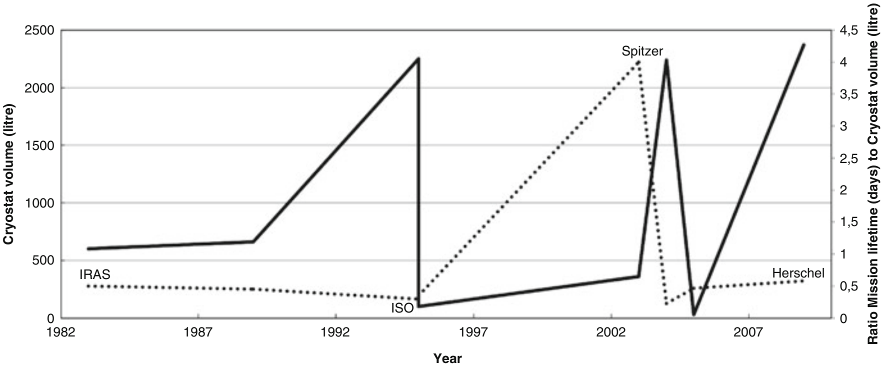


Fig. 9.6 Evolution of cryostat reservoir volume (*black line*) and of the ratio mission lifetime vs. cryostat volume (*dashed line*). Values are based on Table 9.1. Four iconic missions are indicated: IRAS, ISO, *Spitzer* and *Herschel*. © CEA (V. Minier)

9.6 Conclusion

Cryogenic systems are essential for the operation of sensitive far-infrared space missions. The development of complex cryogenic payloads calls for an integrated system approach, involving the complete spacecraft design from start. As noted in Chap. 5, successful technical assessment and design need to find a convergence of functions in a structural unity rather than searching for a potentially poor compromise between conflicting demands. Instruments and cryogenics should be designed together so that each object fits optimally in a circular causality with the environment constituted by the rest of the system. Cryogenics should then provide the necessary operating conditions for high-sensitivity instruments in space astrophysics. The *Herschel Space Observatory* combined radiative cooling (for the telescope), a liquid helium cryostat (for the instruments) and closed-cycle ^{3}He coolers (for the PACS and SPIRE photometer focal planes) in order to attain the proper operating conditions. The *Herschel* telescope was in a 'warm-launch' configuration (telescope outside the cryostat) compared to the 'cold' configuration of IRAS, ISO and *Spitzer* (Fig. 9.7).

Herschel's cryostat resulted from the customisation of ISO's cryostat. As stated in Chap. 5, customisation is the adaptation of a pre-existent technology to a new environment and system requirement and induces a loss of genericity. A customisation is therefore a 'one-shot' compromise rather than a truly innovation. It also limits the concretisation of *Herschel* in the meaning of Simondon's philosophical concept (see Sect. 5.6). The use of a liquid helium cryostat limits the mission lifetime to a few years and therefore breaks down the functional synergy with the instruments. Furthermore, the detectors operate at temperatures well below 1 K, requiring additional coolers within the instruments. Although this choice was economically motivated and scientifically acceptable for *Herschel*, it was a more conservative approach than the original ambition to use Stirling cycle cryocoolers.

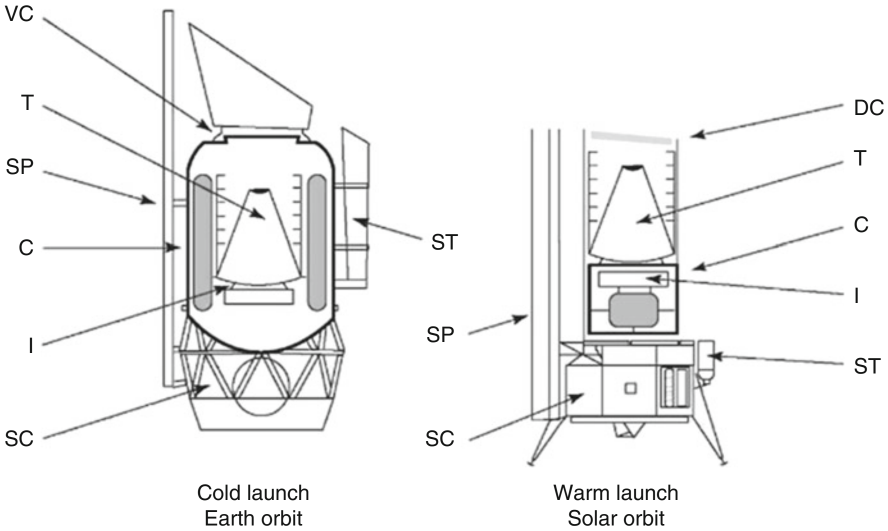


Fig. 9.7 Space observatories with cold launch architectures like IRAS, ISO or *Spitzer* (*left*) and warm launch architectures like *Herschel* (*right*). Components, such as the spacecraft (SC), the telescope (T), instrument payload (I), the solar panel (SP) and the star tracker (ST), are common to both systems. Each satellite also includes a cryostat (C), containing the liquid helium cryogen in a separate helium tank, which is shown shaded. In the cold launch system, the telescope is located within the cryostat and cooled by direct contact with the cryogen tank. In the warm launch system, the telescope cools radiatively to 100 K or below. The sawn-off conical sunshade at the top of the cryostat is required in an Earth-orbiting system by the Sun-Earth-orbit geometry. A much smaller sunshade is needed in the solar orbit system because Earth visibility is not a concern. © NASA

This would have been a more novel concept, with the potential advantage of longer lifetime, but the cost and risk were too great and the technology was not yet ready.

Today, significant development effort is still required to improve the performance of cryogenic systems, to enhance cooling efficiency and reliability and to reduce resource demands as well as cost. The nearly flat pattern of the cryostat technical lineage (Fig. 9.7) shows no fundamental progress since IRAS besides the capacity to adapt pre-existing cryostats to new orbital operation and technical requirements. Future detector technologies will require even colder temperatures of around 0.05 K, and future missions will indeed use mechanical coolers rather than cryostats to reduce launch mass and enable longer mission lifetimes (e.g. this is already the baseline concept for the proposed SPICA mission). Such developments need to be undertaken in the early project phases so as to minimise the risks during the implementation phase when the impact of design changes is largest. The significant efforts devoted in this field by the leading space organisations show that cryogenic systems are going to play an important strategic role in future space missions.

References

Balcet N, Roussel E, Wiertz T (2005) Welding qualification, and realisation, for *Herschel* tanks and cryostat. In: Fletcher K (ed) Proceedings of the European conference on spacecraft structures, materials and mechanical testing 2005 (ESA SP-581), 10–12 May 2005, Noordwijk, The Netherlands. CD-Rom, #33.1

Becker RJ, Bartholomae KP, Heidt FD, Schwille H (1977) Helium technology, infrared measurements, and space travel, Deutsche Gesellschaft für Luft- und Raumfahrt, Jahrestagung, 10th, Berlin, West Germany, 13–15 Sept 1977, Paper 17

Castles SH, Di Pirro MJ (1990) The fountain effect in aerospace cryogenics. In: Fast RW (ed) Advances in cryogenic engineering: Part A & B. Springer Science, Heidelberg, p 210

Duband L (2015) Space cryocooler developments. In: 25th International cryogenic engineering conference and the international cryogenic materials conference in 2014, ICEC 25–ICMC 2014. Phys Proc 67(2015):1–10

Hawarden TG, Cummings RO, Telesco CM, Thronson, HAJr (1992) Optimised radiative cooling of infrared space telescopes and applications to possible missions. Space Sci Rev (ISSN 0038-6308), 61(1-2):113–144

Jedrich N, Zimbelman D, Swift W, Dolan F (2002) A mechanical cryogenic cooler for the hubble space telescope. In: SPIE meeting

Juillet JJ, Trogus W, Schupp J, Cornelisse JW (1997) Far infrared and submillimeter space telescope—a satellite concept based on the ISO cryostat. In: Wilson A (ed) The far infrared and submillimetre universe. ESA, The Netherlands, p 385

Lemke D, Proetel K, Dahl F (1981) The German Infrared Laboratory/GIRL/—A status report. In: Shuttle pointing of electro-optical experiments; Proceedings of the Seminar, Los Angeles, CA, 10–13 Feb 1981 (A81–37177 16–19) Bellingham, Washington, Society of Photo-Optical Instrumentation Engineers, 1981, pp 366–374. Research supported by the Bundesministerium für Forschung und Technologie

Petrac D, Mason PV (1984) Infrared astronomical satellite (IRAS) superfluid helium tank temperature control. In: Advances in cryogenic engineering, vol 29. Springer Science, Heidelberg, p 661

Pilbratt GL, Riedinger JR, Passvogel T, Crone G et al (2010) Herschel space observatory—an ESA facility for far-infrared and submillimetre astronomy. Astron Astrophys 518:1

Rando N (2010) Cryogenics in space. In: Huber MCE, Pauluhn A, Culhane JL, Timothy JG, Wilhelm K, Zehnder A (eds) Observing photons in space, vol 9, ISSI Scientific Report Series. Springer Science, New York, pp 639–655

Chapter 10
Management and Organisation of Science Instruments

Abstract Here we describe how *Herschel* was built, following the usual ESA framework, with industry, consortia of universities and research laboratories, ESA and member states working together. Although the HIFI, PACS and SPIRE consortia organised themselves differently in managing time and resources, planning the use of their guaranteed observing time and exploiting their data, all shared the same science-driven management approach. Here HIFI is chosen as a typical case to illustrate the process and difficulties involved in creating and managing such consortia and to analyse how they organised their respective community and managed their development and their own scientific observation programmes. PACS serves as a typical case to study the interaction between the consortia, industry and ESA, and SPIRE illustrates the success-oriented approach of the instrument teams and subtle differences in the way they organised their observing programmes. The testimony of the instruments' principal investigators offers a unique comprehensive analysis of the management of large and complex space science instruments and analyses the social context for innovation in the light of coopetition and sociality concepts. It suggests that a less risk-averse approach to science instrument management, compared to the current conservative method of estimating technology maturity, could better drive technical innovation and lead to greater scientific success.

Industry and research organisations built the *Herschel Space Observatory* within a framework that was governed by the European Space Agency. Industry provided the spacecraft, the telescope and the cryostat, while research laboratories organised in consortia designed, built and exploited the three science instruments. The HIFI, PACS and SPIRE consortia organised their respective communities, proposed a specific instrument and managed their instrument development and their own scientific programmes. The three consortia organised themselves under different rules but all shared the same science-driven management approach. Through the testimony of the three instrument principal investigators (PIs), this chapter investigates the management of science instruments and the social context for innovation in the light of the coopetition and sociality concepts. It presents some evidence that a less risk-averse approach to science instrument management, compared to the current conservative method of estimating technology maturity, can better drive technical innovation and leads to greater scientific success.

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_10

10.1 ESA/*Herschel* Organisation

The *Herschel-Planck* programme constituted the largest industrial contract ever signed by the Scientific Programme of the European Space Agency in the framework of Horizon 2000 (Chap. 2). The prime contract was awarded to Alcatel, which later became Thales Alenia Space, which led an industrial consortium selected to meet the industrial return rules in member countries of ESA. A total of 95 European companies and industries participated in this dual project. Three industrial entities shared the responsibilities for the construction and delivery of *Herschel* and *Planck*: Thales Alenia Space, Astrium Toulouse and Astrium GmbH Friedrichshafen.

A space science satellite has two parts from the hardware point of view: the spacecraft, consisting of the service module and solar shield with panels, and the payload module that contains the telescope, the cryostat and the instruments manufactured under the responsibility of consortia led by public research laboratories. Thales Alenia Space in Italy was in charge of delivering the *Herschel* and *Planck* service modules. Astrium GmbH in Germany became responsible for the *Herschel* payload (cryostat and instruments). The *Herschel* telescope was built by Astrium, France, under ESA supervision. Thales Alenia Space in France was entrusted with the responsibility for the *Planck* payload module (Fig. 10.1).

The overall industrial effort was under the responsibility of a project manager appointed by Thales Alenia Space, Jean-Michel Reix, and the project manager appointed by ESA, Thomas Passvogel. ESA imposed a set of standards on the satellite construction and made integration and test facilities available to manufacturers. This complex organisation (Figs. 10.2 and 10.3) necessitated organising *Herschel* and *Planck* in the same way as a conventional large industrial project.

On the scientific side, the *Herschel* payload included three different and highly complex instruments, each comprising several distinct, specific components. Taken together the three instruments were designed to carry out photometric mapping and spectroscopic observations covering a wide spectral range from the far infrared to the submillimetre domain. The detection techniques required different instrumental

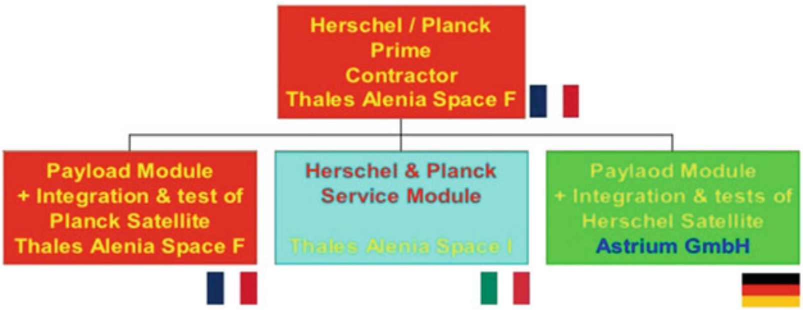


Fig. 10.1 Organigram of industry main contractors for *Herschel-Planck* programme. © ESA/*Herschel*

Fig. 10.2 European industries involved in *Herschel-Planck* programme. © ESA/*Herschel*

approaches relying on photoconductive, bolometric and heterodyne components, each with its particular advantages and limitations (Chaps. 7 and 8).

The resources required for the development and operation of such complex instruments were comparable to those required to build and operate a medium-sized spacecraft. Not surprisingly, each of the proposed *Herschel* instruments involved a large multi-national and multi-institute consortium, with each institute contributing specific components or services developed with public research money and manpower under national funding. Each consortium was under the responsibility of a principal investigator (PI) acting as the single formal interface of the consortium with ESA, through the Project Manager and Project Scientist. Each PI represented an institute and country providing a leading contribution to the instrument and was supported by one or more co-PIs and a set of co-Is responsible for other essential contributions (Chaps. 3 and 4).

In addition to industry and instrument organisations, a third network was responsible for the mission operation procedures, data calibration and data reduction process for optimising science products, known as the 'ground segment', which included teams in charge of the mission operation procedures and of the data reduction process. The *Herschel* Science Ground Segment (H-SGS) consisted of five elements: The *Herschel* Science Centre (HSC), an Instrument Control Centre (ICC) for the three *Herschel* instruments (HIFI, PACS, SPIRE), and the NASA *Herschel* Science Centre (NHSC). The HSC (based at ESAC, Madrid) was the prime interface between *Herschel* and the science community. It provided information and user support related to the entire life cycle of *Herschel* observations, including calls for observing time, the proposing procedure, proposal tracking, data

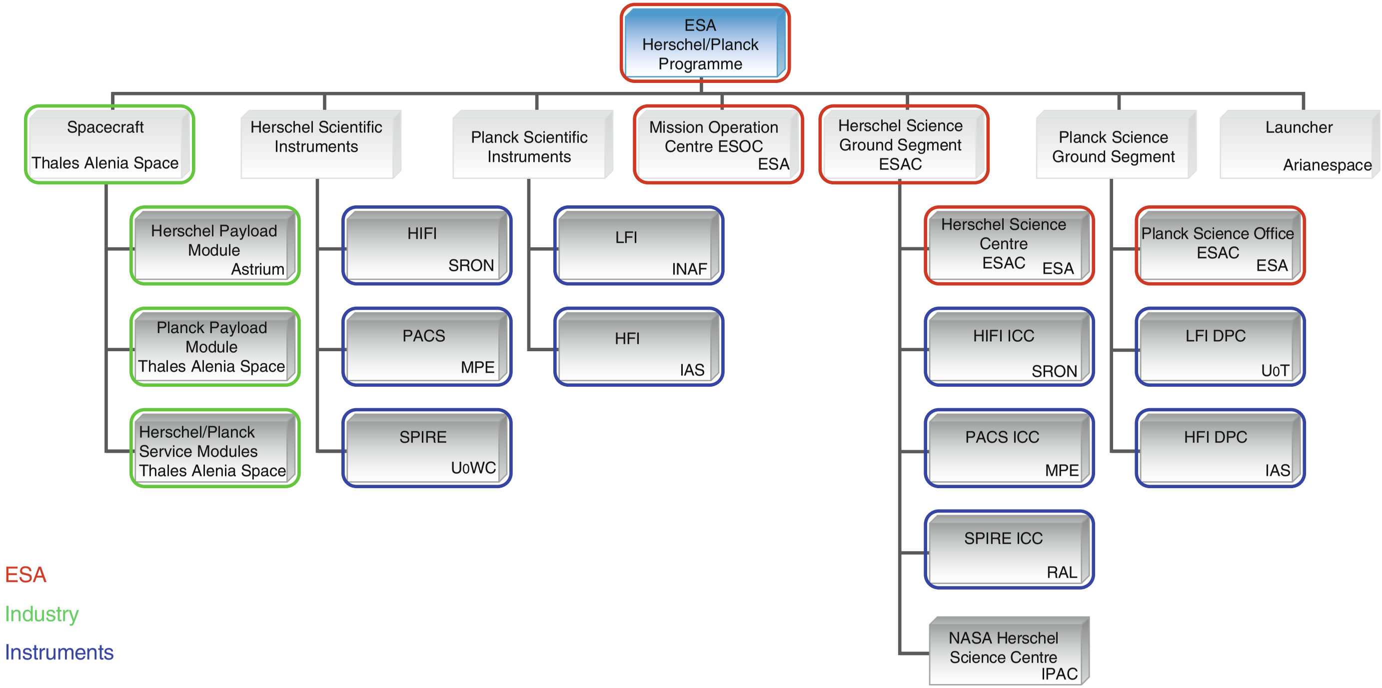


Fig. 10.3 Organigram of *Herschel-Planck* ESA programme showing the share of responsibility between industry (Astrium, Thales…) involved in the spacecraft and research organisations (SRON for HIFI, MPE for PACS, Queen Mary and Westfield College, London (QMWC) and then Cardiff for SPIRE) building the scientific instrument. *Right parts* show the ground segment organisation in charge of operating the satellite, scheduling observations, retrieving, reducing and archiving science datasets. © ESA/*Herschel*

access and data processing, as well as general and specific information about using *Herschel* and its instruments. It also performed scientific mission planning, produced observing schedules and provided these to the Mission Operations Centre (MOC, based at ESOC, Darmstadt) for uplink to *Herschel*. The ICCs were responsible for the successful operation of their respective instruments. They performed instrument monitoring and calibration and developed and maintained instrument-specific software and procedures for generating observations and for the processing of the data generated. The NHSC (based at IPAC, Pasadena) provided additional user support primarily for the US-based users of *Herschel* and offered science exploitation funding for *Herschel* investigators based in the USA.

The complete ground segment for the *Herschel* mission included the staff in ESA Centres at ESTEC (science and technology centre), ESAC (space astronomy centre) and ESOC (operation centre), as well as within the instrument consortia.

10.2 Management of Instrument and Science Consortia

The organisation and approach to scientific exploitation of the HIFI, PACS and SPIRE instruments were under the control of the instrument teams themselves and followed somewhat different paths and management schemes. The next sections outline how the HIFI, PACS and SPIRE consortia arranged themselves to propose and deliver the scientific instruments and to exploit their guaranteed time allocations. It is based on material provided by and interviews with the instrument PIs: Thijs de Graauw and Frank Helmich for HIFI, Albrecht Poglitsch for PACS and Matt Griffin for SPIRE. We do not intend to propose an exhaustive description of how each consortium was managed. The contributions focus on the main management features that emerged within the *Herschel* mission from its conception to science operation and tend to extract general social properties[1] that were important to *Herschel's* innovation and success.

The two first sections describe the HIFI consortium, from its genesis to the science operation. They provide a detailed overview of a consortium constitution and set the scene for the management complexity that is common to all such large space mission consortia as well as an insider view of ‘coopetition’[2] management and its regulation. Then lessons learned from PACS management enlighten the industry—instrument consortium—and ESA relationship and more generally discuss the concept of ‘sociality’.[3] Finally, the SPIRE consortium description exhibits

[1]Many concepts used in this chapter were defined in Chap. 5.

[2]As defined in Chap. 5, this term has been coined by specialists in game theory to designate a strategy which combines cooperation and competition.

[3]As defined in Chap. 5, sociality is different from solidarity which applies only when people share the same values and norms. It implies in particular the development of a form of ‘dialogical communication’ instead of the ‘dialectical communication’.

the common success-oriented approach of each instrument consortium and the subtle differences in the scientific organisation of the observing programmes.

10.2.1 Genesis of a Consortium Instrument: The Case of HIFI

ESA has, since the seventies, implemented a well-defined procedure for the selection of their next space projects and the associated scientific instruments, normally funded by the national agencies. Following an announcement and call for new ideas and missions, there is a competitive phase in which submitted proposals are evaluated with a few selected for further study and development. This bottom-up approach was also the case in 1982 when FIRST was selected for assessment and subsequently for a phase-A study. In 1984 ESA decided to embark on a long-term vision definition process, incorporating a broad representation of the scientific community. This ultimately led to a programme, called Horizon 2000 (see Chap. 2), with four cornerstones and about 6–8 medium-sized missions. One of the cornerstone missions was FIRST, labelled as the 'high-throughput FIR spectroscopic mission'. The long-term European vision programme enabled national space agencies to fund new instrumentation developments with a high degree of confidence that the investments in R&D would not be wasted, and the results would be applied in a future ESA space mission.

Around that time, the field of astrochemistry was born and evolved in parallel with the availability of instruments on ground-based telescopes. Observations started to guide and constrain the astrophysical/chemical models. The need for observations at higher frequencies drove the technical development in that direction, and to higher sensitivities, and led to calls for space-borne observatories. From 1975, ESA took a role as facilitator in the development of far-infrared detection techniques using its Technology Research Program (TRP). Starting in 1980, it issued an invitation to tender for the development of heterodyne components for frequencies up to 1.5 THz. A consortium of interested laboratories was formed, and for more than ten years there was a collaborative effort to develop sensitive THz receivers. This collaboration also turned out to be a very successful community-building mechanism. The participants became aware of each other's specific interests and qualities, and this group of collaborating laboratories ultimately became the nucleus of a consortium for the development and construction of the heterodyne instrument for FIRST (HIFI). This instrument, initially called HET in the model payload, was renamed in the proposal as HIFI (see Chap. 8). All of the participating technical groups were located in astronomical institutes with mm/submm astronomers, and all had access to ground-based mm-wave telescopes like SEST, IRAM or JCMT (see Chap. 2). Some even had the luxury of access to their own institute-owned telescopes where they could deploy their most recently developed receivers and instrumentation. For example, the group in Köln owned the KOSMA telescope at Gornergrat, Caltech had the CSO at Mauna Kea, Onsala/Chalmers the SEST, and astronomers in the Netherlands and UK had access to the JCMT (see Chap. 2). This strong instrumentation and astronomy synergy ensured that the HIFI instrument capabilities were strongly science-driven.

By 1996, the FIRST instrumentation working group had already defined a model payload including the HET instrument (see Chap. 3). HET was going to use SIS mixers and operate up to 1.2 THz. Local oscillators were to be made of Gunn oscillators followed by a cascade of varactor diode frequency multipliers. Acousto-optic and digital autocorrelation back-end spectrometers were the obvious choice (see Chap. 8 for details). During 1995 and 1996, the TRP heterodyne group got together several times to discuss developments and designs for the HET instrument. Interaction with their astronomers resulted in several changes of the scientific requirements for HET: a frequency increase up to 2.7 THz, wider IF bandwidths and extending the existing 1.5 GHz bandwidth to 4 GHz. The enhanced scientific requirements increased the opportunities for the instrument team to participate in very challenging developments but required more resources and expert teams. For the mixer[4] teams, the higher-frequency ranges were considered as the most attractive (noble) work packages, and there was a hidden competition about who would work on the highest frequency bands.

Apart from the TRP heterodyne community, there was widespread interest in the European and US communities in participating in the HIFI instrument development, including its construction and science operations. However, several other conditions needed to be fulfilled in order to have a consortium fully capable and suitable to produce HIFI: interest in the technical development, technical ability and capability to produce the item, national and/or local funding, priorities in the national and local interest in the science leading to competition in funding between *Herschel*/*Planck* instruments, methodology experience and fair geographical distribution of participating laboratories. It needed much negotiation and many frank discussions to ensure that the work packages taken up by the various teams fitted their competencies. There were several cases in which the available funding was limited or precluded participation. For example, there is no UK participation in HIFI although there were several interested groups (RAL, Cambridge), which participated in the 1995–1996 discussions. For the choice of institute to build the acousto-optic spectrometer, between Observatoire de Paris (DEMIRM, later known as LERMA) and KOSMA, the limited budget and the geographical distribution played an important moderation role. Nevertheless, in the end the HIFI consortium had the embarrassing riches of several candidate teams for the same work packages.

In order to come to a rational decision and choose between competing options, an agreed procedure was established. It included a way to present all the technical

[4]A mixer, or frequency mixer, is a nonlinear electrical circuit that creates new frequencies from two signals applied to it. In its most common application, two signals at frequencies f_1 and f_2 are applied to a mixer, and it produces new signals at the sum $f_1 + f_2$ and difference $f_1 - f_2$ of the original frequencies. See Chap. 8.

contributions from all the consortium members in a transparent way and to show what was to be expected of each contributing institute. Each member had to submit a 'Performance and Implementation Plan'. This document had to show a detailed implementation plan with the technical solutions, methods, available resources, etc. In the end there was only one area where there remained competition: provision of the digital spectrometer, offered by both France and Sweden. It was agreed to have a 'shoot-out' demonstration and comparison after two years of preparation. This took place as planned, and France's HRS was accepted as the high-resolution spectrometer for HIFI. Although taking as much as possible a rational approach, the consortium's carefully crafted collaborations were vulnerable to personnel turnover due to short-term contracts imposed by modern research-funding schemes. As the sociologist and philosopher Richard Sennett notices about institutions staffed by high-qualified individuals from different backgrounds, 'in such chameleon institutions, work on short-term projects has the effect of acid attacking authority, trust and cooperation' (Sennett 2012: 214). For example, the job changes of one of the leads of the local oscillator subsystem, early in the project, led to a period of considerable uncertainty until a suitable replacement was offered by Max Planck Institute for Radio Astronomy (MPIfR) in Bonn, Germany. Once the leadership was taken up by Rolf Guesten (MPIfR), the architectural design of the local oscillator subsystem could be completed with a subdivision into a LO Assembly (LOA), Local Oscillator LO Control Unit (LCU) and a LO Source Unit (LSU). This was the basis for the task division between MPIfR (LOA), Poland (LCU) and Canada (LSU) with important contributions from JPL for the LOA. Canada's participation was a special challenge for HIFI and ESA management. Initially the Canadian community wanted to participate in both missions, *Herschel* and *Planck*. The approval process through the committees of the Canadian Space Agency took so long that the Canadian Joint Committee on Space Astronomy, which had meanwhile changed its composition, had now a majority which was only interested in a participation in *Planck*. Finally, the participation in HIFI was secured by an intervention and a financial contribution from ESA that compensated for Canadian's *Planck* contribution.

During 1996 and 1997, the basic design and resulting upgrades of HIFI were still changing and expanding. For example, a conceptual design of the focal plane unit was needed. One design was developed by radio-frequency engineers from Chalmers in Sweden and from Canada. Another design by TNO/SRON in the Netherlands, however, turned out to be the most suitable to allow a wide and uninterrupted frequency coverage and to include several higher-frequency bands. With so many of the technologies not reaching TRL[5] levels above four, the proposal to develop and construct HIFI looked very challenging. However, the HIFI experts and leadership were convinced it could be built, but only after considerable development efforts and with adequate matching resources. The 'marching orders' for the

[5]Technology Readiness Levels (TRL) are a method of estimating technology maturity of Critical Technology Elements on a scale from 1 to 9 with 9 being the most mature technology.

developments were laid out in the 'HIFI Instrument Development and Verification Plan' (IDVP) as described in Chap. 8.

SRON became the PI institute for HIFI, relying on its experience from PI roles in previous missions (ISO, XMM, Chandra), and aimed to have a solidly structured consortium organisation in which the responsibilities and interfaces were rationalised and simplified as much as possible. This led to an organisation mostly mirroring the instrument block diagram for subsystems shown in Fig. 10.4. Note that the HIFI instrument block diagram is a more sophisticated representation of the basic heterodyne receiver concept as is described in Chap. 8. Despite this approach, the subsystems had still quite a few contributors (see Figs. 10.5 and 10.6). The focal plane unit (FPU) had seven contributors, while the LO subsystem involved five institutes plus a few subcontractors. The two back-end subsystems were simpler with mainly contributors from the same country. For the HRS, CESR in Toulouse was supported by the Observatoire de Bordeaux, and for the wide band spectrometer (WBS), support came from MPSS and from Italy. The main reason for these large sub-consortia was to cope with the sheer volume of work, including budget, and with the wide range of expertise needed.

Although the division of tasks was well described in the HIFI proposal, the stability of this division took a few more years to become established. The poor collaboration between two back-end labs caused a complete split of the HRS and WBS subsystems.[6] Instead of sharing the IF processor, the signal chain was split right after the 4–8 GHz IF amplification chains which left the WBS without IF processor as they were missing the required funding. This led to a search for new parties and to an investigation into whether it could be done with only one set of back-end spectrometers. This was strongly encouraged by ESA, who were anxious to decrease the complexity of HIFI. After the review, which took place at Schiphol airport, the panel confirmed this possibility but also recognized the advantage of having two completely different technologies covering the critical functionality of the spectrometers. So the outcome of the review did not help much to decrease the tension in the co-PI team. In the end the problem was graciously solved by DLR (the German space agency) providing extra funding for the WBS intermediate frequency processor.

Among the other serious issues that occurred in the consortium were the HIFI 're-scope', resulting in the abandonment of band 7 (a 2.7 THz channel that had been strongly promoted by Tom Phillips[7]). The main scientific driver for this channel was to observe the hydrogen deuteride (HD) line at 112 μm wavelength. However, the development of the multiplier chain for this frequency was taking up so much effort that lower frequency bands, also still requiring much development work,

[6]High resolution spectrometer (HRS) and wide band spectrometer (WBS).

[7]Thomas G. Phillips is a Professor of Physics at Caltech and one of the pioneers of submillimetre astronomy. In 1982 he became Director Designate for the Caltech Submillimeter Observatory, to be constructed in Hawaii and, in 1986, on successful completion of the construction, became Director. He is also the PI of the US contribution to the HIFI instrument for *Herschel* as a final result of his work as PI of LDR and the Submillimeter Explorer.

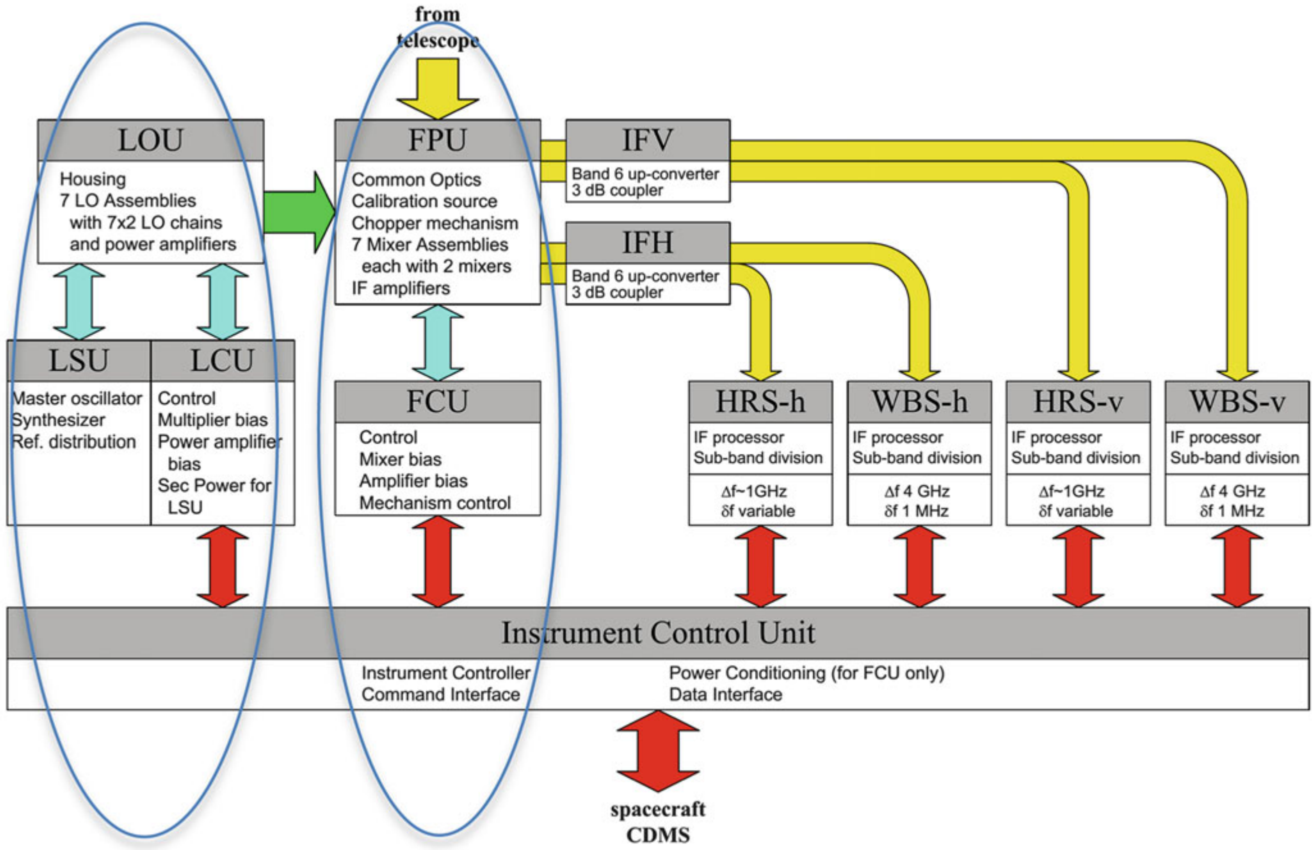


Fig. 10.4 HIFI subsystem hardware block diagram. The *circles* show the subsystem organisation units for the LO and focal plane unit subsystems. The two HRS units and the two WBS units, each were independent subsystems. The Instrument Control Unit was the fifth subsystem for HIFI, while it was also a subsystem for all the other *Herschel* instruments. © ESA/*Herschel*/HIFI consortium

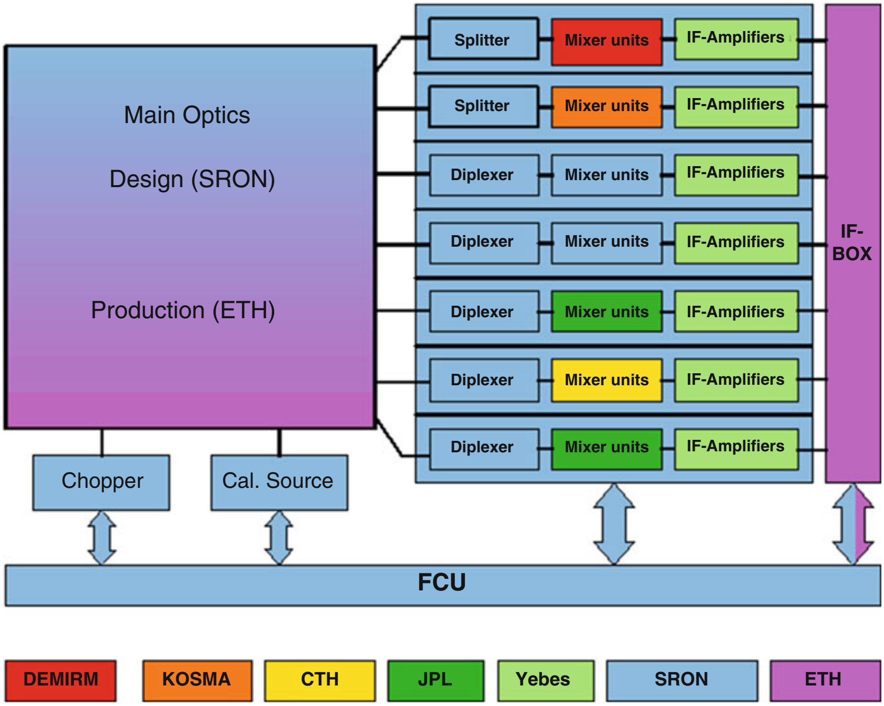


Fig. 10.5 Organisation of the focal plane unit subsystem. *Colours* indicate the organisation and which components it provided. SRON was in the lead of this subsystem. © ESA/*Herschel*/HIFI consortium

were suffering. After a review that took place at JPL, it was decided that the resources for band 7 would go to band 6—not only its finances but also its 'real-estate' (volume, mass, power) within the focal plane unit. Another shift occurred when the provision of the band-6 mixers was partly moved from JPL to the Chalmers group in Sweden, which had developed superior hot-electron bolometer mixers. This eased the financial situation at JPL, again in favour of the LO chains development.

In the end 23 institutes from 11 countries contributed to HIFI: the Netherlands, USA, Germany, France, Spain, Italy, Switzerland, Sweden, Poland, Canada and Ireland, with a small contribution from Taiwan. The HIFI consortium was indeed very large (Fig. 10.7). At the top level were the interactions with ESA and the ESA contractors. Communication management became of highest importance for the project, which involved over 300 people around 2004–2005. The HIFI consortium meetings twice a year became *the* place for exchanging information among consortium members. HIFI's top-level governance came from the HIFI PI, the HIFI project manager and the PI team consisting of the PI, the three Co-PIs and the project scientists. The HIFI Steering Committee consisted of the PI, the co-PIs and lead co-Is, one for each of the participating countries. Its task was to support the PI

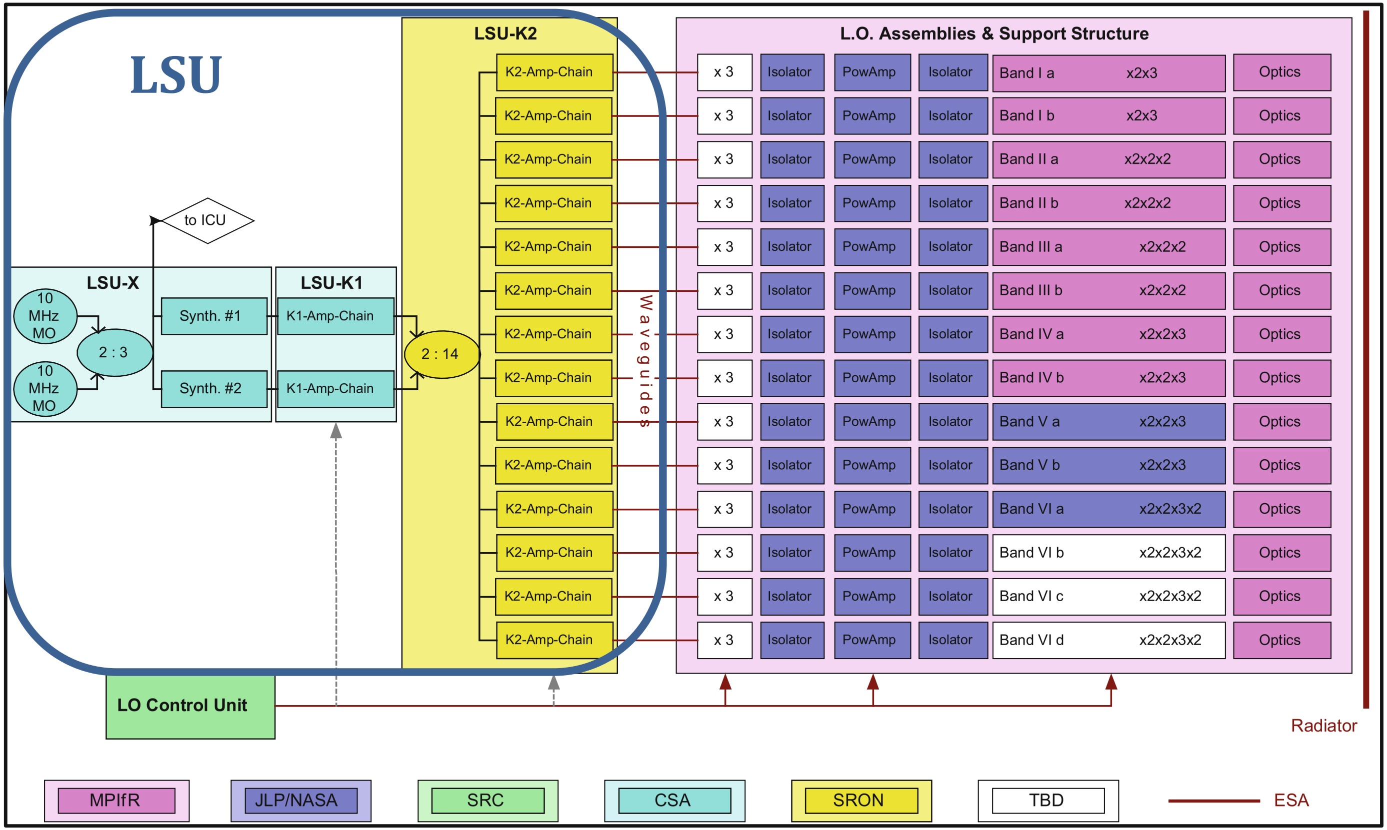


Fig. 10.6 Organisation of the local oscillator subsystem. *Colours* indicate the organisation and components provided. MPIfR was in the lead of this subsystem. Note: all band VI multiplier chains were provided by JPL. © ESA/*Herschel*/HIFI consortium

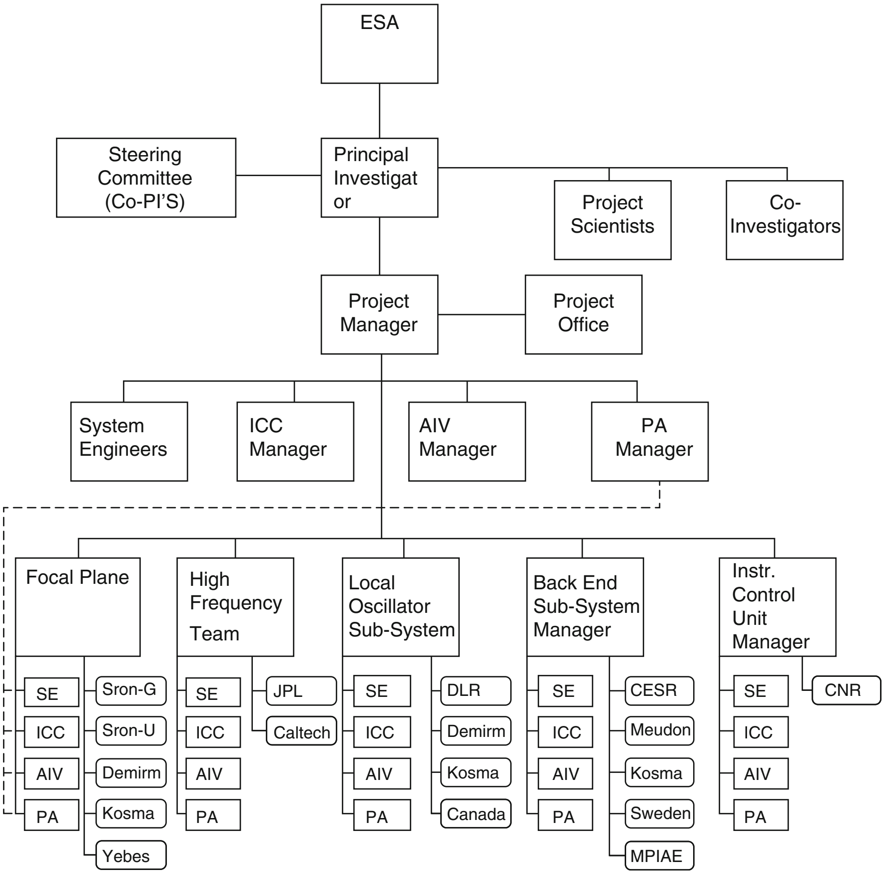


Fig. 10.7 Organigram of the HIFI hardware consortium (taken from the HIFI proposal). © ESA/*Herschel*/HIFI consortium

on policy issues, work distribution and funding issues. The members of the steering committee were also the national point of contact to their national space-funding agency and for the distribution of the guaranteed science time. In HIFI these funding agencies were kept outside the steering committee but were welcomed to participate in the meetings as observer. This approach is probably one of the last cases with a steering committee with non-funding agency memberships only. While the idea of steering committees to keep the link between science and funding agencies for nationally funded instruments should be fostered, by now, all space instruments are formally governed by steering committees dominated by funding agencies where risk avoidance has high priority. This evolution of the management appears at least problematic, and its implications for the future spatial missions are discussed in the conclusion of this book.

10.2.2 Science Organisation Within the HIFI Consortium

Overall there were 127 scientists associated with HIFI as co-investigators (co-Is), almost all from the hardware- or software-contributing institutes. Their co-I status was left to the leads of the contributing institutes, and their number was expected to be in proportion. Some were invited because of their special expertise. The regulation of the number of co-I's was self-limiting as the data rights per country were to be shared among its scientists. In any case all were interested in observing with HIFI and in participating in one of the guaranteed time programs. HIFI choose to devote almost all of its guaranteed time to Key Programmes (i.e. programmes using more than 100 h of observation). The science consortium therefore organised itself in various Key Programme teams. The construction of the Key Programmes was led by the HIFI project scientist, Xander Tielens, in close collaboration with the PI team and the steering committee. They took care that there was a balance between the science topics and that the most important hardware-contributing partners got leading positions in the science themes and/or key programmes, which sometimes was not so easy, as the distribution of participants over the key programmes did not follow nationality.

In order to manage all the science matters, the HIFI PI team (the PI, the three Co-PIs and the project scientist) wrote a document that described the rules of engagement in the science exploitation of the HIFI guaranteed time. This HIFI policy was based on the creation of a HIFI science core team, consisting of the PI team and the PI's of the HIFI Key Programmes. Its duties were to coordinate and direct the activities preparing HIFI observations and data interpretation, which included preparatory work on laboratory experiments, theoretical modelling, calibration, special HIFI data analysis software, etc.; to encourage and foster collaboration between the consortium members; to coordinate between the HIFI Key Programme teams and support coordination with the PACS and SPIRE teams, the *Herschel* mission scientists and the *Herschel* Science Centre team; to provide input and comment for the HIFI web pages; and to contribute to other PR activities. In order to constrain the proposals, all members of the steering committee (national representatives) got a number of hours based on their hardware contribution. In a special HIFI meeting, they distributed their shares over the Key Programmes in which they and their co-Is were involved. Before the launch it became known that HIFI only had 1894 hours, rather than the expected 2000 hours, and that each HIFI guaranteed time key program had to be cut by ~5%, causing quite some interaction in the consortium. Each co-I generally participated only in one or two key programmes, so as not to spread their efforts and observing time too thinly. Coordination with the other hardware consortia was attempted, but it was clear very soon that the scientific overlap between HIFI and the other instruments was quite small, and coordination was either incorporated immediately or soon forgotten.

The HIFI policy document also mentions the following about authorship on the first papers: 'Institutional as well as individual needs will be considered. The institutions of those partners, which put in guaranteed observing time need to get appropriate credit'. Building HIFI and ensuring its functioning in both hardware and software took many people away from science for years. The steering committee members and the PI assigned people to one or more first release papers according to their investment. This was especially important for young people, since their careers depended on the number of papers they published.

Another important matter was the data calibration and analysis. The preparation of the data analysis capabilities was mentioned up front in the HIFI policy document (October 2005). From experience with similar ground-based data, it was clear that reducing HIFI data itself could be fairly easy. However, the interpretation would be far from simple and straightforward. Fundamental data, such as collisional excitation numbers and line frequencies, were generally missing. As these data were indispensable, several preparatory workshops were organized by the HIFI teams to reach out to all *Herschel* observers including the PACS and SPIRE consortia and the *Herschel* Science Team members. These workshops were held in the Leiden Lorentz Centre. Experts and specialists from theory and lab experiments were invited and took part. They were very interested as they saw a new group of customers and applications of their specialised effort emerging. Besides ensuring a well-prepared *Herschel* spectroscopy community for data analysis, the workshops had a community-building effect, which assisted the communications and coordination between and inside the consortium.

On the instrument calibration front, HIFI decided early on to base the science ground segment in Groningen but operate it in a distributed fashion to keep travel costs in the consortium as low as possible. To have the ICC staff located at their PI institutes helped to effect a 'smooth transition' between development and operations. The concept supported the interaction between the people in the ICC and in the instrument development, but communication across the partners still had to rely heavily on electronic communication (email and wikis), teleconferences and the consortium meetings with their various splinters. Usually these were held in Groningen. The ICC started out with software engineers to define the requirements and to make the first architectural designs. It was soon after recognized that coding in Java was a specialist job, and qualified programmers became part of the ICC. During the development the software and pipelines evolved, but not always at the same pace as did the instrument. This mismatch led to some tension, when results from instrument-level tests (ILT) could not be analysed or displayed online or in real time. The database generated in the ILTs was subsequently used extensively when the HIFI prime instrument had a failure. Later on, a fast turnaround could be achieved and was available for the commissioning.

During commissioning and early operations, the collaboration between consortium members and HIFI ICC was intensified. PhD students and postdocs from the consortium came to Groningen to help with data reduction, which fed into analysis of the instrument data. A big side-effect of this was that they became highly trained in the data reduction software and took that expertise back to their home institutes. During mission operations the ICC continued to work with many teleconferences, co-locations and various splinter meetings dedicated to science ground segment aspects, and this continued into post-operations until March 2016.

10.2.3 Herschel*, Industry and Innovation: Lessons Learned from the PACS Consortium*

The previous section described the emerging phases of a consortium and its organisation for science operation and return through the study case of the HIFI consortium. It emphasized first the needs for organised competition and cooperation (i.e. coopetition) to optimise the instrument design and secondly the management organisation to facilitate the instrument development and to optimise the key programme selection and scientific return. The description focused on issues internal to the HIFI consortium. Within this section, we intend to analyse, through the case of the PACS team, the relationship between a consortium team, ESA and industry, focusing on the concept of 'sociality' introduced in Chap. 5. This sociological analysis, including coopetition and sociality concepts and confronting them with the management description, identifies social characteristics of the innovation process within the *Herschel* mission.

PACS stands for a Photodetector Array Camera and Spectrometer. PACS has been designed and built by a consortium of institutes and university departments from across Europe under the leadership of Principal Investigator Albrecht Poglitsch located at Max Planck Institute for Extraterrestrial Physics, Garching, Germany. Twenty research institutes within six countries were members of the PACS consortium: Austria (UVIE), Belgium (IMEC, KUL, CSL), France (CEA, OAMP), Germany (MPE, MPIA), Italy (IFSI, OAP/OAT, OAA/CAISMI, LENS, SISSA) and Spain (IAC) as shown in Fig. 10.8. In that respect, the PACS instrument genesis and team structure resemble those of HIFI or SPIRE. All three consortia were science driven and consisted of many science institutes in Europe (and beyond in the cases of HIFI and SPIRE) having the capability of building instruments and exploiting them for astronomy. As written in Poglitsch et al. (2010), 'The PACS instrument was designed as a general-purpose science instrument covering the wavelength range ~60–210 μm. It features both, a photometric multi-colour imaging mode, and an imaging spectrometer. Both instrument sections were designed with the goal of maximising the science return of the mission, given the constraints of the *Herschel* platform (telescope at $T \approx 85\ K$, diffraction limited for $\lambda > 80$ μm, limited real estate on the cryostat optical bench) and available FIR detector technology'.

The following interview between the authors and Albrecht Poglitsch concentrates on the lessons learned from the PACS consortium management during the instrument realisation, integration into the *Herschel* spacecraft and science return. Let us start with an overview of the mission success perception over 15 years from 2000 to 2015.

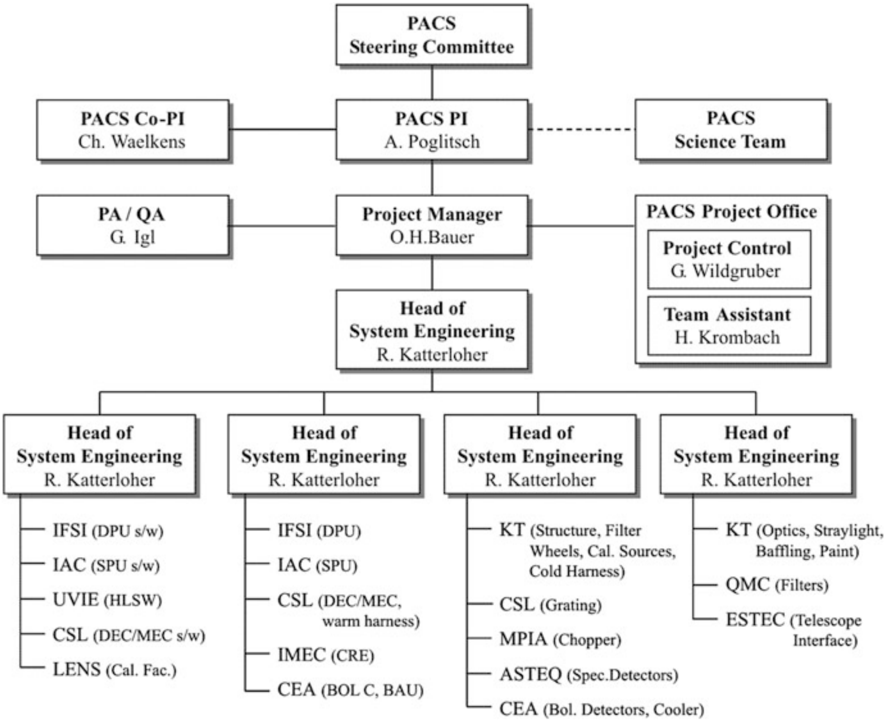


Fig. 10.8 Organigram of the PACS consortium. © ESA/*Herschel*/OACS consortium

Authors Nowadays the history of the *Herschel* mission is presented as a success story by ESA. Was it a straightforward process, or did such a complex mission trigger any tension?

AP After a number of years of operation, analysis and exploitation, when comparing *Herschel* with other space missions and judging it by the range of publications and the rate of citations, it looks very respectable (see Chap. 4 for details). This is a great success and we may conclude that we did the right things. However, the process was not so straightforward. Originally *Herschel* and *Planck* had respectively been planned as a Cornerstone mission and a medium-size mission. Then there was the ad hoc decision to introduce the Mars Express into the ESA program and it, of course, costs money. So, to achieve both *Herschel* and *Planck*, they had to be done for a lower amount of funding: instead of having a budget equivalent to a cornerstone and a medium-size mission, we were supposed to do it both together for the cost of a large mission minus 10% (see Chap. 3 for details). There were various creative solutions, including launching them together or merging them into one mission. To make it cheaper, it had to be done faster, because in this type of project the cost scales with the duration. The schedule was compressed to a level that was simply impossible to follow.

Then some people took it literally, and it was also at the time the science director changed, i.e. when Roger-Maurice Bonnet's turn was over and David Southwood came in, for whom *Herschel* and *Planck* were mostly a concern. They were not on schedule and obviously overrunning in cost. The *Herschel* and *Planck* programme was perceived as a threat to the science programme of ESA. The *Herschel* and *Planck* teams had to find a way of working together. Then, when the next science director [Alvaro Giménez Cañete] came in while *Herschel* and *Planck* deliver their results, he was in the comfortable position of not having been involved in all of the previous crises, and *Herschel* and *Planck* were now presented as 'two textbook examples'. *Herschel* was implemented for the cost of a large mission; *Planck* was implemented for the cost of a medium mission. Success depends on the metrics you apply: if one expects to do both for the price of one, then of course it might be seen as a disaster, but if one expects them to cost what they were supposed to cost at the beginning, then it looks like a great success from the 'design-to-cost' point of view. Since it has been a scientific success, people tend to forget some of the other aspects that happened before.

Remark While the importance of designing to real cost appears as one of the key markers of the success of a space mission for all involved social organisations (ESA, industry, science consortia), A. Poglitsch continues and comments on how industry interacted with the instrument consortia and its management model. 'Sociality' as a dialogical communication between industry engineers and instrument scientists was maintained and mediated by a third person, the ESA Project Manager.

Authors What was the relationship between industry and the instrument consortia? And what was the benefit to industry to be involved in such a mission?

AP Industry was supposed to manage the interfaces between instrument and satellite, and the satellite was built by the very same companies. They had no reason to treat the instrument side favourably. Whenever there was a difficulty, of course, their attitude was that the scientists had to do something about it because they could not change any industrial parts. Industry does not make so much money in a programme such as *Herschel-Planck*, so they need to prevent any possible extra cost. Whenever there is a request from the industry or from the instrument, a neutral arbiter who is fair to both was needed, and only ESA was in such a position.

It soon became clear that this tension was very real sometimes. The Project Manager [Thomas Passvogel] had a very important role to play. He had to maintain the pressure on the payload and spacecraft providers to deliver on schedule, but on the other hand, he had a deep understanding of what was possible and what was not and of what was necessary and what was not. Also he understood how different parties [industry, research laboratories, etc.] worked and how you had to communicate with them to move forward. For example, with industry one can use contractual arguments, but with the instrument teams, this did not work the same way, and science requirement arguments were more powerful. The important point was communication. The Project Manager had the right combination of understanding the scientific side but also the habits of dealing with the 'real world'.

The third part of the dialogue concerns the concept of innovation, its industrial and managerial conditions.

Authors Was there a common culture of breakthrough innovation between industry and instrument consortia?

AP Industry had to use systems that were demonstrated technologies, except for the telescope story (see details in Chap. 6). For example, the cryostat was essentially ISO's cryostat. One needs to distinguish between the main satellite (the spacecraft) on one hand and the telescope and the instruments on the other. The satellite reused the building blocks of ISO and XMM and put them together again. All of that existed. Most of the novelty and associated risk was in the instruments. At the time, we did not even know what a Technology Readiness Level[8] (TRL) was. The construction of the instrument was a wild mix between work done inside research institutes and work done by industrial contractors. It has also to do with the funding model, with the fact that the instruments were financed by the national agencies, not by the ESA. Since the funding was coming from the national agencies, they wanted to have influence over where the money went and wanted to be involved in these decisions, and we had sometimes to debate with them about the choice of a particular contractor.

Nowadays, the science instruments are no longer allowed to work along a 'success-orientated' approach, but they have to comply with the TRL business. This induces a big delay because one has to demonstrate that a technology is working before it is allowed to be considered for a space mission. There is now a very extensive study phase in order not to run into problems later on. ESA studies several mission candidates in parallel and then selects at the end the one that has the least risk of failing in operation. Of course, the national agencies have figured out that it will cost them a lot more because one has to study several in parallel, but only one will be selected, so basically they are going to pay a lot for things that will never fly. The risk reduction will be funded by them.

Authors With the TRLs, would a very innovative mission like *Herschel* be possible today?

AP No. That is the very simple answer: no, absolutely no. Many of instrument components did not even exist, not even in a lab, when the mission was selected. There was a sort of faith and determination that under the pressure and with success, we would get enough enthusiasm and funding to make it work. Indeed, the whole concept of TRL was not applicable to what we did. The telescope did not exist either. For PACS, we did not really know how to build the detector arrays, nor the

[8]Technology Readiness Levels (TRL) are a method of estimating technology maturity of Critical Technology Elements.

read-out for the spectrometer. Of course, the optics for our complicated instrument had not been rigorously demonstrated. It was based on a sort of intuitive design.

Authors During the design and construction phases, did you observe ‘last-minute innovations’ or ‘sudden death of a technology’?

AP For the PACS instrument consisting of a photometer and a spectrometer, we had optimised our detectors for spectroscopy. There was a problem when we wanted to use the same detectors then for the broadband photometric mode, which involved the detectors receiving something like a thousand times more light. All involved parties, including myself, had been over-optimistic about the dynamic range that the read-out could accept and, with the same read-out at least, the same detectors could not make it. We were really in a dead end. Then, at the same time, CEA[9] had developed a detector technology for the SPIRE photometer camera, which, however, was not found mature enough or optimal for this instrument. The CEA team was relatively sad after losing the SPIRE selection, and Suzanne Madden, who knew that I worked on another instrument for SOFIA (another flying observatory), convinced Laurent Vigroux to call me up. In these days we were still confident that SOFIA would fly relatively soon, and we could not change our SOFIA instrument anymore at that stage. However, we had to do something serious about the PACS photometer; otherwise it would have been a failure. If the CEA detectors, which had been developed for the broadband mode in SPIRE, could be adopted in our instrument and if we could tune them to our wavelengths, they might be a perfect application for the CEA expensive development. We had two choices: a guaranteed failure or at least a certain chance for success. Everybody worked very hard and very fast to make it work, and in the end it was the ideal outcome for both parties because we had the most sensitive spectrometer and a very well working photometer and science would be done. A potential disaster turned out into a sort of a dream come true for several parties.

Authors What has been learned from the experience out of the *Herschel* mission?

AP The Horizon 2000 program was good, in many ways, because it was determined by setting our scientific priorities, with some realism about what could be done but also with the idea that if you wanted to do something new and interesting, one may need a lot of development along the way. Bonnet said exactly the same thing as what we concluded in our PACS consortium: if we had been forced to go by the TRL, it could have been only a very mediocre instrument at best. The approach ‘we want to do our science and we make whatever effort is necessary to do it’ was a good approach and for sure a better approach than what is done now. We have shown with five instruments, between *Herschel* and *Planck*, which it could be done right for each for them. So it is not a singular lucky success; in these five successful instruments, there is clear statistical evidence that this kind of risky approach led to

[9]Commissariat à l’Energie Atomique et aux Energies Alternatives in France was already engaged in the SPIRE consortium.

success and success which would have not been possible if you had gone only by playing it safe. So, it should have not been too hard to learn that lesson.

10.2.4 Confronting the Management Organisation in HIFI, PACS and SPIRE

In the previous sections, many features have emerged from the general description of the HIFI consortium and from specific focuses on the PACS consortium: international organisation; complex organisation associating scientific, technical and industrial responsibilities; coopetition with clear rules to select the instruments and the observing programmes; return on investment; well-prepared instrument calibration; and needs for constant communication within each consortium. But also the fact that each consortium has to be ready to face some crises and that sociality is built during these crises since (as long as everybody works hard to overcome it) authority is increased for the managers, mutual respect is acquired between the different groups and coopetition is reinforced. The international and complex characteristics are common to any large 'big science' instrument consortium such as in astrophysics (e.g. XMM-*Newton*, Rosetta in space, ALMA on the ground) or in particle physics (e.g. the Large Hadron Collider instruments) and more generally to any large-scale technical work. Our analysis reveals, using the example of HIFI, that coopetition is important in the decision-making, and, using PACS as an example, that sociality is an important foundation for success. It might then reveal generic features of the scientific field when designing a complex scientific instrument. Coopetition and sociality regulate the complex collaboration between participating members of a consortium. They require common rules (e.g. the HIFI policy document) and a dialogical communication between science and industry in order to set the trust of fair scientific or technological return for each participant.

Let us now enter the SPIRE consortium and investigate whether coopetition and sociality were also at play. The SPIRE consortium was formed in the late 1990s in preparation for submission of a proposal to ESA to build a long-wavelength camera and spectrometer for FIRST. It included scientists, engineers and managers from five European countries (France, Italy, Spain, Sweden, UK) as well as the USA and Canada. The SPIRE organisation chart resembles those of HIFI and PACS (Figs. 10.7, 10.8 and 10.9). They present similarities as all consortia were coordinated by a principal investigator and co-PI(s) and were ruled according to clear policy documents. By agreement between the leading participants from each institute, a SPIRE Constitution was formulated setting out the general responsibilities and rights of all participants, and a more detailed management plan was written to describe the various tasks and management practices and procedures that the consortium would adopt. The consortium was led by the Principal Investigator, Prof. Matt Griffin (then of Queen Mary, University of London, subsequently of

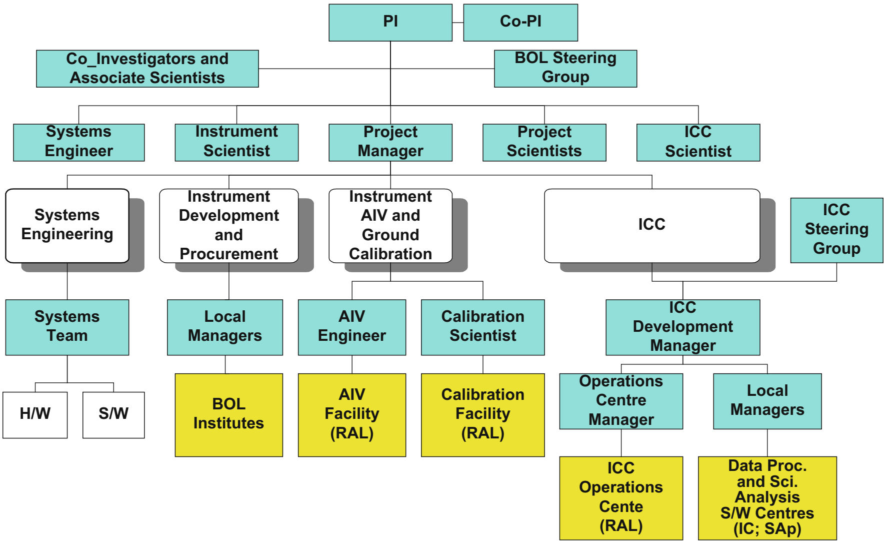


Fig. 10.9 Organigram of the SPIRE consortium (note that at the time, early in the project, when this was devised, the instrument was known as 'BOL'). © ESA/*Herschel*/SPIRE consortium

Cardiff University) and the Co-Principal Investigator, Dr Laurent Vigroux, of CEA, Saclay (see Fig. 10.9). This joint leadership reflected the major financial contributions of the UK and France to the project.

At lower levels in the organisation charts, instrument consortia differ in the responsibilities expected and given to the scientific leaders and co-investigators. In SPIRE, a single co-investigator was appointed to represent the scientific interests and the technical programme of each participating institute. At a later stage, the National Astronomical Observatory of China (NAOC) in Beijing joined the project, and a Chinese Co-I was also appointed. The co-investigators were the main signatories of the proposal submitted to ESA, and they were the individuals with formal responsibility for delivering their institutes' contributions to SPIRE. Such contributions were of two forms: (i) the provision of instrument hardware and (ii) participation in the SPIRE Instrument Control Centre (ICC). ICC work included all aspects of instrument command and control and data processing and the ICCs for all three instruments participated in the joint endeavour of the *Herschel* Common Science System (HCSS) together with the ESA *Herschel* Science Centre (HSC). One of the main reasons for *Herschel*'s scientific success was the recognition that careful and thorough preparation for instrument operations and data processing would be essential, and this motivated the significant investment that the instrument teams and ESA made in the ICCs and the HSC, right from the start of the project. In the case of SPIRE, for instance, ICC members worked closely with hardware designers and builders to develop and implement the systems used for instrument commanding and testing. During the ground-test campaigns, ICC and hardware specialists acted as integrated teams bringing together all necessary expertise to plan, execute and interpret the tests. Data pipeline working groups were formed early and worked to develop and test data processing software for the spectrometer and camera prior to launch and post-launch, continuously improving and enhancing the data quality.

One co-investigator from each participating country was also identified as the national representative, with the task of liaising with that country's funding agency to assist in resolving any significant financial or programmatic problems that might affect SPIRE. In the SPIRE proposal, an outline was presented of the contributions that each country and institute was envisaged to make. Subsequently, after the project started, some adjustments were made to take financial and technical factors into account. For example, Canada was originally foreseen to provide a shutter mechanism for the instrument, but it was later decided not to have a shutter as it was technically not essential and could be a single-point failure for SPIRE. An alternative Canadian contribution was therefore defined, including support of the instrument ground testing and the development of software for the SPIRE spectrometer. In addition, at the time of the proposal, the final detector technology for SPIRE had not yet been selected, and three institutes were working on developing different options (see Chap. 7). The overall technical direction of SPIRE was carried out by the SPIRE project team, which included the PI, project managers for the hardware and ICC, the instrument scientist and the system engineer. During the design, construction and testing of SPIRE, the technical programme of each participating institute was carried out under the direction of an institute project manager, reporting to the SPIRE hardware or ICC project manager.

On the scientific side, the SPIRE Constitution laid out the principles for how the consortium would use its guaranteed time. The SPIRE Science Team was identified as the group of scientists who would be involved in the definition of the SPIRE guaranteed time (GT) programme and its scientific exploitation. It comprised scientists, from within the SPIRE Consortium or invited from outside, who contributed to the work of the consortium and wished to use SPIRE GT or open time data for scientific research. Three membership categories were defined: co-investigator, associate scientist and consultant, which recognised the role of scientists involved in the project from scientific leader to PhD student. The co-investigators were the formal owners (in the custodial sense) of SPIRE GT observational data. The science consortium team was divided into six Specialist Astronomy Groups (SAGs), each one responsible for a particular scientific area (e.g. SAG 3 covered star formation in our galaxy). Each SAG had two coordinators to organise the work and share the workload, and produced a proposal with an observing time request. The co-Is acted as an internal SPIRE time allocation committee and distributed the available time among the SAGs. Each SAG then formulated its detailed programme fitting within its time allocation. Co-I and Associate Scientist members of the SAG then had equal rights to the data. The SAG coordinators were responsible for optimising SAG activity and implementing publication policies fairly.

When comparing with HIFI and PACS, the scientific organisation in SPIRE appears 'bottom up' rather than 'top down'. The lead Co-Is coordinating the six

SAGs defined the use of the guaranteed time in key programmes collectively. Non-co-Is (associate scientists, consultants) were involved in the process. Then each SAG organised its allocated time for key programmes, and the SAG appeared as the main scientific decision entity. In HIFI, the number of co-Is for a given institute was proportional to its investment. The construction of the key programmes, therefore, the scientific legacy of the consortium, was led by the HIFI project scientist in close collaboration with the PI team and the steering committee (including the lead co-Is). Here the top organisation was the main scientific decision framework. One might say that the science-driven approach in each consortium was top driven in the case of HIFI and collectively driven in the case of SPIRE. However, these differences are quite subtle and probably refer to some specific 'cultural approach' of each scientific/instrument community (Sect. 4.5). In PACS, the initial investment of a participating institute was converted in observing guaranteed time at the disposal of each co-I institute to run a PACS only observing programme or bring it to a common PACS-SPIRE programme. Some of the SPIRE-led key programmes then benefited from Co-Is' PACS time when planning observations using both instruments. However, two different sets of rules applied to the data whether retrieved with SPIRE or PACS. In summary, scientific debate within the SPIRE SAG has led the organisation of Key Programmes, while institute ownership applied to PACS guaranteed time and top-level consortium organisation prevailed on distributing and using HIFI time.

Communication and discussion within the consortia were fundamental to maintain the success and science-oriented organisation. Besides innumerable small meetings and video/teleconferences to pursue the development of the instrument and ICC, the SPIRE consortium held full consortium meetings roughly every year, with meetings held in rotation in the different participating countries. At such meetings the overall status of the project would be reviewed, as would the consortium's plan for using its allocation of GT. ESA also convened a series of major reviews for all of the *Herschel* instruments, also held at approximately one-year intervals. These reviews constituted an organised framework for making and monitoring progress and in general worked very well. This was because they were honestly conducted and thorough, with conscientious engagement from the instruments, ESA and industry. Problem areas were highlighted rather than hidden, with the aim of identifying them clearly and putting in place plans to address them—in keeping with the principles outlined by Albrecht Poglitsch in the interview above, whereby the instrument teams were driven by their scientific ambitions and determination to confront technical challenges and risks and solve them in the course of the development programme. The sequence of reviews also constituted an extremely effective management tool for the instrument teams: they helped focus attention on key issues, they were invaluable in sorting out documentation, and they helped convey key messages to the teams and to the national funding agencies.

Early on in the project, the standard reporting mechanisms (monthly paper reports at all levels) were adopted but were soon found to be inadequate. Due to the large consortium size and the complexity of the project, information was inevitably incomplete, of variable quality and out of date, and there was a consequent general lack of clarity and control at all levels. Eventually a more flexible and responsive system was adopted with regular structured and 'minuted' teleconferences becoming the main reporting forum, and this was much more effective, more disciplined, faster and more reliable with a natural emphasis on identifying and starting to address problem areas in time.

The initial development plan for SPIRE had to be adapted and revised a number of times to take into account technical problems and delays. Such revisions required careful consultation with all concerned, considering the complex and interlinked nature of the programme. Engagement with the project by the national agencies was poor in the early years, and this was a problem for all of the *Herschel* and *Planck* instruments. No specific forum for national agency engagement had been set up, and the project was complex and continued to evolve in a way that bypassed the visibility of the agencies. This lack of connectivity and up-to-date information led to poor understanding of problems which needed attention at agency level. The solution finally adopted, to hold quarterly face-to-face working meetings involving the ESA project team, the instrument teams and national agency representatives, was immensely beneficial in generating common understanding, maintaining momentum and common focus on priorities and ensuring that all stakeholders understood what they needed to do.

The SPIRE, PACS and HIFI consortia were large by necessity; no single country had all of the expertise or resources needed to build and operate these instruments. Despite the inevitable managerial challenges and difficulties, the experience of *Herschel* instrument management shows that such large teams can be made to work effectively, with subtle differences in their organisation depending on the historical and cultural background of each scientific/instrument community. One of the major problems, not unique to large teams, but certainly not helped by the operation of a big consortium, is schedule credibility and schedule slippage. Everyone expected someone else to be late, and that became the default assumption for their own schedule planning. Firmly monitored and managed 'success-orientated' schedules were necessary to maintain schedule discipline. Coopetition regulated by specific scientific community rules and sociality triggered by meetings and routines of dialogical communication has certainly favoured the success of each instrument consortium.

10.3 Management of Data Products

The *Herschel* Space Observatory was an ESA facility available for the worldwide astronomical community. It was designed to provide three years of routine science operations consisting of guaranteed and open time (see Chap. 4). The GT fraction was 32% and the remainder was OT. The GT was owned by contributors to the *Herschel* mission, mainly the science payload consortia. The OT was awarded in a standard competitive proposal procedure, reviewed by the *Herschel* Observing Time Allocation Committee (HOTAC) in response to Announcements of

Opportunity (AO) issued by the *Herschel* Science Centre (HSC) managed by the *Herschel* Project Scientist on behalf of the Director of Science and Robotic Exploration. The initial Announcement of Opportunity (AO) process for *Herschel* observing time took place in February 2007 to February 2008 and concerned Key Programme (KP) proposals only, which were large programmes in terms of requested observing hours (>100 h). The three instrument teams decided to use up nearly all of their GT in the form of Key Programmes at this stage, keeping very little back for the later stages of the mission. The reasons for this were twofold: first, the teams tended to believe that large organised programmes were best suited to *Herschel* science and, secondly, the rules meant that if they kept time and science programmes in reserve, they could be scooped by OT proposers whose turn it was next.

The KP guaranteed time (KPGT) and open time (KPOT) programmes represented approximately 5500 h of observing time each, constituting 11,250 h in total or 57% of the available *Herschel* routine mission science nominal time. It is interesting to note that the instrument use share differs from HIFI to SPIRE. PACS was the most popular instrument in the KPs (see Fig. 10.10). This is a major departure from the original FIRST concept presented in the Horizon 2000 programme as a spectroscopic mission with millimetre-wave radio astronomy at the forefront (see Chap. 2). Moreover, the Key Programmes show that *Herschel* science was dominated by infrared astronomy rather than millimetre radio astronomy (Fig. 10.10). As written in Chap. 4, the share between spectroscopy and photometry at all wavelengths was about 53–47%. *Herschel* has extended the astronomical objectives of ISO to longer wavelengths as much as it has opened the submillimetre domain to molecular spectroscopy with HIFI.

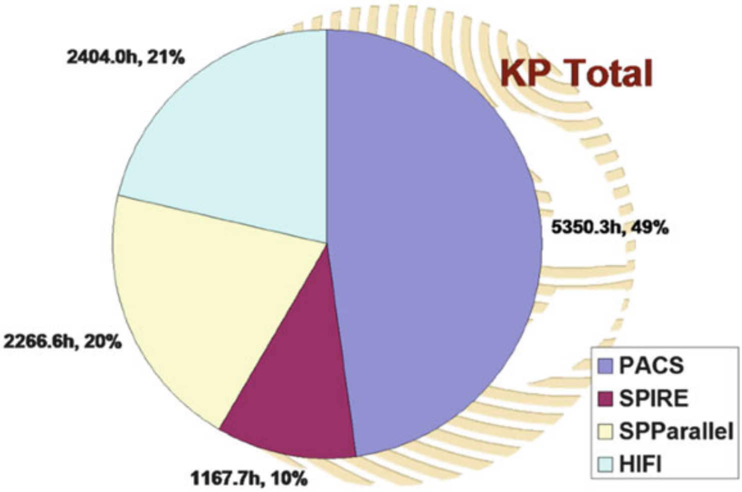


Fig. 10.10 Time allocation for Key Programmes relative to the instrument. SPP parallel mode involved SPIRE and PACS operating simultaneously to maximise data collection efficiency. © ESA/*Herschel*

The first in-flight AO process (AO-1) took place between February and December 2010 and comprised guaranteed and open time phases, labelled GT1 and OT1. There are 33 GT1-observing programmes representing about 555 h and 241 OT1-observing programmes representing almost 6600 h of observing time. OT1-observing time was awarded as priority 1 (almost 5000 h) and priority 2 time. The second in-flight AO cycle (AO-2) consisted of an initial guaranteed time phase (GT2), followed by an open time (OT2) phase.

Among all the ground segment (Sect. 10.1) procedures and operations, the top priority was scheduling observations as efficiently as possible for the liquid helium consumption (see Chap. 9) and hence for making the best possible use of the lifetime of the cooled instruments. Scheduling cycles contained sequences of instrument usage, adhering to various constraints and reflecting the nature and amounts of observations available for the part of the sky currently visible. There were no specified priority differences between the various Key Programmes (KPs); however, the GT KPs had been available for a longer time and had a higher average completion percentage than the OT KPs did. The *Herschel* mission lifetime was extended by a few months with respect to the 3.5-year expected lifetime. *Herschel* ceased operation in April 2013.

Herschel data products are available in various formats from raw data to calibrated data. Most of them are automatically generated by systematic pipeline processing performed at the *Herschel* Science Centre (HSC) using the *Herschel* data processing system and made available through the *Herschel* Science Archive (HSA). The *Herschel* products consist of observational products (scientific data), auxiliary products (non-science spacecraft data), calibration products (behaviour of the satellite and the instruments), quality-control products (quality of the executed observation and the products generated) and user-provided data products (interactively reduced data provided to the *Herschel* Science Centre by the observers). All observations made in the first year of the routine phase had proprietary times of 12 months, while for all observations made later, the proprietary time was 6 months, with a simple ‘bridging scheme’ so that no observation was made public before observations that were executed earlier became public as well. The proprietary time applied to each observation individually, counted from the day when the data were made available to the data owner. However, a scheme was put in place whereby the *Herschel* Project Scientist and the HOTAC Chair in consultation could grant additional proprietary time to certain large programmes, in order to prevent the release of improperly or inhomogeneous calibrated or processed data.

10.4 Conclusions

Herschel was managed in different ways: top down with ESA as the agency in charge of the mission and bottom up with the three consortia of research laboratories in charge of delivering the payload consisting of three instruments (HIFI, PACS and SPIRE) and of ensuring the maximum scientific return. Industry was in charge of the payload integration to the spacecraft leading to a sometimes tense relationship with the instrument consortia. Despite a common ESA framework, the instrument consortia were not organised in an identical fashion, reflecting different cultures within the instrument and scientific communities that merged their interest within *Herschel*, but still concentrated on common science topics as overviewed in Sect. 4.5. However, they all had to manage both collaboration for bringing together a large team—an 'interstitial community'[10] gathered around their respective instrument (Shinn and Ragouet 2005)—and competition between the scientists within the team for influence and control over the observing programmes and for leadership of scientific publications. They all had also to collaborate with industrial partners and to develop sociality in order to do so.

All consortia devised their own way of organisation in order to make instrument design choices among many concepts in competition and ensure the most beneficial return to their scientific community and to the national agencies and ESA. The HIFI policy document tending to 'rationalise' the competition and allocate clear instrument subsystems to given laboratories or the SPIRE Constitution organising the general responsibilities and rights of all participants were two examples of ruling coopetition (Chap. 5). Coopetition operates in three steps: it involves *complex relations* between actors who are potential collaborators and competitors (within instrument consortia but also with industry), it occurs only when there is a relative *separation between the phases* of cooperation and competition through a sharing of tasks and responsibilities and/or with an alternation of the phases (hence the need for policy or constitution documents clearly defining these phases and the rules of the game), and it requires a *regulatory body* whereby the actors can confront their claims and can develop a code of conduct (here ESA). HIFI rules pacified competition for hardware selection, while SPIRE regulated and organised an open competition in such a way that the bolometer camera of the second financial contributor was not selected. Coopetition appears thus as a general condition of such a complex process of innovation: 'participants need to cooperate at the beginning of a competition phase in order to set the rules; winners must accept to concede something to the losers at the end if they want the competition to continue' (Sennett 2012: 117).

The lessons learned from the PACS consortium emphasise the sociality norms between industry and science instrument teams. The concept of sociality was defined in Chap. 5 as a way of collaborating, for a common goal, by people who have different values and norms. It implies in particular the development of a form of 'dialogical communication' instead of the 'dialectical communication', which prevails in the scientific field: when scientists discuss together, they try to prove their point and to discard other points of view because truth is going to emerge from this conflictual process, but when scientists and industrialists have to interact, first they must translate the arguments of the other side in their own words in order to

[10]'Interstitial community' recognises that an instrument does not arise from single institutional, disciplinary or industrial uses but from within an interstitial arena that lies between the usual poles of interest and organisation—university, firms, the state, military and so on.

avoid misunderstandings and, second, they must set the condition for compromises. Clearly, the *Herschel* project manager was the ruler of sociality and a key actor to cross boundaries in order to circulate between these two different social worlds. However, industry building the satellite, integrating the payload and designing interfaces between the spacecraft and the science instruments on one hand and the consortia delivering the instrument on the other hand created an unbalanced organisation and tension.

Finally, scientific planning and exploitation were also managed in different ways by the three instrument teams. HIFI and PACS operated a top-down approach, PACS was less collaborative and each co-investigator could use its guaranteed time as wanted, and SPIRE implemented a more bottom-up approach with the observing programme defined communally and optimised via a competitive time allocation process. But all three instruments used most of their GT in the form of Key Programmes, taking the view that *Herschel*'s core science was best implemented in that way, and also being aware that they would not get a chance to propose smaller programmes until after the first OT round and so would risk being scooped if they reserved some of their time to be used later.

Despite these differences in managing instruments and science, all consortia and ESA shared common objectives: maximising *Herschel's* efficiency to make best use of its limited lifetime and achieving all of its major scientific objectives.

In conclusion, the approach consisting of proposing ambitious science goals and making the necessary effort to build the required instruments led to three successful instruments on board *Herschel* (and likewise two *Planck* instruments). There is thus strong evidence that a managed-risk approach with respect to Technology Readiness Level leads to success and that coopetition and sociality play a key-role in this approach.

References

Poglitsch A, Waelkens C, Geis N, Feuchtgruber H, Vandenbussche B, Rodriguez L (2010) The photodetector array camera and spectrometer (PACS) on the *Herschel* Space Observatory. Astron Astrophys 518:2

Sennett R (2012) Together: the rituals, pleasures and politics of cooperation. Yale University Press, New Haven

Shinn T, Ragouet P (2005) Controverses sur la science. Pour une sociologie transversaliste de l'activité scientifique. Raisons d'agir, Paris

Chapter 11
Conclusion: Risk-Based Innovation and Knowledge Management

Abstract This book has considered the dynamics of innovation in space science missions. Making *Herschel* an outstanding success was not a simple task and was not the result of a smooth and flawless process. We have addressed the essence of innovation and of technicity from a philosophical standpoint including tools and concepts borrowed from humanities and social sciences (sociology, economy, C-K theory, etc.). Specific patterns were applied to characterise some technological and social features of the creativity of scientists and engineers. The notion of an 'instrumental community' has been coined and adopted in order to understand the effect *Herschel* had on the scientific field, using an original representation based on the bibliometric analysis of a co-citation network. While confirming the value of the approach taken by ESA in formulating its long-term plans, we conclude that the level of technology readiness and qualification was not a relevant discriminator for deciding to select FIRST/*Herschel* as a cornerstone of ESA's scientific programme, demonstrating that an approach to technology readiness based on managed risk can lead to success and to an optimised use of resources in terms of scientific outcomes.

This book has considered the dynamics of innovation in space science missions. Because of its complexity and its technical and scientific ambition, *Herschel* was our choice for this exercise. After analysing the strategic framework, the evolution of the mission assessment and definition, the science made possible, key technological innovations, and finally the management at ESA and scientific community levels, what conclusions can we draw from the *Herschel* mission and its risk-based innovation management?

The history of *Herschel* could seem like an obvious and smooth success story when summarised in few lines: 'Building on the several innovations and the success of the group of small infrared telescopes, in 1982 the European Space Agency decided to construct a major space telescope for the infrared. (. . .) The launch was executed perfectly, and both projects were an outstanding success'.[1] The analysis of its design and its achievements leads us to conclude that indeed this mission, which included the largest telescope ever flown at the time of its launch and a set of

[1]Francis Graham-Smith, Eyes on the Sky - A Spectrum of Telescopes, 2016, Oxford University Press.

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6_11

complex scientific instruments aiming at exploring a poorly known part of the electromagnetic spectrum, has been an ‘outstanding success’. But our analysis also demonstrates that this achievement was neither a simple task nor the result of a smooth and flawless process: there are particular organisational and social conditions underlying the success of such an innovative mission, which, as always, involved many vicissitudes. Our goal was to understand where innovation in both technology and management occurred throughout all the phases of development, from the first concepts to the observations in orbit and the exploitation of the data, and how such multifaceted inventive capacity can be maintained or improved for the benefit of future missions.

For that purpose, it was necessary to identify the key criteria necessary for such an analysis. Chapter 5 defines a typology of innovations according to the evolution of the ‘technical lineages’ proper to each critical *Herschel* components, and these were illustrated in Chaps. 6, 7, 8 and 9. Our method has drawn on the work of the French philosopher Gilbert Simondon, addressing the essence of invention and of technicity. It also included tools and concepts borrowed from social sciences (sociology, economy, C-K theory, etc.). Specific patterns were applied to characterise some technological and social features of the creativity of scientists and engineers. They include strong scientific motivation, non-linearity of the innovation process, relative independence of the concept from the knowledge base, radical changes and respect of the pace of different types of innovation. The notion of an ‘instrumental community’ has been coined and adopted in order to understand the effect *Herschel* had on the scientific field; this effect was characterised using an original representation based on the bibliometric analysis of a co-citation network. All these efforts were aimed at building a coherent ‘analysis toolkit’ to facilitate the implementation of knowledge management supporting innovation for, and leading to, the development of future space missions.

Let’s be honest: most of these concepts, such as ‘coopetition’ and ‘sociality’, were not at the forefront in the minds of the scientific and technical actors of the *Herschel* story—they were busy on the practicalities of defining and developing the spacecraft and its payload! Neither did we think in such terms at the beginning of our study: these tools were introduced along the way and eventually proved to be more accurate than the conventional categories used in previous traditional studies of earlier space missions. They were particularly relevant to the revelation of implicit knowledge. As the sociologist Pierre Bourdieu noted, the explanation of social facts may cause the agents to discover the truth of what they lived in a conceptual way that is not the implicit experience in which they lived.[2] This new way to perceive their own action may surprise them at first but also has a practical effect in the long term by increasing their freedom of action. Three concepts, which make explicit the know-how of innovators, can be identified:

Coopetition combines cooperation and competition. It is adopted in economics when the actors must adapt to a very competitive environment but still have to

[2] Pierre Bourdieu, Cours de sociologie générale. Volume 2, page 1073, Paris, Raison d’agir/Le Seuil, 2016.

collaborate in collective action (either to prevent other competitors from emerging or to pool their resources and exchange knowledge), a situation also often found in scientific activities. It was useful to apply this concept because it touches on the interactions between science managers and research institutions when establishing social division of the scientific work between the scientific community, industrialists and agencies such as ESA. Coopetition involves complex relationships between the various agents who, in their relationships with each other, have characteristics of both collaborators and competitors. Cooperation helps in pooling resources, while competition often assists in optimising the choice of leaders, in ensuring that key decisions arrived after thorough scrutiny and in the exploitation of the results. Coopetition needs political or constitutional agreements or documents clearly defining what are the rules of the game. It obviously requires the existence of a regulatory and arbitration framework enabling stakeholders to confront their demands and develop and observe a code of conduct. In the case of *Herschel*, that body was ESA at the top level, but the instrument consortia also developed their own internal mechanisms for moderating important decisions such as choice of competing technology and the plans for use of observing time and publication of the results.

Taken from the work of the philosopher and sociologist Richard Sennett, *sociality* expresses the complementary idea that big science projects, which involve agents with quite different backgrounds, require a specific mode of communication across the boundaries of their different social worlds. It involves in particular the development of a form of a ‘dialogic’ communication instead of the dialectical one, which has traditionally prevailed in the scientific field: when scientists discuss together, they try to prove their point of view and argue against those of the others because truth is supposed to emerge from this conflictual process. But when they have to interact with industrialists and funding agencies, they must first translate the arguments of the others into their own words in order to avoid misunderstandings, and, secondly, they must aim at, and be ready for, agreement on inevitable compromises. The concept of sociality led our analysis to identify within the various scientific and technical groups involved in *Herschel* the agents who were the key players capable of doing such a ‘diplomatic’ task. That approach may help the responsible actors of future missions to establish an effective mode of communication between the specific areas and the different players involved in the management and development of their mission in order to accomplish their duties with maximum clarity and visibility—avoiding either scientists or the industrialists creating ‘black boxes’ that might be impossible to understand for their partners on the other side.

Another fundamental reason why ESA could claim success for *Herschel* is based on the long-term approach adopted by its Scientific Programme in 1984 and embodied in the Horizon 2000 concept. M. Harwit emphasised, in the 1980s, that ‘short contracts prevent researchers from starting the kind of ambitious development program that leads to true innovation’.[3] The fundamental strategy of Horizon 2000 was meant to prevent that shortcoming. Quoting A. Poglitsch: ‘The Horizon

[3]Martin Harwit, *Cosmic Discovery*: *The Search*, *Scope*, *and Heritage of Astronomy*, Cambridge (Mass.), MIT Press, 1984, page 259.

2000 program was good, in many ways, because it was determined by setting our scientific priorities, with some realism about what could be done but also with the idea that if you wanted to do something new and interesting one may need a lot of development along the way' (Chap. 10). Extending over two decades, it was based on scientific ambition, balance between the main domains of excellence of the European scientific community and cost control through the 'design-to-cost' approach. This last aspect deserves a special analysis.

Design-to-cost is one key concept for understanding the reason of the success of *Herschel* mission. It was an explicit principle of Horizon 2000. However, what we have in mind with this concept is probably different to what was the original idea. Design-to-cost is not meant to justify a purely financial management but to encourage the respect of a principle of budgetary stability, staying in line with real scientific ambitions through regulatory mechanisms and feedback between ESA, national space agencies, instrumental communities and industrialists. It is as much 'cost to design' as a 'design-to-cost' in the strict sense. This point has not always been clear during the mission evolution. At times, attempts to drastically reduce costs almost jeopardised this policy (see Chap. 10). Nevertheless, implemented in Horizon 2000, the rather stringent discipline of design-to-cost had an essential role in triggering innovation at all levels of the mission, with designers intent on achieving the highest possible performance levels within the constraints of available resources.

The *Herschel* SiC telescope is a radical innovation, i.e. an invention that dramatically increases informational system performance. This creation of a new technical lineage enabled the construction of a large telescope for the FIR with major improvements in angular resolution and sensitivity. The future of FIR and submillimetre astronomy will require other important innovations to move to the next level, such as a deployable 20-m class telescope dish or space interferometry with a few satellites in a free-flying or tethered configuration.[4]

However, the question of which technical lineage is the most effective does not have always a single answer. *Herschel* illustrates how the choice of detectors for a space instrument is a complex exercise in which many technical, scientific and financial factors have to be taken into account. The optimum technical lineage for SPIRE bolometers ended up being an extension of the traditional and proven solution of feedhorn-coupled NTD germanium detector arrays. Newer technical lineages involving micromachined filled array silicon bolometers have been adopted, on the other hand, quite unexpectedly, for the PACS camera, although initially they were developed with SPIRE in mind. So, the SPIRE bolometer programme triggered two types of successful innovations: an incremental innovation (the SPIRE bolometers) and a breakthrough innovation (the PACS bolometers).

Innovation requires also a new 'associated environment'. Cryogenic systems are essential for the proper functioning of high-sensitivity instruments in space astrophysics. *Herschel* combined radiative cooling (for the telescope), cryostat with

[4] As stated in the White Paper 'Sub-arcsecond far-infrared space observatory: a science imperative' submitted to ESA in May 2013 by a large consortium of submillimetre astronomers.

superfluid liquid helium (for instruments) and in addition cryocoolers (for the bolometric detectors of PACS and SPIRE). The cryostat was based on customisation of the ISO cryostat. It was a one-shot compromise rather than a real invention. Although this choice was economically and scientifically motivated, it was a step back from the initial ambition of using Stirling cryocoolers to avoid having the large launch mass and limited lifetime associated with a cryostat. In addition, sensors operating at temperatures well below 1 K required additional coolers. Future detector technologies require colder temperatures of about 0.05 K, and future missions will use mechanical coolers rather than cryostats to reduce launch mass and allow longer operational lifetimes. Such developments must be undertaken in the early stages of the project to minimise risks during the implementation phase when the impact of design changes is greatest. The effort made in this area by large space organisations clearly shows that cryogenics will play a strategic role on future space missions.

Innovation is indeed a complex non-linear process. In the case of *Herschel*, it needed management at multiple levels: top-down with ESA as the agency in charge of the mission; bottom-up within the three consortia of research laboratories charged with providing the HIFI, PACS and SPIRE instruments; and transversal because industry was responsible for integrating the payload within the spacecraft (which resulted in a sometimes tense relationship with instrument consortia). The HIFI and SPIRE consortia adopted different approaches to the management of coopetition. HIFI rules pacified competition for equipment selection, while SPIRE organised and regulated a tougher open competition (to the point that the bolometric technology of the second financial contributor was not selected). The study of the PACS consortium shows strong emphasis on the sociality standards between industry and the scientific instrument. The ESA Project Manager and his team were key to crossing borders to circulate information between these two different social worlds. The continued motivation of the team leaders ensured the agreement of the consortium teams during the long implementation phase. The HIFI consortium implemented a regime of coopetition (alternating phases of cooperation and competition), which has demonstrated its effectiveness in producing many scientific results and training dozens of people at the highest levels in engineering, instrument science and project management. Similar conclusions could be drawn for PACS and SPIRE consortia.

The management of time and resources, i.e. the planning of the use of their guaranteed observing time and the exploitation of the data products, was also approached somewhat differently by the *Herschel* instrument teams. The HIFI and PACS approach was top-down, with PACS being less cooperative since coresearchers could use their allocated time as they wished, while SPIRE implemented a more bottom-up approach with a common observation programme that was optimised via a competitive time allocation process within the team. The three instruments used most of their guaranteed time in the form of Key Programmes, believing that *Herschel*'s basic science was being implemented optimally in this way, and incentivised by the time allocation rules which prompted them to put the bulk of their time into guaranteed time to avoid being scooped by

Open Time observers. Despite their differences in management, all consortia and ESA shared indeed a common goal: to maximise the scientific effectiveness of *Herschel* in order to make the most of its limited duration.

These considerations confirm the value of the approach taken by ESA in formulating its long-term plans. More than thirty years passed between the FIRST/*Herschel* proposal submitted to ESA in November 1982 and the completion of its science observing in April 2013 and the post-operations phase at the end of 2017. *Herschel* delivered what it promised, and more, thanks to a set of complex and advanced instruments whose technological availability was not at all guaranteed when the development of the mission began. The approach of proposing ambitious scientific objectives and making the necessary efforts to construct the necessary instruments has thus led to three successful instruments on board *Herschel* (and two *Planck* instruments). In particular, the development and exploitation of a unique space telescope and of an innovative set of instruments combining photometry and high-resolution spectroscopy offered major discoveries through a newly explored spectral window. In the Milky Way, *Herschel* discovered the ubiquity of the 'filaments' at the origin of the formation of the stars. It revealed that, for most galaxies, the amount of raw material available for star formation is also the constraining factor in their evolution. Its high-resolution spectral capabilities have made it possible to study the presence of water in all parts of the Universe. In terms of publications, the citation and impact classify this mission, along with *Planck*, as the most productive of all ESA space missions so far. *Herschel* also helped to structure the community of submillimetre astronomy by bringing together research traditions derived from radio-millimetre and infrared astronomy into a new instrumental community.

In summary, we have investigated innovations within *Herschel* in order to maintain a trace, a living memory of the scientific inventiveness, and we have decided to display it through many diagrams, which define our approach of knowledge management. We conclude that the Horizon 2000 approach adopted by ESA, national agencies, scientific community and European industries, including coopetition, fair sociality ('no black boxes') and design-to-cost discipline, was essential in stimulating and driving innovation in all areas of technology, scientific instrumentation, spacecraft equipment and management.

Overall, the *Herschel* experience has shown that, in a relatively short time, initially low Technology Readiness Levels (TRL) systems—and most important payload systems—can be developed and enabled for use in a complex space mission if all parties involved are organised, mobilised and motivated by the intention to achieve the best science within clear financial, technical and schedule constraints. A less constrained TRL framework probably explains the major innovations achieved during the development of the *Herschel* payload. This is facilitated if long-term funding by the responsible space agencies is ensured. Furthermore, the preparatory phase of the scientific payload elements played an important role in community building, leading to an efficient and smooth consortium formation. Our main conclusion is therefore that the level of technology readiness and qualification was not a relevant discriminator for the green light (or not) of a mission concept (see Fig. 11.1). This is an important lesson, and one at

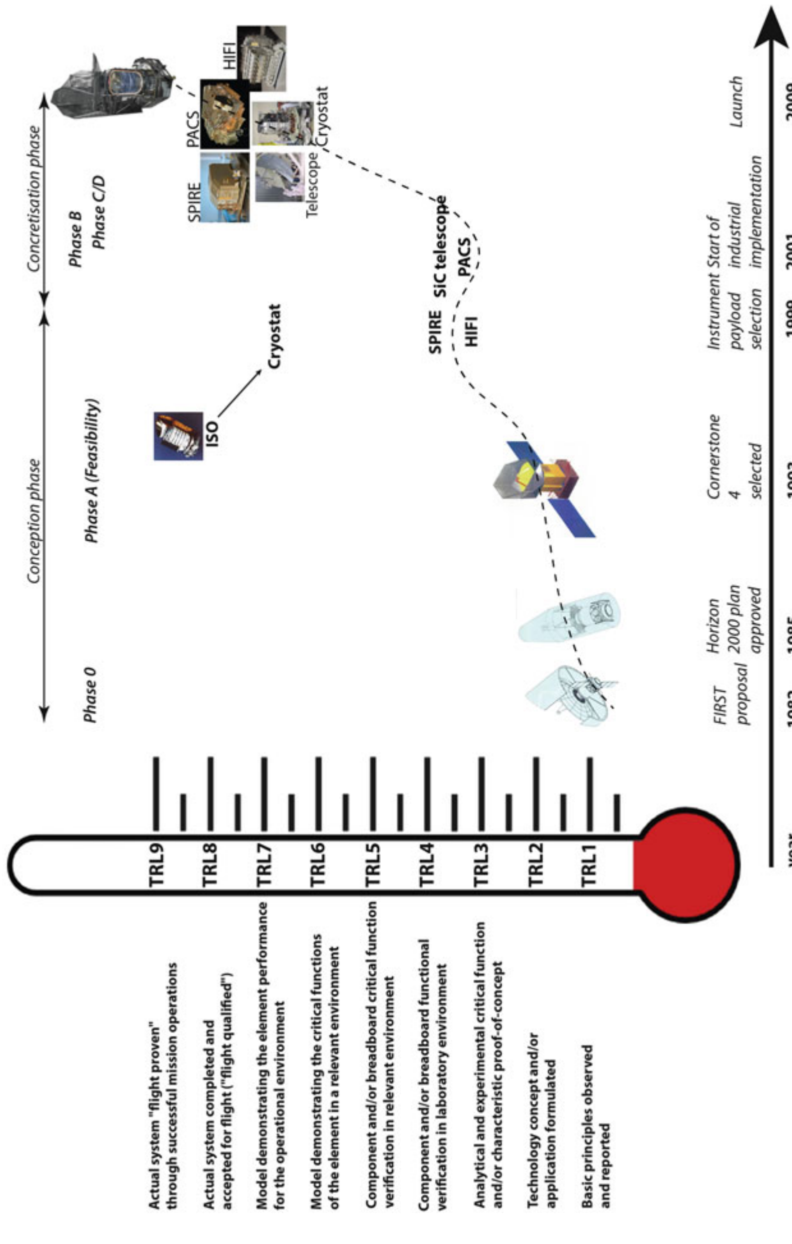


Fig. 11.1 TRL or Technology Readiness Levels of the *Herschel* mission concept over time, starting with the initial FIRST concept to the delivered payload and telescope before the space observatory launch in 2009. Phases 0 to C correspond to ESA mission phases: mission analysis, feasibility, preliminary

variance with today's more conservative practice, demonstrating that a managed risk-based approach to technology readiness can lead to success and to the optimised use of resources in terms of scientific outcomes.

In ordinary life, the further ahead we look, the smaller are our commitments. Even in big institutions, firm commitments are rarely made more than 10 years ahead of time. Because large space missions require all types of innovations, often including radical ones, their planning needs to be handled with a different approach to the management of resources over time. Instead of decreasing into the future, the commitment of ensuring budget stability must be reinforced in the long term. True innovations always begin as undecidable concepts which it may not be possible to evaluate with the initial knowledge base. Thus, the success of future ambitious space missions will rely on our ability to set new horizons, to establish a long-term strategy based on scientific goals and the principle of coopetition rather recurrent short phases of competition, strong innovation management and, very importantly, the commitment to 'never give up'.

Fig. 11.1 (continued) definition, detailed definition and qualification and production. These can be included in two phases: the conception and the concretisation phases (see Chap. 5). The TRL evaluation was not applied at the time of the mission assessment and definition. The values provided here are retrospective educated guesses based on the current definition of the TRL by ESA as given on the left hand side of the *y*-axis. Note that TRL were probably below 4 and close to 3 at the time of the payload selection consisting of HIFI, PACS and SPIRE instruments. The TRL of each instrument is an average of the TRL for each subsystem with the local oscillators within HIFI (see Chap. 8) or the bolometer array within PACS (see Chap. 7) probably below TRL 3. Neither was the telescope a well-defined option. Only the cryostat, originating from ISO, was probably high in TRL. At launch, all instruments were at TRL 8. © CEA (V. Minier)

Acronyms

A&A	Astronomy & Astrophysics (Journal)
AAAF	French Association for Aeronautics and Astronautics
ADS	(NASA) Astrophysics Data System
AFCRL	Air Force Cambridge Research Laboratory
AGN	Active Galactic Nucleus
AIAA	American Institute of Aeronautics and Astronautics
ALMA	Atacama Large Millimeter Array
AO	Announcement of Opportunity
AOS	Acousto-Optical Spectrometer
AOT	Astronomical Observation Template
APEX	Atacama Pathfinder EXperiment
ARC	Ames Research Center
ARTEMIS	ARchitecture de bolomètres pour des TElescopes sub-MIllimétriques au Sol
AS	Aerospatiale
ASP	Astronomical Society of the Pacific
ATHENA	Advanced Telescope for High-ENergy Astrophysics
AU	Astronomical Unit
AWG	Astronomy Working Group
BICEP	Background Imaging of Cosmic Extragalactic Polarization
BLAST	Balloon-Borne Large Aperture Submillimeter Telescope
BOOMERanG	Balloon Observations of Millimetric Extragalactic Radiation and Geophysics
BWO	Backward Wave Oscillator
CaC	Cost at Completion
Caltech	California Institute of Technology
CCD	Charge Coupled Device
CDR	Critical Design Review
CEA	Commissariat à l'Energie Atomique et aux Energies Alternatives
CESR	Centre d'Etudes Spatiales des Rayonnements

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6

CFRP	Carbon Fiber Reinforced Plastic
CIB	Cosmic Infrared Background (sometimes CIRB)
C-K	Concept-Knowledge
CMB	Cosmic Microwave Background
CMOS	Complementary Metal Oxide Semiconductor
CNES	Centre National d'Etudes Spatiales
CNSR	Comet Nucleus Sample Return Mission
CO	Carbon monoxide
COB	Cosmic Optical Background
COBE	Cosmic Bacground Explorer
COI	Composite Optics Inc.
Co-I	Co-Investigator
CoP	Commissioning Phase
Co-PIs	Co-Principal Investigators
CS	CornerStone
CSO	Caltech Submillimeter Observatory
CW	Continuous Wave
DARA	Deutsche Agentur für RAumfahrtangelegenheiten
DASA	Daimler-Benz Aerospace
DIRBE	Diffuse Infrared Background Experiment
DLR	Deutsches zentrum für Luft- und Raumfahrt (German Aeronautics and Space Research Centre)
EC	European Commission
ECA	(Ariane 5) Evolution Cryotechnique Type A (*Herschel* (& *Planck*) Launcher)
EM	Electromagnetic
ESA	European Space Agency
ESAC	(ESA) European Space Astronomy Centre
ESLAB	European Space Research LABoratory
ESO	European Southern Observatory
ESOC	European Space Operations Centre
ESRO	European Space Research Organisation
ESTEC	European Science and Technology Center
FCU	Focal-Plane Control Unit
FIR	Far Infrared Radiation
FIRAS	Far Infrared Absolute Spectrometer
FIRST	Far Infra-Red and Submillimetre Space Telescope
FTS	Fourier Transform Spectrometer
GaAs	Gallium-Arsenide
GHz	GigaHertz (10^9 Hz)
GIRL	German Infrared Laboratory
GOODS	Great Observatories Origins Deep Survey
GOT C+	Galactic Observations of Terahertz C+ (*Herschel* KP)
GREAT	German Receiver for Astronomy at Terahertz Frequencies

GSFC	Goddard Space Flight Center
GT	Guaranteed Time
H-ATLAS	*Herschel* Thousand Degree Survey (*Herschel* KP)
HEB	Hot Electron Bolometer
HELL	*Herschel* Explanatory Legacy Library
HEMT	High Electron Mobility Transistor
HEO	Highly Eccentric Orbit
HET	Heterodyne Instrument
HGBS	*Herschel* Gould Belt Survey (*Herschel* KP)
HIFI	Heterodyne Instrument for the Far Infrared (*Herschel* instrument)
Hi-GAL	*Herschel* infrared GALactic Plane Survey (*Herschel* KP)
HIPE	*Herschel* Integrated Processing Environment
HK	House Keeping
HOBYS	*Herschel* Imaging Survey of OB Young Stellar objects (*Herschel* KP)
HOM	Heterodyne Only Mission
HOTAC	*Herschel* Observing Time Allocation Committee
HRS	High Resolution Spectrometer
HSA	*Herschel* Science Archive
HSC	*Herschel* Science Centre
HST	*Hubble* Space Telescope
ICC	Instrument Control Centre
ICU	Instrument Control Unit
IDVP	Instrument Development and Verification Plan
IF	Intermediate Frequency
InSb	Indium antimonide
IOCR	In-Orbit Commissioning Review
IPAC	Infrared Processing and Analysis Center
IR	Infrared
IRAM	Institut de Radio Astronomie Millimétrtique
IRAS	Infra-Red Astronomical Satellite
IRSA	Infrared Science Archive
IRT/Spacelab	Infrared Telescope Experiment on Spacelab
ISM	Inter-Stellar Medium
ISO	Infrared Space Observatory
ISS	International Space Station
ISSI	International Space Science Institute
ISSTT	International Symposium on Space Terahertz Technology
ITT	Invitation to Tender
JCMT	James Clerk Maxwell Telescope
JPL	Jet Propulsion Laboratory
JWST	James Webb Space Telescope
K	Degree Kelvin
KAO	Kuiper Airborne Observatory
KID	Kinetic Inductance Detector

KL	Kleinmann-Low
KP	Key Programme
KPI	Key Performance Indicator
L(DN)	Lynds (Dark Nebulae)
L2:2nd	Lagrangian Point
LCU	Local Oscillator Control Unit
LDR	Large Deployable Reflector
LEO	Low Earth Orbit
LEOP	Launch and Early Operations Phase
LERMA	Laboratoire d'Etudes du Rayonnement et de la Matière en Astrophysique
LETI	Laboratoire d'électronique et des technologies de l'information
LIRTS	Large Infrared Telescope for Spacelab
LO	Local Oscillator
LOA	Local Oscillator Assembly
LOU	Local Oscillator Unit
LPAC	Launch Programme Advisory Committee
LSP	Legacy Science Phase
LSU	Local Oscillator Source Unit
MAMBO	Max Planck Millimeter-Bolometer
μm	Micrometre (10^{-6} m)
MAU	Millions of Accounting Units
MBB	Messerschmitt Bölkow Blohm
MDL	Micro Devices Lab
MFH	Multi-Frequency Heterodyne
MIPS	Multi-band Imaging Photometer for *Spitzer*
MIT	Massachusetts Institute of Technology
MLI	Multi-Layer Insulation
MMIC	Monolithic Microwave Integrated Circuit
MMS	Matra Marconi Space
MOC	Mission Operations Centre
MPE	Max Planck Institute for Extra-terrestrial Physics
MPIfR	Max Planck Institut Für Radioastronomie
NASA	National Aeronautics and Space Administration
NEFD	Noise Equivalent Flux Density
NEP	Noise Equivalent Power
NFM	Neyrpic Framatome Mécanique
NGC	New General Catalogue
NGST	Next Generation Space Telescope
NHSC	NASA Herschel Science Center
NICMOS	Near-Infrared Camera and Multi-Object Spectrometer
NIKA	Néel IRAM KID Array
NRAO	National Radio Astronomy Observatory
NTD	Neutron Transmutation Doped

OD	Operational Day
OECD	Organisation for Economic Co-operation and Development
OMC	Orion Molecular Cloud
OT	Open Time
PACS	Photodetector Array Camera and Spectrometer
PbS	Lead Sulfide
PDF	Probability Distribution Function
PDR	Preliminary Design Review
PI	Principal Investigator
PILOT	Polarised Instrument for Long-wavelength Observations of the Tenuous ISM
PSP	Priority Science Phase
PSR	Precision Segment Reflector
PVP	Performance Verification Phase
PWG	Payload Working Group
QMC	Queen Mary College
QMWC	Queen Mary and Westfield College
QO	Quasi Optical
RF	Radio Frequencies
RPG	Radiometer Physics GmbH
RSP	Routine Science Phase
SAC	Science Advisory Committee
SAFARI	SpicA Far Infrared Instrument
SAG	Science Advisory Group (SAG)
SAG	Science Advisory Group
SCUBA	Submillimetre Common User Bolometer Array
SDP	Science Demonstration Phase (*Herschel* mission phase)
SDS	System Definition Study
SDS	(FIRST) System Definition Study
SED	Spectral Energy Density
SEST	Swedish-ESO Submillimetre Telescope
SFR	Star Formation Rate
SHARC	Submillimeter High Angular Resolution Camera
SiC	Silicon Carbide
SIRTF	Space Infrared Telescope Facility
SIS	Superconductor–Insulator-Superconductor
SMP	Science Management Plan
SOAP	State-of-the-Art at the time of the Proposal
SOFIA	Stratospheric Observatory For Infrared Astronomy
SoHO	Solar and Heliospheric Observatory
SPC	Science Program Advisory Committee
SPC	Science Programme Committee
SPICA	SPace Infrared telescope for Cosmology and Astrophysics
SPIRE	Spectral and Photometric Imaging REceiver
SQUID	Superconducting Quantum Interference Device

SRON	Stichting Ruimte Onderzoek Nederland (Netherlands Institute for Space Research)
SSAC	Space Science Advisory Committee
SSD	Space Science Department
SSWG	Solar System Working Group
STSP	Solar-Terrestrial Science Program
Submm	Sub-millimetre
SWAS	Submillimeter Wave Astronomy Satellite
TALC	Thinned Aperture Light Collector
TB/TV	Thermal Balance and Thermal Vacuum (test)
TES	Transition Edge Sensor
THz	TeraHertz (10^{12} Hz)
TRL	Technology Readiness Level
TRP	Technology Research Programme
TWG	Telescope Working Group
UC	University of California
UKIRT	United Kingdom Infrared Telescope
ULIRG	Ultra-Luminous InfraRed Galaxy
US(A)	United States (of America)
UTC	Universal Time Coordinated
VLT	Very Large Telescope
WBS	Wide Band Spectrometer
WIRE	Wide Field Infrared Explorer
WISH	Water in Starforming regions with *Herschel*
XMM-*Newton*	X-ray Multi-Mirror Mission

Index

V. Minier et al., *Inventing a Space Mission*, ISSI Scientific Report Series 14,
DOI 10.1007/978-3-319-60024-6

W

X