planck

esa

European Space Research
and Technology Centre
Keplerlaan 1
2201 AZ Noordwijk
The Netherlands
www.esa.int/Planck

DOCUMENT

Planck Lessons Learned

Phase 1 report

Prepared by: Jan Tauber
Reference: SCI-S/2016.236/jt
Issue: 1.0
Revision:
Date of Issue: 22-09-2016

APPROVAL

Title: Planck Lessons Learned – Phase 1 Report	
Issue: 1.0	**Revision:**
Author	
Approved by: Planck Science Team	**Date: 22-09-2016**

CHANGE RECORD

Issue	**Revision**		
Reason for change	**Date**	**Pages**	**Paragraph(s)**

planck

esa

Table of contents

1 INTRODUCTION 4
1.1 Objectives 4
1.2 Audience 4
1.3 Process 4
2 THERMAL ASPECTS 5
3 LFI DETECTORS 11
4 HFI DETECTORS 15
5 OPTICS 20
6 POINTING 25
7 OPERATIONS AND SCIENCE GROUND SEGMENT 28
8 MANAGEMENT ASPECTS 41

1 INTRODUCTION

This document constitutes the report of the first phase of a Lessons Learned exercise for Planck. The exercise was initiated by the Planck Project Scientist, and was led and supported by the Planck Science Team.

1.1 Objectives

The objectives of the exercise were to:

- Capture (good and bad) experiences which could be relevant and useful to other experiments
  - Create a repository of relevant documentation for future use
- Formulate Lessons Learned and Recommendations based on these experiences

Note: the emphasis of this exercise was on the "sharp end" of Planck, i.e. the aspects that set it apart from other experiments. Aspects of Planck that are "standard" (e.g. solar cell technology) were not included in the scope of the exercise.

1.2 Audience

The target audience for this exercise consists of:

- ESA and other partner agencies
- Engineering and/or scientific teams preparing similar experiments who may profit in some way from the Planck experience.

1.3 Process

Initially a procedure was drawn up for the exercise (attached as Annex to this report). The procedure contained a preliminary identification of topics to be covered. Topics were grouped into two broad categories, to be covered in two sequential phases: the first phase concentrated on "technical" aspects and the second on "scientific" ones.

A list of individuals who had participated in and were cognizant of each topic was drawn from members of:

- The Herschel/Planck Project Team
- Industry involved in Planck
- The instrument development teams
- Satellite and instrument operations teams
- The Planck Science Office
- The Data Processing Centres

The individuals were requested to provide "bullet-type" items that they considered could provide experiences and lessons of interest to other experiments. Expert members of the

Science Team organized the numerous items provided into more general topics, and selected those topics of highest potential interest.

A two-day workshop was then organized in November 2015, where all those who had provided bullet-type items were invited to participate in open discussions. The presentations and minutes of that workshop (which are available on request) form the main basis for this report, which was prepared by the Project Scientist and approved by the Science Team.

Once the report is approved, it will be distributed to representatives of the target audience. Feedback and questions will be invited and additions to this report will be made as appropriate.

Note: when the second phase of the Lessons Learned exercise is completed, the corresponding report will be included and the final report will be reissued.

2 THERMAL ASPECTS

System design, development, and testing

Planck achieved its extremely challenging goal of providing detector temperatures of 100 mK, using technologies which had not yet been space qualified. The cooling chain flown on Planck was quite complex and risky, containing four stages of mutually dependent subsystems, and at least one single-point failure. However in the end it performed extremely well and largely as predicted from the ground. It shows that such a system can be implemented within the cost and risk envelope of an ESA "medium-sized" mission.

In such a complex chain, flowing system requirements to individual units becomes particularly important, and all parties need to have clear interfaces. Communication between all subsystems needs to be in place to adapt to inevitable changes. The Planck Project team did not use a tool to manage this process flexibly and this led to interface problems, in spite of the existence of "tiger teams" focusing on cryogenics. For example, shock levels and testing requirements came particularly late and caused problems for some units e.g. the 4 K cooler. These issues need to be addressed early and systematically.

The design of Planck was based on the approach that the SVM should not induce thermal loads on the cryo-chain. However some unexpected interactions between SVM and cryo-chain were discovered in flight: notable examples are the 4K panel temperature which was controlled via an SVM bang-bang control loop, and the power load of the RF transmitter which was turned on and off once a day – both caused thermal instabilities which required special measures to be taken in flight. In fact in Planck the "payload" was not confined to the PLM, but invaded much of the SVM. The traditional design separation between SVM and PLM led to some lack of attention in the interface design, meaning that some thermal requirements were not adequately set; and communication between SVM and PLM developers was not providing enough information to all parties.

A specific issue that arises in such a system is how to manage performance margins. Each of the elements of the long cooling chain needs margins in temperature and heat flows. Typically the overall performance is estimated and tested in terms of nominal and worst-case conditions. In the case of Planck, most subsystems worked at goal level which means that most if not all margins were realized. However, during testing on the ground it was realized that when some stages were too cold, even when working at lowest helium flow, the system oscillated. In flight a second similar instability was found.
Fortunately there were a number of heaters present which allowed to control the different stages, e.g. not having had the extra heater at 1.6 K would have made the system unstable.

In as complex a thermal system as Planck, testing becomes a crucial aspect of the development. The overall test plan included partial tests on dedicated facilities of different combinations of elements of the chains (HFI, LFI, SCS), including simulators for the rest of the cryo-chain. These tests were important and useful, but often the results were difficult to interpret and extrapolate to system level.
Therefore, system-level testing undoubtedly became the most relevant – though also the most difficult - element of the cryo-testing activities on Planck. A special facility at the Centre Spatial de Liège was developed for this purpose, and was a major effort and cost element in the development of Planck; this highest-level stage of testing was extremely important, though it came rather late in the overall schedule. Necessarily the facility included items not required in flight but which had an impact on in-flight performances, e.g. a special Helium circuit in the focal plane to accelerate cooling times during testing. It is worth noting that the nominal purpose of the facility did not cover cryo-operations of cooler or instruments, however it became a key element for understanding many operational effects which could not have been discovered in other settings (for example, thermal oscillations due to interactions between stages). In addition, simulations of failure and recovery of elements of the cooling chain were executed and were very useful to plan the HFI operations (daily monitoring of the cryo performance).

Operation of the combined cryo-chain in flight was foreseen to be complex and a special group was put together well before launch. This group contained members of all the elements of the chain and defined all the operations (including contingency) in great detail. The existence of this group was a key factor to the smooth and uneventful operation of the complex cryo-chain.

Lessons learned:

- It is possible to achieve 100 mK in orbit within the cost and risk envelope of an ESA medium-sized mission.
- When estimating overall thermal performance, include “best cases” in addition to nominal and worst-cases.

- Beware of traditional separation between independent SVM and PLM. Design principles, e.g. of thermal isolation between SVM and PLM, should be supported by actual requirements.
- Use a tool to manage requirements and their flow-down from system level to sub-units, and to maintain visibility to all parties.
- Design the system-level cryo-testing facilities simultaneously with the satellite, so that any performance impacts are understood. Carry out system-level testing early enough to correct the design if needed.
- Allow for adequate operational testing of cryo-coolers and instruments at nominal operating conditions.
- Operations in flight may be complex and need to be planned well in advance of launch.

Passive cooling

The initial goal of a passively-achieved temperature < 60 K was considered extremely challenging in the initial stages of Planck development. This first stage of the Planck cooling system consisted mainly of V-grooves and large honeycomb radiators on the PLM. The latter included paint especially developed for the purpose. The 3 large V-grooves carried heat exchangers for all three active coolers. Planck was the first satellite to implement the innovative V-groove system and initially there was considerable skepticism that it would perform well. However, the combined passive cooling system worked extremely well and its performance was as predicted from the ground, to the extent that V-grooves are now implemented or planned in many other missions. Nonetheless, there were many issues encountered along the way which required special attention – passive cooling may be "cryogenic classic" design but at the level required by Planck it is certainly not easy to achieve.

Lesson learned:

- V-grooves are efficient, predictable and reliable systems for thermal insulation and passive cooling.

20 K cooler

The sorption cooler on Planck was initially designed as a redundant system, carrying two identical units in cold redundancy. Eventually, the lifetime for a single individual cooler was not sufficient to cover the operational baseline, and it was clear that both would be operated in flight. Both coolers were indeed operated in sequence, but as it turned out their individual performances were quite different. Part of the difference was due to the mode of operation, which was optimized for the second (FM1) based on in-flight experience with the first (FM2). Note that only one of the two coolers was operated as an integrated system on the ground. Nonetheless it became clear in flight that the two coolers had intrinsic differences which had an impact on their individual operation, performance and lifetime.

A large part of the performance difference between FM1 and FM2 was likely due to their having been built by different teams, leading to hardware difference between the two units, i.e. larger variations in allowed powder distribution in FM2. The importance of the powder distribution was understood, but there was no requirement on equivalence of the two units – verification of such a requirement would certainly have led to discovery of the issue on the ground. The differences could have been found on the ground (e.g. by extensive use of neutron tomographic imaging) but were not, and remedied either by better process control or design changes (e.g. possible use of thin separators every ~10 cm along the beds).

In the end the primary lifetime limitation was due to contamination by hydrogen and methane of gas-gap switches included in the units. Design changes have been identified which are required to fix this problem.

Regeneration of the bed material was tested in flight and shown to work although at the expense of degradation of the heat switches and thus performance. Any future user of similar coolers should design this operational feature in. The temperature needs of regeneration, of course, need to be taken into account when defining the bakeout temperature to be applied for cleanliness of components (failure to do this is the assumed reason for the undesirable methane production in the FM-2).

The performance of the coolers in orbit did meet expectations from the ground, except at the level of temperature fluctuations at the cold end. It is interesting to note that the ability to predict fluctuation levels in the zero-g environment of space was considered a major problem issue (almost a show-stopper) during development, particularly because of the impossibility to do zero-g testing on the ground. It did not turn out to be a major issue, and temperature fluctuation levels were manageable though undesirably high (especially post-switchover from FM2 to FM1). The cause of these anomalous fluctuations remains not well understood, though it is likely related to the non-ideal two-phase flow design, and to thermal links between the two parallel units. To avoid similar problems in future use, it should be considered to eliminate the redundant cooler, and achieve equivalent redundancy by adding one or two extra sorbent beds and the ability to regenerate in flight.

As noted, operations in flight (e.g. adjustment of cycle times) were not initially optimal, and part of the faster degradation of FM2 is due to this. Because of the slow timescale for changes in performance to be measurable, it took a long time to acquire experience on the effects of operational parameters, and to optimize operations accordingly. Significant operational experience on the ground would have helped to reduce the time needed to achieve optimal operations. In flight, it was very helpful to have developed a control software with many “knobs”.

Lessons learned:

- The sorption cooler demonstrated good performance and operability in flight. The operational experience of Planck leads to a number of recommended design changes which should improve performance even more.
- Redundancy and lifetime are closely linked and requirements have to be made on both. It should not be assumed that redundant units are identical, instead this should

be verified.

- It is important to acquire operational experience of cryo-coolers on the ground in flight conditions. Allow for the possibility to adjust as many parameters as possible in flight.

4K cooler

The 4K cooler flown on Planck profited from a long heritage starting with laboratory models and prototypes for FIRST, which really helped to gain confidence in the cooler. Joint operation in a laboratory setting with a development model of the dilution cooler was also very important in this respect, particularly in learning how to deal with gas contamination. Heritage went as far as re-use of a (refurbished) pair of compressor units – although it is possible that it would have been easier/cheaper to build new models.

The lengthy demonstration programme within scientific institutes was very successful and most of the problems that arose occurred in the transfer of expertise to flight development by industry. There were significant problems in particular with the drive electronics. An active vibration damping control had been developed at RAL, which had proven to be very effective. Unfortunately the actual Planck manufacture was contracted to a different company which had little or no experience in space applications of power electronics. As a consequence there were many problems, ranging from mechanical issues (EEPROM legs breaking) to the grounding scheme not working, which had significant schedule and cost impacts. The mechanical units also suffered from some problems which were found only during actual testing (e.g. coil wire breaking). Part of these problems can probably be attributed to the local management (by PPARC) of the 4 K development, which was not following the refurbishment closely enough, nor reacting fast or thoroughly enough to problems (e.g. the grounding scheme).

Overall the 4 K cooler – initially considered a high-risk item - worked very well in flight, with only one shutdown in the earliest operational phases, probably due to particle contamination. Continuous operation over ~4 years is a remarkable record for a cryo-mechanical device.

In terms of the performance of the 4 K cooler in flight, the temperature stability of the cold end was extremely good, being mainly driven by the 20 K stage. The major issue encountered was related to export of EMC interference to the HFI detectors at levels far (30 times) out of specification, which can be traced back to a faulty design of the grounding scheme of the drive electronic unit. This had a negative impact on the ultimate sensitivity of the HFI. The operation of the cooler in flight was fairly uneventful, if anything the tuning of the harmonic control may have been somewhat overdone.

Lessons learned

- problems arise when the development of critical units is taken over by industry for flight development, especially if they have little experience – these cases need tight oversight and management to spot problems early and take fast action.

- Vibration damping by electronic control of harmonics works well, but as much care needs to be taken with EMC as with mechanical vibration.

0.1 K cooler

The 0.1 K cooler provided the temperature needed for the HFI bolometers to achieve their high sensitivity and was therefore a key element in Planck's success. This was a new development and a single-point failure for the mission, and therefore constituted an important risk factor for Planck. Recognising that, the local funding agency (CNES) invested very significant resources in industry on this development, and managed it very tightly. This investment in resources and management paid off, as the cooler worked in flight predictably and in specs from the cryogenic point of view.

As for the other elements of Planck's cryo-chain, a key element for success was testing at many different levels, up to system level. The system-level operation allowed to understand adequately the thermal interactions of all elements of the cooling chain. In fact, it is only when operating the full cryo-chain that oscillatory behavior (unstable liquid helium evaporation in pipes between the 4 K and the 0.1 K stages) was discovered in certain ranges of the operational parameters, although a substantial safe domain remained in the operational parameter space.

It is worth noting the failure of a pressure regulator which took place right before the flight model was shipped to the launch site. It emerged that this regulator had been suspected of failing on a different spacecraft, but had never been identified as suspect in connection with HFI. Planck was extremely fortunate that this unit failed before launch.

The main problem encountered in flight was related to the effects of cosmic rays (CRs). CRs had been expected but not at the high levels encountered – this was clearly an oversight as the effects of solar modulation were known and could have been predicted (or indeed estimated from e.g. SOHO SEU levels). Had this been predicted, probably better use would have been made of the on-board Solar Radiation Monitor (SREM).

CRs not only induced direct response on the bolometers (glitches discussed under HFI detectors), but also injected significant heat at 100 mK. CRs are the driver of the temperature fluctuations at 100 mK and they also affected the thermometers controlling the temperature regulation PID. As a result, it became impossible to fulfil the temperature stability specs – however, temperature fluctuations were well monitored by the blind bolometers present on the 100 mK plate, which allowed a very good a-posteriori removal of these fluctuations.

At the same time, high energy cosmic rays induce showers of lower energy particles which affect simultaneously many detectors, induce temperature fluctuations, and lead to correlated noise between detectors. These are more difficult to remove as the thermal propagation would need to be modeled.

It may be worth to note that solar proton enhancements (particularly those with strong 100 MeV enhancement) can also disrupt a 100 mK stage. This happened a couple of times in the mission for a few hours/day each time.

Lessons learned

- Important risks in the development of this unit were mitigated by significant investment in industry and management.
- Operational testing at system level of low-temperature units is extremely important to avoid important surprises in flight,
- The level of the thermal impact of CRs need to be assessed thoroughly in the design phase of any experiment using <100 mK detectors, and mitigating measures taken, as they could be the determining factor for the temperature and thermal stability of very-low-temperature stages.

3 LFI DETECTORS

Definition of requirements and early phases

The definition of requirements on the LFI evolved significantly during the development. One specific example is polarization: initially it had very low profile in the mission scientific case and therefore not much attention was put on it in the initial design phases. The very concept of the LFI receivers (differential pseudo-correlation with external reference loads) was optimized for total intensity, not for polarization. Also, during the mission development, the thermal structure of LFI was changed substantially, introducing active cooling for the front end at 20K and separation between front and back end. In addition, 4K reference loads, implemented on the HFI 4K box, were introduced in the design to reduce the dynamical range spanned by the LFI, at the expense of adding constraints on the integration of the two instruments.

The identification of what were the main systematic effects to be dealt with started early on. An LFI systematic error budget was produced and regularly updated.

One aspect which received a lot of attention, as it was identified early on as a critical area, was the thermal stability of the LFI interfaces. Stringent requirements were placed, resulting in good stability in operation and in systematics that gave a subdominant contribution. On the other hand, bandpass measurements and knowledge (which have impact on optical response and polarization leakage for polarization), while part of the ground calibration plan, for programmatic reasons were not allocated enough time in the measurement campaign, which resulted in sub-optimal characterization. Similarly, the effect of far-sidelobes was probably under-estimated.

Seen a-posteriori, the most important systematic effects known today (gain calibration, ADC non-linearity,) were not expected to be the dominant contributors in the early phases, and were not identified as such during early testing. This is not very surprising, considering that it is very difficult to create a test environment on the ground which is representative for subtle systematics.

An "LFI instrument model" was developed early on to support the design and development phase. However a posteriori it is clear that the level of detail, accuracy and validation of such model was not sufficient to effectively support the development and calibration campaign. It was also recognized in later phases that it is important to keep performance margins for "data processing".

Much of the analysis of flight performances depends on the availability of housekeeping information. A-posteriori, it is clear that the LFI would have benefited from: (a) better monitoring of temperatures throughout the radiometers, reference loads, and telescope; (b) better monitoring of the LNA biases. Because the HK requirements are set fairly early on in the mission, it is difficult to change them if/when testing reveals additional needs; therefore they need to be given careful thought in the earliest stages.
Similarly, the ability to analyze very raw (uncompressed, unprocessed) scientific data is very beneficial to understanding subtle effects, but with the pressure to reduce telemetry bandwidth this ability is often compromised.

It is important to remember that the LFI was a very innovative instrument at the time it was conceived, it broke ground in many technical areas, and in the end it performed very well in flight.

Lessons learned:

- It is very difficult for missions like Planck, which depend on exquisite control of systematics, to know very far in advance which will be the most significant problems in flight. An instrument systematics error budget should be developed already during the design phase and should be updated throughout the development, operation and data analysis phases.
- It is important to develop very early on an instrument model (a simulator of scientific data) which includes all possible systematic effects, even those which are considered negligible at the time. The model can be used to set requirements which must be verified on each one of the systematic effects. The model parameters should be validated with unit-level tests and kept up-to-date with test results.
- As long as the data processing pipelines are not fully developed, keep a "margin" for them in the performance budgets.
- Do not underestimate HK requirements even in early stages.
- Ensure that it is possible to access in flight scientific data in a very raw state for diagnostic purposes.

Development phase

The LFI had very complex technical interfaces, both internal (e.g. waveguide connection between front-end at 20 K and back-end at 300 K) and external (e.g. to HFI, mechanically and radiatively via the 4 K loads; to the 20 K cooler; to the three V-grooves at approximately 50K, 100K and 150K; optically to the telescope). Furthermore the LFI subsystems were provided by many different institutes. The management aspects were therefore heavy (documentation, reporting etc) and not part of the usual working practices of many of the participating institutes.

One aspect of interfaces which is particularly important is the management of margins, which depends on building trust in the system leads and keeping open books among all parties.
An illustrative example, and a particularly difficult interface, was the division between receiver front- and back-ends for the 30 and 44 GHz channels. The subdivision was made for balancing the contribution from individual countries, but was technically sub-optimal. The very high interdependence of the detailed characteristics of the front and back end amplifiers made the overall optimization process far from ideal. In spite of the extra effort made by the programme manager and the general collaborative spirit of all parties, , insufficient communication between providers, and slow turnaround times resulted in somewhat sub-optimal overall receiver performance. Seen a-posteriori, it is possible that a better power dissipation margin management could have improved the front- to back-end split (especially after the 100 GHz channel was removed from LFI).

Another example of a complex interface are the 4K loads provided by HFI for use by LFI. Temperature monitoring of these loads was run by the HFI team and data made available to the LFI team.

The LFI complex interfaces also involved thermal interaction between the SVM and the instrument. As an example, the LFI was sensitive to the imprint of the power cycling of the telemetry RF transmit signal in the early phases of the mission. Fortunately there was a simple solution of leaving the transmit carrier on all the time, though the initial part of the mission was affected.

The LFI detectors were each essentially treated at individual level during development, rather than as arrays. Quality and Product Assurance practices were followed, but quite unevenly across the different suppliers. Two cryogenic amplifiers of the flight Front-end Modules failed during pre-delivery testing, and could fortunately be repaired – though not fully re-characterized on the ground. It highlights the high level of risk taken at detector level. Another example is in the design of the harness and electronics. Optimal grounding practices were not followed, which resulted in significant ground instabilities being discovered during system-level testing; this had a big impact on the tuning procedures during commissioning, whose complexity and duration (in commissioning phase) increased dramatically as a consequence.

A major issue for LFI during its development was the loss of the 100 GHz channel, due to a combination of technical and programmatic issues. After this traumatic loss, the Focal Plane Unit was not re-optimized, though it could have been. One of the important rationales for the 100 GHz detectors was to enable a direct comparison of LFI and HFI data at the same frequency. It is not clear to what degree the direct comparison would have helped in the analysis of systematics. Nevertheless, 70x100 GHz cross spectra were used to demonstrate that the systematics between the two instruments were uncorrelated and thus could not have been used to identify systematic effects.

Lessons Learned:

- Avoid to the extent possible the multiplication subsystems and suppliers, especially when highly interdependent interfaces are involved which drive performances
- System leadership has to be clearly established, and must be based on the trust of all sub-system suppliers
- Evaluate carefully the potential for S/C operations and interfaces to impact the instrument and vice-versa. In general the split between S/C and Instruments tends to isolate and hide some of the technical issues even when there is potential for interaction.
- Agree early among all suppliers on how to manage margins.
- Strong Quality and Product Assurance practices are required throughout the development.

On-ground testing

(Note that some of the points raised here are applicable to both LFI and HFI).

It is always the case in experiments like Planck that more testing on the ground would have been more useful. Overall, it is felt that not enough time was spent testing LFI (both at subsystem and system level). Part of the reason for this was that the time needed for testing was often underestimated, causing stress on the overall schedule.

In the case of LFI, the main areas where significant shortcomings have been identified are: bandpass measurements, and cryo-performance testing.
Regarding the bandpass measurements, they are most important for polarization. The complete measurement of spectral response was part of the test plan, but programmatic and schedule constraints resulted in reduced accuracy of the measurements. It is possible that with a stronger early recognition of the criticality of bandpass effects, a more complete ground measurements and/or means to measure in flight would have been required and put in place.

More generally, it is in practice not possible to measure end-to-end polarization properties on the ground at the level of Planck sensitivity, so this must be achieved using flight data. Accuracy of instrument parameters using the sky data redundancy should be established by simulation and taken into account when planning ground test requirements.
It is possible (if feasible) that the availability of in-flight internal polarized sources could be a significant advantage and would off-load the need for sophisticated ground testing. Similarly, testing for ADC non-linearities to the level of accuracy required by experiments such as Planck is also complex, and requires long testing times in very stable conditions; this is another area which it was not foreseen to verify either on ground or in flight.

Overall, it seems that more thinking of in-flight "calibration" in the early phases would lead to a clearer understanding of what are the quantities that need to be measured on the ground and in flight respectively, and help to optimize the way in which these measurements are made.

The test at system level, i.e. of the fully integrated satellite in the cryo-facility at the Centre Spatial de Liège, is considered to have been crucial for LFI. This was the only time on the ground where any interactions between coolers, LFI and HFI could be tested, and the instruments operated in a flight-representative way. It should be noted that this test was not originally intended to test the instruments and their operations, but was supposed to be limited to validation of the satellite thermal behavior. However, it was extremely important to assess a global end-to-end performance verification and to learn to operate the instruments during this time.

The model philosophy included spares for most LFI subsystems. However, a full working model was not kept in operation after launch. It is felt, however, that doing that would have helped to manage better operational aspects after launch, as well as to achieve detailed understanding of the detector performances.

Lessons Learned:

- Keep ample time margins for testing. Ensure that the schedule for test set-up preparation is properly accounted for, again with the due margins.
- System-level tests should be included that allow the full operation range of the instrument in flight-representative conditions
- Include also "in-flight calibration" as an integral part of the on-ground instrument model, to help optimize the means to measure calibration parameters both on the ground and in flight. The accuracy of estimation of instrument parameters using the data acquired in flight should be established and taken into account when planning ground test requirements.
- Keep a working spare of the detector assembly on the ground for testing during flight operations.

In-flight operations

Often the teams who have developed the instrument and tested it on the ground move on to different projects after launch. It is very important that this experience is transferred to the in-flight operational teams. It was crucial for LFI to have a number of key people of hardware development participate throughout the operations and data analysis phases.

Lessons learned:

- Ensure that the on-ground development/test teams continue to support the flight operations and data analysis.

4 HFI DETECTORS

The selection of the Spider-web bolometers and Polarisation-Sensitive Bolometers as HFI detectors was based on their high sensitivity, low heat capacity (i.e. providing fast time response), low sensitivity to vibration, and low cross-section to cosmic rays. While the HFI detectors generally lived up to expectations (being back-ground limited), a number of performance-related surprises were found in flight. In the end, the performance of the HFI is limited by these systematics rather than by the background only for polarization and for multipoles lower than 30.

Time response

The bolometer time response is a major and unavoidable systematic effect for a measurement of the angular power spectrum, as it determines the calibration transfer from multipole ell=1 (dipole frequency) to higher multipoles, and generates low frequency noise ($1/f^2$). The detector requirements were written as a maximum value of a single time constant, and the detectors tested at a subsystem level met those requirements. However, once HFI was integrated and tested as an integrated instrument it became clear that the time response was more complex and a major effort was undertaken to measure it. However, it is very difficult to make measurements in the lab to the 0.1% level of the time response across a band between ~0.016 Hz and ~100 Hz. It was only in flight that it was clearly recognized that the physical processes that lead to the longest time constants are more complicated than expected, and they are present over a wide range of times. To constrain these processes, on-ground measurements on similar devices were made after launch; during these measurements it became clear that the time response was influenced significantly by the links between the wafer and the detector housing. In flight, it is possible but difficult to separate the beam response from the time response in the data acquired from the sky; only one test had been planned in flight for this purpose (the satellite was spun up to a higher rotation rate). Nonetheless, the high redundancy of the data permits to invert the problem to solve for the time constants. In particular it became possible to measure and remove a very long time constant fairly late in the data processing.
In the end, the best knowledge on the time response of the detectors is derived from a combination of data acquired both during ground testing and in flight.

Lessons learned:

- The time response characteristics of bolometers are complex and influenced by the device itself as well as the structure in which it is embedded.
- Plans should be made to clean flight data empirically, i.e. relying on data acquired in flight, including where possible dedicated in-flight measurements.
- It takes a long time and a lot of data redundancy to identify and characterize systematic effects affecting large angular scales.

Ground measurement of bandpass and polarization characteristics

Although the bandpass response was measured on the ground, the estimated accuracy of the measurements was at the edge of what could be tolerated for Planck, e.g. for ground-based correction of temperature-to-polarization bandpass leakage. Similarly the polarization angles and cross-polarisation leakage levels were measured on the ground with low accuracy. This type of measurement is very difficult to realize accurately on the ground. Ground measurements of a strong polarized source (Crab nebulae) were used to check the leakage parameters. This was not as accurate as expected. Using the data redundancy from the scanning strategy was shown to provide the best accuracy. This also showed that the ground measurement accuracy had been underestimated.

Lessons learned:

- Future missions clearly need to consider very carefully the tradeoff between the ground measurement of bandpass and other polarization characteristics and the extraction of these parameters from the data as they could become very demanding, or alternatively (even better) develop accurate methods based on in-flight data. Note that this also applies to other parameters such as time transfer functions.

Cosmic-ray glitches

In flight, the HFI detectors were found to suffer from a higher glitch rate than expected. In particular, there was an unforeseen family of long-tailed glitches that contaminated a very significant fraction of the data and created an effective low-frequency noise in the detectors. The radiation environment seen by the detectors at L2 was correctly estimated for interaction with the bolometer themselves. The interaction with the bolometer silicon wafer were severely underestimated because they were computed assuming thermalization. Ballistic phonons dominate and create a large number of small glitches with longer time constants. Thus the use of the modeling tool GEANT was not adapted in this context. On-ground radiation testing with energetic particles was not done . The modulating effect of the solar wind during the solar cycle on the flux of cosmic rays was expected (note that Planck was launched as the solar maximum activity phase was starting; most of the flux it experienced was from Galactic particles, except during strong solar flares). As mentioned earlier, the modulation of the solar wind extends to low frequencies and generated fluctuations in $1/f^2$ of the temperature of the bolometer plate.
Through intensive analysis of in-flight data and a campaign of ground tests (that helped to identify the "zoology" of glitches), the effects of cosmic ray induced glitches on the data were very significantly reduced, though at a small cost to the final science products (i.e. rejection of 10-25% of the data in each detector; introduction of some correlated noise). It is likely that residuals of long-tail glitches remain in the HFI data and cause part of the 1/f residual noise.

Although Planck carried a radiation monitor on board (the SREM), it does not measure all the particle energy ranges responsible for HFI glitches; data from the SREM could have been used more effectively. Fast sampled 100 mK thermometers which have larger volume than bolometers were more sensitive to particles and dark bolometers were better suited for correcting the baseline temperature drift.

Lessons learned:

- The instrument radiation responsivity model should be developed to high level of detail
- On ground tests of radiation sensitivity of detectors or more generally the detector environment using particle accelerators (e.g. as done a posteriori for HFI) must be carried out in an in-flight representative (e.g. cryogenic) situation.
- Future detector design should take into account heat deposited by cosmic ray hits in all areas of the detector and support structure. For example, the design could mitigate ballistic phonon propagation from support structures to the detector. As a second step, increasing the heat capacities of support structures could absorb better the cosmic ray energy.

ADC nonlinearity

The Analog-to-Digital Converters in the HFI detector readouts have non-linearities just around the zero bit, which is exactly where the HFI signal appears. The non-linearity was measured on the ground but only to check specifications. This became – quite unexpectedly – one of the most damaging systematics for HFI because of insufficient accuracy of this measurement
If recognized early enough, the instrument tuning could have been adjusted to minimize the problem, i.e. shifting the signal away from the central range of the ADC. Furthermore, a more accurate ground-based measurement of the nonlinearity of the ADC devices would have allowed to correct the flight data. However, the problem was only recognized well into routine operations, as the main originator of apparent "gain variations" which could not be due to the bolometers themselves.
It was fortunate that Planck continued to operate the LFI instrument after the HFI detectors warmed up due to the exhaustion of cryogenic fluids. Over many months within this extended period, it was possible to make specific detector-by-detector measurements of the ADC characteristics that allowed to make a first-order correction of the effect of the non-linearity on the data. A better empirical correction has later been implemented using the redundancy in the sky data.
A secondary effect of the ADC nonlinearities is that they are modulated by the pickup of harmonics from the 4K cooler ("4K lines"), complicating their correction in a subset of detectors. This effect could possibly have been mitigated if access much more raw telemetry had been available.

Lessons learned:

- Ground testing and verification of device specs (e.g. ADC) has to be very thorough and any strange features followed up.
- Readout electronic design should take into account the intrinsic nonlinearities of all ADCs.
- It is crucial to plan the ability to occasionally download raw, uncompressed data for diagnostic purposes.
- Distribution of telemetry should remain flexible to help with such problems.

"Random Telegraphic Signals" (RTS)

RTS, a randomly occurring jump in the signal between two baseline states, rendered two detectors completely unusable during the mission and others intermittently so (it may contribute at a low level to remaining 1/f noise). The cause of this problem is not very well understood but it is probably related to contact parasitics in harness connectors (changes were seen after unplugging/replugging the focal plane unit). This effect was seen on the ground in the fully integrated system and. Fortunately this effect did not spread to additional detectors after launch. It is not clear if any measures would have helped to reduce this problem, e.g. very strict EMC requirements.

Lessons learned:

- It should be anticipated and accepted that a small fraction of detectors will not be useable in flight.
- Perform tests for contact noise on the ground.

4K lines

The detector wiring picked up electromagnetic noise originating in the drive electronics of the 4K cooler, due to a suboptimal grounding scheme (which led to pickup being a factor 30 out of specifications). Early in the mission design phase it had been decided to phase-lock the bolometer AC readouts to the same discrete frequency spectrum used by the cooler driver. This turned out to be the right decision, allowing to remove most of the contamination from the 4K lines by direct notch filtering.

Lessons learned:

- Phase-locking critical frequencies to the same spectrum of harmonics allows to manage electrical pick-up effectively.

CO line contamination

The bandwidth of the HFI detectors is very wide and contains strong lines of interstellar CO (the J=1-0, 2-1, and 3-2 rotational lines of CO are among the strongest emitters from the galactic interstellar medium). These lines were expected to be (and are) weak at high galactic latitudes, away from molecular clouds. The tradeoff between using more bandwidth and narrower filters between strong lines was not considered in depth. On the other hand, the all-sky maps of CO, including the weak emission at high-galactic latitudes, are one of the astrophysical legacies of Planck.
Several effective methods were developed during flight to extract the line emission within each HFI band from the continuum. Using these methods, Planck confirmed the (statistical) weakness of CO lines at high latitudes by comparison to deep but sparse ground-based surveys. To get the absolute intensity, some of the methods depend on the

knowledge of the bandpass shape available from on-ground measurements.

A-posteriori, it remains unclear whether it would have been more advantageous to exclude these lines from the detector bandpasses or not, even though excluding the lines from the CMB bands (100-353 GHz) would have reduced significantly the sensitivity. The final effect of these lines on CMB intensity science is small. For polarization at low multipoles, the leakage has to be estimated from the data themselves and removed. Furthermore, as it is possible that CO emission is weakly polarized, the effect on CMB polarization science at sensitivity levels higher than Planck could be increased.

Lessons learned:

- When planning future CMB experiments, strong line emission from CO and potentially other species should be considered as a distinct, weakly polarised, and significant component of diffuse galactic emission. Consider also unresolved lines.

100 GHz PSBs

The PSBs flown at 100 GHz were not part of the initial design of HFI, as they were considered immature technology at the time. They were introduced very late in the development (2003) as a mitigation measure when the 100 GHz channel of LFI (which was polarised) was removed. This was quite a bold and risky measure at the time, taken in recognition of the importance of polarisation measurements at 100 GHz. This decision did pay off though as they performed very well in flight.

Lessons learned:

- Though the baseline technology for science missions must be mature and safe, it should remain possible to include riskier but more highly performing elements, as long as they do not imply single-point failures.

5 OPTICS

Requirements

In a classical approach, high-level scientific requirements (i.e. what is needed to achieve the science objectives) are formally and then flowed down into subsystems (telescope, reflectors, focal plane, detectors) in a traceable and consistent manner, and with a clear view of all phases of the development and operation, i.e. manufacture, on-ground verification and characterization, transition into orbit, in-orbit characterization and use of optical information for scientific analysis. The first rung in this approach was formally missing in Planck, i.e. although high-level scientific requirements did exist for both performance and knowledge, they were not formally recorded and accepted by the Project. Instead, an optical design was arrived at based on ad-hoc criteria and imposed on the manufacturers without a clear view of the consequences on the rest of the programme. The technology for the telescope was decided based on the wish to exploit development efforts made previously for Herschel. Manufacturing requirements for the telescope were

set largely based on what was believed that the technology could achieve rather than on real needs. Different parties were responsible for reflectors, structure, and verification.
In addition, a confusing mix of optical and radio requirements was adopted, partly driven by the simulation tools available. Much energy was expended in sorting these requirements out, and in fact quite a few errors and misunderstandings were uncovered in the process.

The choice of technology, and the contractual setup, had quite drastic consequences as a number of manufacturing requirements were not formally met, and the verification programme had to be significantly curtailed and adapted during the development.
On the positive side, the lack of formal high-level requirements did provide flexibility to adapt the various parts of the programme when driven by circumstances.
Fortunately the real performance requirements on the Planck telescope were fairly permissive, and – contrary to initial expectations - the required knowledge could be largely achieved in orbit e.g. due to mission extensions which allowed to observe planets many more times than initially planned.

Lessons learned:

- Do not bypass the classical approach to set formal scientific requirements at high level and flow them down systematically and traceably to subsystems, taking into consideration all phases of the mission (from manufacturing all the way into scientific analysis)
- Requirements on knowledge may be more important than on performance and need to be specified early
- agree from the start on the formal terminology (radio or optical) to be used and stick to it

Corrugated horns.

The Planck instruments both used corrugated horns to collect radiation from the telescope. This technology is very mature and their design is straightforward. The LFI optics were “simple”, with a corrugated horn feeding two orthogonally-polarized receivers via an Ortho-Mode transducer. The HFI optics were much more complex, with a triple horn system and filters at different temperatures feeding radiation to a bolometer in a cavity; while the first corrugated horn dominated the optical response, the other elements had non-negligible effects. Manufacturing imperfections in the corrugations are known to have an effect on the horn patterns, and these imperfections are difficult to control or measure at the highest Planck frequencies (HFI). Manufacturing imperfections are likely to be the cause of higher-than-expected far-sidelobe levels at the highest HFI frequencies).
The on-ground verification of the horns was not complete. Both LFI and HFI horns (front ones for HFI) were measured at ambient temperature only, verifying that their co-polar response was as designed (though note that even for LFI the cross-polar response did not behave as modeled). The measurements were monochromatic, whereas the actual detectors are very broadband. Neither LFI nor HFI detector patterns were measured at operating (cryogenic) conditions on the ground. In particular for HFI it would have been very difficult to develop an adequate measuring scheme including the entire detector package. In actual practice, theoretical feedhorn patterns were used as input to higher-level optical modeling.

On the other hand, the detector feedhorn patterns almost directly determine the system angular responsivity in key sidelobe regions, at levels which are similar to the current overall uncertainties. In particular for polarization, it is essential to measure their properties on the ground because these levels could not be measured (by Planck) in flight.

Lessons Learned:

- the optical response of individual detectors in the focal plane should be specified and measured on the ground in operating conditions and with real detectors. For CMB polarization experiments, some leakage parameters of detector optics will probably have to be extracted from the in-flight data.

Multi-moded horns

HFI elected to use multi-moded horns at the highest frequencies because they would provide more SNR for point source in these channels, whose purpose was to measure thermal dust emission from the Milky Way. Although at the time of selection, the propagation of modes through multi-mode waveguide structures was understood, nevertheless the complexity of the pixel configuration and its large relative electromagnetic size actually meant that accurate modelling of the field propagation through the entire detector package and its coupling to the bolometer were not possible at the same level as for single-moded packages. Even though the multi-mode pixels as designed proved highly efficient with the expected beam widths, nevertheless the detailed predictions of the beam patterns were not accurate enough to use the models to reduce the data. This was partly due to the manufacturing uncertainties and partly due to the uncertainly in the coupling efficiencies between the front back-to-back horn and the detector horn through the filter stack gap, and also the coupling of the modes to the bolometer, with unknown proportions. The evolution and improvement in the modelling that has taken place since first selection has also been hampered by the difficulty of making accurate ground based laboratory measurements.

The trade-off worked for Planck due to the fact that it is not critical for CMB science to know the beam pattern of the dust channels to high accuracy. At the same time, the high SNR achieved in these channels allows to transfer the CMB-dipole-based calibration from lower frequencies to the highest ones. This will be an important legacy of Planck science to astrophysics.

Lessons learned:

- Although mode propagation in multi-moded waveguides and horns per se is well understood, for HFI-like detector assemblies the accurate modeling of beam patterns is not yet possible. The need for their knowledge must be carefully assessed and there must be a clear approach to measure them on the ground and/or in-flight.

CFRP technology implications

The selection of CFRP technology for the reflectors and structure of the telescope had quite important implications for the performance of the optical system, its verification and

characterization on the ground, prediction of in-flight performance, and measurement (and/or estimation) of performance in flight. These implications were not fully understood when the technology was selected, and in fact the problems encountered measuring the CFRP-based telescope at low temperatures were seriously under-estimated.
The formal consequences were quite serious for Planck as the estimated in-flight geometry of the telescope did not meet formal requirements, and the medium-scale reflector surface characteristics (and small-scale for the primary) could not be measured as initially planned. In fact, the initially planned system-level characterization was unfeasible and there was no time to implement a more suitable one (e.g. based on mm-wave measurements).
Fortunately, the geometrical requirements were found a-posteriori to be tighter than actually required, and the characterization of the optical response could be largely carried out in flight. So the shortcomings of CFRP technology did not result in a significant reduction of the science of Planck. However, future CMB missions may not be in as fortunate a situation as Planck.

Lessons Learned:

- The technology to be used should be selected taking into account real requirements and the verification/characterization possibilities

Optics and scanning strategy

The scanning strategy of Planck did not take into account explicitly any beam-measurement requirements, with the exception of some specific instances which were adopted during operations (higher density sampling of some planets; one instance of spun-up observations of Mars). Generally there was strong consensus within Planck that it was more important to maintain survey stability than to obtain additional information on the beams.
On the other hand, the HFI bolometers exhibited more complex time response than initially expected, and it is significantly degenerate with the "optical" response from the planet measurements. One instance of spun-up Mars observations provided some ability to separate time-response and optical-response, limited to the DC "offset" between the two. Until now it has not been possible to determine in flight the optical main beams of the HFI detectors accurately. This has not been detrimental to HFI analysis for which knowledge of an "effective" response (the "scanning beam") combining time and optical aspects is adequate, e.g. map-making. It is however an issue for all aspects where the true optical response needs to be known, e.g. photometric calibration, estimation of the polarization window function at high accuracy levels. For example, initially the importance of including the far sidelobes in the overall calibration of the Planck data was not realized (its importance depends strongly on frequency). The amplitudes can be predicted with only moderate accuracy from the ground and are in most cases too faint to be measured directly in flight during the surveys.

Lessons learned:

- the scanning strategy should take into account the requirements to measure the optical response of detectors, and if needed make provision for specific observations to achieve this (e.g. for measuring far sidelobes, time response, etc).

Modelling tools

The principal (though not the only) RF/optical modelling tool used for Planck was GRASP. Within GRASP, it is possible to carry out calculations using either PO (Physical Optics) of multi-GTD (multi-ray Geometrical Theory of Diffraction) options. To limit the computational time, multi-GTD is the only practical solution to calculate far-sidelobe patterns at high frequencies. The number of multi-GTD orders used determines the precision of the calculation – for Planck up to seven orders were used, but for some uses it was sufficient to limit to the first 2-3 orders.
The fidelity of GRASP was verified by extensive ground-based testing on the RFQM. However it is recognized that the inputs to flight predictions were not always optimal (e.g. theoretical detector/feedhorn patterns).

GRASP is complex to use and is computationally very expensive, especially at high frequencies and with complex geometries. It should be kept in mind during the design phases that the shape of surfaces impacts the complexity of the model, and in the case of GRASP using simple shapes often implies considerable savings.
Overall a huge investment was made across many parts of the project in modelling the Planck optics using GRASP, and they were used extensively in all phases of the mission - but not as much as wished for, due to its complexity and cost. In particular during the analysis stages, it is very important to have a modelling tool which allows to simulate repeatedly observing with different kinds of beams.
One specific aspect of optical modelling/analysis which was initially under-appreciated was the broad-band nature of the detection system. At least partly this was because GRASP is a monochromatic design tool and it is computationally expensive to extend simulations to many frequencies.

The ground measurements of an RF copy of Planck ("RFQM") show that GRASP (as used by Planck, i.e. using multi-GTD for far sidelobes) is able to model 4π beams up to 350 GHz with an accuracy of a few % in total power (integrated over 4π) - this difference is spread over the entire beam such that around the main beam the prediction is very accurate, but for example peaks in the far sidelobe levels differ from measurements by ~a few dB (depending on frequency). This large uncertainty is due to a number of factors: uncertainties in the RFQM measurements, in the knowledge of the detailed geometry input to GRASP (e.g. alignment, horn corrugations etc), approximations in the use of multi-GTD, and effects not taken into account (e.g. bandpass). Polarization predictions are even more difficult, and the Planck RFQM campaigns were not able to validate the fidelity of the GRASP polarization models - although this was mainly due to poor knowledge of the CATR transmit horns, and there is no reason to think that GRASP is less accurate for polarization than for total intensity. Although all these issues can be improved, it should be clear that ground predictions of the optical patterns are currently not very reliable where sub-percent features are concerned.

On the other hand, Planck has shown that it is possible to determine some of the pattern parameters much more accurately based on flight data, for example the level of far-sidelobe features can be fit to null maps to precisions of order a tenth of a percent in total power.

Lessons learned:

- a tool such as GRASP is needed in all stages of the mission, but it needs to be very agile: which means to run it on efficient enough computing facilities, to reduce the computational time, especially for broad-band models
- Design the geometry of critical optical elements with GRASP in mind (e.g. using simple shapes which simplify the GRASP geometrical model)
- The multi-GTD approach has been validated on Planck but remains complex – it would be useful to reduce this complexity (but note that in early design phases low-order mGTD calculations may be sufficient)
- validate GRASP for experiments focused on polarization
- use verified inputs to GRASP (e.g. detector patterns)
- use of a tool such as GRASP requires experts – ensure that they are available
- Although GRASP is a good and necessary tool (in practice the only one available) for the prediction of beam patterns, its ability to predict sub-percent features in the sidelobes remains limited, especially for polarisation. Experiments requiring sub-percent knowledge need to find ways to use the in-flight data to acquire the necessary knowledge.

Using optical information in science analysis

The path from optical modelling (e.g. using GRASP) and/or measurement (whether on ground or in flight) to using this information for science analysis is quite long and convoluted. In Planck this included transformations from optical beam to scanning beam to effective beam to a scientific representation (window function) and its uncertainties (eigenmodes). The latter part of this chain was developed largely after launch, testing different methods along the way, and was not particularly optimized. The early part of the chain relied almost entirely on GRASP, which is an engineering tool for expert users, not particularly adapted for fast iterations. Onsite expertise is required to achieve this, and was not always available. The overall result is that it was quite difficult and time-consuming to go from end to end of the entire chain. In the HFI, beams were measured on planets. It took the combination of Mars, Saturn, and Jupiter measurements to obtain hybrid beam profiles down to 50-60 dB. Especially after launch, data analysts were confronted with poorly understood beams at low levels, and the need arose for iteratively and rapidly producing different types of beam models and extracting -from the sky data - eigen mode corrections to the transfer function

This was only possible in selected cases, and did limit the range of possibilities which could be explored. Overall one has to be careful with the confidence in the assessment of uncertainties using such models.

Lessons learned:

- Establish early (well before data analysis starts) the complete chain leading from optical modelling/measurement to the use of this information in scientific analysis, and optimize it using simulations. Put in place the ability to experiment rapidly and

iteratively with optical modelling during the analysis phases.
- The ability to extract optical information from the sky data should be evaluated already at a very early stage.

6 POINTING

Planck was a rather unusual system with regard to pointing: a fairly slow spinner (weakly stabilized by the low spin rate), and with the periodic repointings of the spin axis actuated by thrusters. Its main sensor was a redundant star tracker with Time Delay Integration capability (a key feature for Planck) to provide good signal-to-noise on stellar sources moving across the CCD. The on-board software was fairly complex, as

- Attitude estimation had to separate the nutation motion from the angular momentum orientation
- Slewing was a mixture of
  - an open-loop planner which built short series (three) individual actuations to achieve combined repointing and nutation-damping
  - with a triggered repetition of the open-loop planner, as a way to achieve longer slews under a pseudo-closed-loop approach

In addition to the main set of sensors and actuators, Planck carried a Fiber-Optic Gyro (FOG) package which was not meant to be used in flight, except as a self-standing technology demonstrator.

Planck operated autonomously, having only three hours contact daily with the ground. In this sense the FDIR design was quite stringent. It is worth pointing out that the in-flight pointing performance of Planck was very good, and it met all requirements set on the ground (see e.g. Zorita et al. GNC-2011, which also includes some detailed technical lessons-learned).

Thermoelastic effects on the apparent pointing

During the mission, large variations in apparent pointing (~20" peak-to-peak) were observed and eventually the major part of these variations were traced to thermoelastic deformations of the structures supporting the star trackers (not real motions of the spin axis). These deformations were due to switching on-off and/or thermal control of units near the star trackers; seasonal effects also caused variations on 1-year timescale. Most of these effects had not been predicted from the ground, or were much larger than predicted. In flight some of the impact was mitigated by operational measures (e.g. adapting thermal control loop parameters, avoiding switch on-off operations). It is worth noting that it took a long time to measure, analyse and understand these effects, and partial mitigation measures were undertaken well into the routine operational phase (disturbing the long-term stability). Part of the long time until understanding was achieved is due to the thermal timescales involved (e.g. seasonal effects require at least two full seasons to characterize), but also to the difficulty of the measurement and analysis, which exceeded significantly what could be done during the commissioning phase.

The fact that long timescales of stable thermal behavior are required to establish thermoelastic effects on the pointing also implies that any important changes in the thermal baseline must be avoided. For example, switching to the redundant star tracker (which fortunately Planck did not require) would imply an important thermoelastic change and the partial loss of the knowledge acquired until then.
Overall, it seems clear that a satisfactory methodology for understanding the implications of thermoelastic effects on pointing (especially dynamical ones), and the means to manage their impacts in flight, either does not exist or was not made applicable.

Lessons learned:

- Thermo-elastic effects – both static and dynamic - are an important part of system-level performance and they need to be assessed carefully during the design phase. The assessment should include also the impact of FDIR measures (e.g. switching to redundant units), especially wrt star trackers.
- It is very advantageous if thermal loads and control loops are built to be flexible enough to allow adaptation in flight (e.g. allowing permanent switch-on of high power units, or to move from bang-bang to PID control in specific cases).
- Mission planning should take into account that long timescales are needed to measure and analyse subtle pointing impacts of thermoelastic effects.

Poorly understood changes in pointing

In addition to the apparent changes in pointing due to thermoelastic effects, some real changes were observed over the mission, some of which remain poorly explained. Some of these were generically expected since there is gradual depletion of on-board consumables (fuel, Helium), and potentially also migration of consumables (fuel, Helium) between tanks (note that PID-based thermal control of the fuel tanks had been put in place to manage this). However, the observed motions of the spin axis do not entirely fit these causes. It is currently believed that the behavior of fuel in the tanks was significantly modulated by two effects which are very difficult to model and predict: (a) ribbed membranes in the tanks, most probably had complex mechanical behavior over time (folding, with sliding/blocking of adjacent surfaces); and (b) overdamped PID thermal control on the tanks caused time-delays in fuel migration rebalancing following fuel use. It is worth noting that the tank membranes caused significant development and qualification problems due to their potential to develop cracks under folding, but these issues were not at that time linked to the potential consequences of non-linear behavior on pointing in flight.
Additional smaller influences on real spin-axis motions were due to: regular sorption cooler pulses, depletion of Helium and migration between tanks.
Overall it was quite a complex task to measure and diagnose all of the above motions (and to separate them from the apparent thermoelastic-induced motions), and in this process it turned out to be quite useful to be able to use data from the FOG as an independent but also complementary set of data. Unfortunately the FOG, being a technology demonstrator, was not part of the attitude baseline, and was poorly supported in terms of user manual, procedures, and database (e.g. no parameter decommutation nor out-of-limits). Indeed a consequence of this black-box provision was an extended period following a SW reset

where the FOG was running but communication with the onboard computer was not established leading to data-loss.The eventual utilization was inefficient (e.g. the FOG channels need periodic bias calibration, and no plans had been made for this).

Lessons learned:

- Consumables cause gradual changes in pointing which are difficult to predict and diagnose, e.g. non-linear behavior of ribbed tank membranes or fuel migration (in spinners) modulated by thermal control
- It is very helpful to have access to pointing data from independent sensors (Star tracker and FOG for Planck). If such sensors do exist, adequate plans for their utilization should be made.

Science during slews

Planck actively slewed the spin axis, and overall about 9% of the time was spent slewing. Although the instruments were continuously on and gathered data also during slewing time, it was not initially foreseen to use this data. Instead, it was felt that a-posteriori processing might allow to use it at a later time and no particular requirements were set on the attitude recovery during slews (except those derived from the need to determine the attitude on board to calculate the nutation damping manoeuvers). Attempts were made during the mission to improve the attitude knowledge during slews, by implementing different filters and experimenting with the use of FOG data. These attempts required significant effort by the Flight Dynamics team at ESOC, in collaboration with scientific teams, but this was late in the game and as a result the slew data has not yet been used. It is not excluded that this may still happen at a later time, but in any case not while the teams which developed the new filters and hold the expertise are still available. The feeling remains that if this work had been foreseen from the start, features could have been designed into the on-board and on-ground software to simplify it.

Lessons learned:

- It is useful during the design process to consider (as “goals” rather than requirements) the potential use of slewing pointing periods for scientific analysis, and the implications of trying to meet the goals.

Pointing reconstruction work

The driving requirements for Planck were related to a-posteriori reconstruction (knowledge) rather than to in-flight pointing performance. The pointing reconstruction was a result of the work of three groups of people spread over the lifetime of the mission, each producing unique information, and having somewhat different objectives: (a) engineers in industry who designed the on-board attitude measurement and control system, and developed the on-board software generating the attitude data transmitted to ground; (b) engineers in the Mission Operations Centre, who designed the ground filters reconstructing the satellite pointing based on the output of the on-board software; (c) scientific teams who transferred the MOC attitude products to the instrument field-of-view, correlated them with scientific data from instruments e.g. to reconstruct the focal

plane based on planets and other bright sources. There was reasonably good contact between (a) and (b) during the development phases, and between (b) and (c) during the operational phases. Contact between (a) and (c) was almost nonexistent, and this caused some issues e.g. poor understanding of requirements (some of this due to the use of different "vocabularies"), leading to less-than- useful verification tests. More importantly, some of the key data produced by industry in the development phases (e.g. PDFs of on-board attitude recovery errors) was not provided in a form in which they could be used. Contact between (b) and (c) was quite close, constructive, and fruitful, particularly when it came to disentangling apparent and real motions of the spin axis, and developing new attitude reconstruction filters adapted to the scientific needs. However much of the work "required" by (c) from (b) had not been foreseen and was not clearly defined, causing some unhappiness on all sides. It is clear that the quality of pointing reconstruction depends both on the individual and the joint work of all three groups, and fostering interactions between them at all stages of the mission is very beneficial to the final outcome.
It is worth noting that in the case of Herschel, similar work to improve the pointing solution was required after launch – however, the existence of an ESA-based Herschel Science Centre which had a keen interest in the pointing quality, and close connection to MOC, facilitated very much the work.

Lessons learned:

- The quality of pointing reconstruction benefits significantly from close interactions of the AOCS engineers in industry with MOC engineers and with the scientific users of the pointing data, starting from the design phase onwards. Each of these groups have their own specific skills and objectives, and it is crucial to keep them all in sync and to plan adequate resources at the different phases of the mission.

7 OPERATIONS AND SCIENCE GROUND SEGMENT

The term Science Ground Segment (SGS) refers to the system built to ensure that the spacecraft – including the payload - was operated optimally and resulted in data products that could be exploited for science. The Ground Segment consisted of a Mission Operations Centre (MOC, developed and run by ESA), which carried out all direct communications with the spacecraft; the Planck Science Office (PSO, developed and run by ESA, which was responsible for coordination of scientific operations, survey planning and monitoring, and the final data distribution system); and two Data Processing Centres (DPCs, developed and run by the LFI and HFI Consortia, responsible for operation of their respective instruments, and for the generation of the data products). The Science Ground Segment covered all these elements except the MOC, but included the interfaces of PSO and DPCs to MOC.

Distributed design

The initial design of the SGS called for a distributed system, with two IOT/DPCs[1]

[1] IOT = Instrument Operations Team; DPC = Data Processing Centre

(sharing a common infrastructure), a very minimal (initially non-existent) PSO, and a MOC which communicated mainly via downlink with DPCs (as the uplink for this "survey" mission was expected to be minimal). During the development, the connections between the entities evolved significantly, PSO becoming a more substantial and influential unit, the IOT/DPCs sharing less infrastructure, and MOC needing to strengthen more the uplink (as the complexity of the CPV phase was recognized).
Overall, the final distribution of tasks in the operational SGS, and the interfaces between them, worked quite well, but it has to be recognized that the evolution into a working system was not easy and required additional (initially unforeseen) resources, in particular for MOC and PSO.

Lesson learned:

- Distributed SGS systems need a lot of interface definition work at a very early stage to ensure that all functions of the system are adequately carried out. This is traditionally the task of a Science Operations Centre, but work in this sense has to be done even before a SOC is formed.

Importance of CPV and Contingency phases relative to Routine phase

An important goal of a survey mission such as Planck is to maintain the stability of observations and instrument calibration, and therefore to minimize changes in instrument settings once the Routine phase has started. This makes survey instrument operations relatively “simple”, i.e. reduced to monitoring instrument health and stability, and making adjustments only if really needed.
One of the points that was recognized rather late was that Planck operations were more complex during CPV phase and in Contingency scenarios than during Routine operations phase. Initially Planck was often referred to as a “simple survey mission” – especially by MOC - which reflects the early lack of understanding of the situation based largely on the routine phase only. As the operational design became more advanced, it became clear that (a) the CPV phase would be quite long (~3 months), require long-time-scale and delicate operations (cooling down of the cryo-chain, tuning of LFI, etc), interactions between instruments and spacecraft, a complicated imbrication of operations between instruments, and ultimately a lot of real-time control of the spacecraft; (b) contingency scenarios deriving from the instrument data (due to very stringent requirements on data completeness originating from the instrument teams) and needing quick reaction time were difficult to implement, due to the conjunction of a narrow daily contact window (3 hr), combined with delays in the chain of evaluating data quality (IOTs), re-planning observations (PSO), and implementing a new plan (MOC). Re-planning scenarios became important system drivers for the MOC (e.g. the “small-gap recovery” scenario), and required regular transfer of information between DPCs and PSO (the Daily Data Quality Report - DQR, and Weekly Instrument Health Report - WHR) to minimize the long (3-day) response chain. These scenarios, with the assumption that the instruments would not be 100% stable, and high uncertainty on the ability of the full cooling chain to complete the minimum survey, resulted in very stringent requirements on survey gap avoidance during the first part of the routine phase.

In fact, the instruments and spacecraft behaved extremely well in flight and the total mission duration was extended several times. As a consequence the initial requirements on data completeness became obsolete after the first year of operations. In fact, minor contingency re-planning scenarios were applied only twice during the entire mission.

As a consequence, a-posteriori there remains a feeling that the contingency procedures, and periodic information exchanges such as DQR and WHR constituted a lot of "wasted" development effort. However, it is recognized that this feeling is a result of the high stability and reliability of the instruments/spacecraft in flight, which can not be guaranteed a-priori – unstable and/or short-lived instruments could have resulted in a very different situation.
At the same time, it is clear that some of the detailed information provided in DQR and WHR (on data "quality" and instrument "health") was not useable operationally (for example, noise levels or glitch rates per ring), i.e. to trigger contingencies. Even in the early phases of the mission, it would have been feasible and more productive to rely on the IOT's daily assessment (by a human) to trigger any contingencies. And, as soon as the minimum survey conditions had been completed, it would have been reasonable to switch to a simpler operational scheme.

It is also noted that the IOT monitoring (QLA – Quick Look Analysis) of instruments and data quality was to a large extent decoupled from similar functions carried out in the DPCs. This led to duplication of effort.

Lessons learned:

- A generic feature of "survey" missions (as opposed to "observatory" missions) is that operations during CPV phase and Contingencies can be complex and drive the design of the operational SGS. They should be considered in detail during the earliest stages of the project. Operations during Routine phase could be reduced (as done for Planck HFI) to monitoring and instrument adjustments by the IOTs. In such a case, a goal in survey should be to minimize changes of instrument settings.
- For survey missions like Planck, instrument operations during routine phase should be targeted to maintaining stability and minimizing changes of instrument settings.
- Interfaces which transmit scientific information for operational purposes should be carefully designed. It is very difficult to automate the transmission of useful scientific information – if critical decisions have to be made, humans must be involved.
- Operational interfaces should be reexamined once the actual stability and performance of the in-flight systems has been established. There should be enough flexibility to allow simplification of interfaces wherever circumstances allow it.
- If possible, avoid duplication of QLA tools (e.g. IOT and DPC).

The role of the PSO

The Planck SGS was based on the concept that the entire data processing would be carried out by the teams who developed the instruments and would exploit the data scientifically (the Data Processing Centres or DPCs). The data products would be generated by the DPCs and delivered to ESA to be ingested in an archive for distribution to the public. In addition, early on it was believed that the survey nature of the mission implied "simple" instrument operations. Within this line of thinking, it was more efficient for the instrument operations to be carried out by the DPCs in direct contact with MOC (who managed the uplink to the satellite). This scheme was consistent with ideas at that time on operations of survey missions and did not require a Planck SOC (the main function for a SOC would have been to develop and operate the archive near the end of the mission and therefore this function could be delayed to fairly late in the mission lifetime). However, during the development, it became apparent that a significant amount of coordination of (instrument, survey) operations would be required between the two instruments and MOC. Nonetheless, MOC was not suited nor funded to take on this coordinating role, mainly because it requires "scientific" understanding and occasional moderation of the – not always aligned – requirements of the two instruments.

This situation led to the creation and gradual strengthening of the Planck Science Office at ESAC, which developed and coordinated operational interfaces between MOC and IOTs, and provided to MOC the survey-planning information that it needed.
This evolution is now seen as necessary and beneficial. However, as it came rather late in the development, some aspects were under-developed. For example, the PSO had no direct interface to the data delivery from MOC, and had to fabricate a rather sub-optimal one – it would have been more effective for the PSO to have its own "Quick-Look Analysis / QLA" tool, as this could have simplified considerably the "DQR"-type of interfaces.

It is important for a coordinating entity such as PSO to have hands-on experience in the operational aspects and performance of the instrument, as well as in scientific operations (e.g. survey planning) and data processing. For the former, it is most effective if some of the PSO's staff participate deeply in the instrument development and testing process – this is something that was not done enough in Planck, or was lost due to unexpected loss of PSO manpower at critical times. An important opportunity to obtain operational experience is during on-ground operations, both at instrument and at system level. Participation in ground testing also builds important personal links between the teams which will later operate the satellite in flight. Unfortunately PSO personnel were only marginally involved in on-ground operations, and their first real experience came during CPV phase.

Lessons learned:

- Even in survey missions, the coordination of instrument operations, survey planning, and management of the interface to MOC requires substantial resources in an entity (for Planck the PSO) which is "science-knowledgeable".
- A PSO or MOC -like entity should be present during at least some instrument testing, and should also support scientific operations on the ground, to gain hands-on experience.
- If a PSO-like entity requires access to spacecraft and instrument data from MOC for whatever reason, it should build a direct interface to the MOC's data distribution system.

Smooth transition between phases

ESA supported strongly the concept of Smooth Transition between phases, which was to a

good extent based on the required use of SCOS2000 (developed and used by MOC) for operational procedures. This tool was indeed adopted by the instrument teams, although it had a number of shortcomings in terms of documentation, maintenance, version control, etc, which persisted until ESA provided dedicated support[2]. The main problem was that industry was not similarly required to use the same tool, so that the procedures used during system-level testing on the ground were not the same as used during flight.
In general it would have been more beneficial if industry had been more closely connected and gave higher priority to flight operations; as an example, the User Manual was produced and delivered to MOC far too late (one year before launch).

Overall, the implementation of Smooth Transition was considered beneficial, and helped to reduce problems in flight operations. At the same time, it should be recognized that it is impossible to fully achieve seamless transition, and the SGS systems should be flexible enough to adapt to situations where it is less costly to adopt individualized solutions.

Lessons learned:

- The requirement to use a single operational tool for both ground-based and flight operations is very beneficial. However, the requirement should be extended to all parties participating in ground testing and flight operations (industry, MOC, instrument teams), and adequate support must be provided for the use of the tool.

Shared infrastructures

In the initial design of the Planck SGS, it was considered very important that there is more than one end-to-end independent analysis carried out, to be able to cross-check the most important results. The natural division was along instrument lines – since clearly an intimate knowledge of the instruments was an important component of success – leading to two instrument-linked Data Processing Centres (DPCs). In order to be able to compare DPC results meaningfully, it had to be done at many points along the complex data processing pipelines.

To enable this comparison effectively, and to save costs, the initial schema proposed was that the two DPCs would share (the development and utilization of) their computing infrastructure very deeply. The elements of the infrastructure to be shared included: (a) a Data Management Component (DMC); (b) a software language and pipeline environment (the Builder); (c) a "Federation layer" which was designed to make sharing transparent and seamless.

It is fair to say that this ambitious (but ineffectively managed) development succeeded only to a very limited extent, for many different reasons: (a) it was realized rather quickly

[2] It should be understood that using the same tool is not the same as using the same release version of a tool. Practically-speaking it is very difficult to impose a single release across all parties, because (a) the support lifespan of an individual release is finite. AIT work typically starts well in advance of the flight-ops development, and ends many years earlier before the end of operations; (b) AIT and flight-ops operate in different contexts with different needs; (c) managing bug-fixes, priorities and work-arounds in an common way (single release) across users in different organisations with different schedules is challenging.

that instrument expertise could not really be shared, and it would be very difficult if not impossible for one DPC to make meaningful analysis of the other's data at a stage where instrumental effects dominate; (b) the data processing needs and requirements were very different for each DPC, partly due to differences in data characteristics (e.g. total data size, algorithmic approach, etc), and partly to a different "vision" of what were the key data handling priorities (e.g. "global" vs "local") – it meant that it was difficult to agree on a common set of requirements; (c) some of the infrastructure envisioned was very ambitious and relied on new technologies not easily acceptable to coding scientists; (d) some of the key groups developing the infrastructure were not part of the core development teams within the DPCs making communication rather difficult. "Political" priorities also played a role.

At the end of a fairly long development period it became clear that some of the building blocks of the joint infrastructure were not meeting the needs of some of the parts, and that trying to enforce the originally envisaged plan was counterproductive, so much of it was abandoned or eventually simply not used. Simpler interfaces were devised to exchange data between the two DPCs, and only a few features of the common development remained (e.g. a basic simulation tool – "Level S").
In the end, only a small fraction of the infrastructure which was developed in the early stages was actually used. In particular the elements that provided tracking and logging of data and software were considered most useful; whereas "sophisticated" tools such as databases and graphical interfaces were barely or not at all used.

There certainly was a cost to this late change of course: in effect the data from each other's instrument was not thoroughly cross-checked until a fairly advanced stage of processing (i.e. map level), leading to difficulties in basic understanding of the properties of each other's data products (even related to vocabulary, but more importantly to what had been done in terms of cleaning of systematic effects), calibration, beam correction, pointing reconstruction, etc) and therefore integration in scientific pipelines. Some common products, e.g. catalogues of compact sources, were in effect separately produced. A very important impact of the late change in approach was on the production of massive Monte Carlo simulations, which are very computer intensive and crucial for the interpretation of data; an infrastructure independent of both DPCs had to be created to enable credible simulations of the products of both instruments.

At the same time, it is now recognized that an infrastructure that would allow to compare data from the two instruments at a low level (e.g. timelines) would not have been used as much as initially believed, considering that they were subject to different systematic effects and quite different sensitivity levels. And the common simulations are only useful as long as the data are detector noise dominated, whereas in fact they are dominated by systematic effects which are different; this implies that some elements of simulations should be common, but not all.

A-posteriori, it may be said that management recognized too late that the original concept of a common infrastructure would not succeed; on the other side, management did put a lot of effort into trying to make it happen, the potential advantages of which were certainly numerous.

Lessons learned:

- Integration of computing and analysis infrastructure across entities with important differences in scientific objectives and constraints is difficult, and a set of requirements for the common elements should be developed and agreed before it is decided to go ahead. This set of requirements should give enough flexibility to allow for evolving understanding of the data processing needs. It should be kept in mind that such an infrastructure may carry a cost which is higher than the advantages it brings, and overheads should be carefully weighted against the system's advantages. Overall, the Planck experience shows that many factors play a role in this type of mixed scientific/engineering development, and management has to be very flexible to find an appropriate solution.

Interfaces in the SGS

The coordination of operational interfaces between MOC, instrument teams, and PSO, and the actual planning of scientific operations, was a task of PSO. This task was mainly carried out within several topical working groups (ICWG, COWG, etc), under the umbrella of a group where all parties in the SGS were present (PGSSG). The proliferation of working groups was a significant overhead and initially not considered very useful by IOTs. Nonetheless, the level of interactions afforded by the presence of these groups was very detailed, and they provided a good forum for discussion of all operational issues. In particular, the possibility for the instrument operations teams to discuss directly with MOC was seen as very useful and positive. As a consequence, the level of preparation which was achieved was quite good.

Lessons learned:

- Coordination of operational interfaces by a PSO-like entity works well, as long as it enables direct discussions between all parties (MOC/IOTs, PSO/MOC, and PSO/IOTs).

SGS testing

A comprehensive suite of SGS testing was planned for Planck and implemented, from instrument-level tests all the way up to end-to-end tests (including the spacecraft in the cryogenic test facility). The a-posteriori feeling was that these tests were all essential, and if anything, more time should have been spent on some of them. Especially the most comprehensive one (End-to-end tests carried out during the full spacecraft cryogenic testing) took place very late in the development of the mission, but was considered very useful as preparation for in-flight operations. Indeed such tests are essential for "fast-start, hard-life-limit" missions (e.g. cryogenic).

Lessons learned:

- SGS testing, all the way up to tests including the full in-flight operational system, are crucial preparations for in-flight operations.

Instrument operations and mission planning

Although it is clear that there are many similarities between operating instruments during ground tests and operating them in flight, there are also significant differences. There is a feeling that LFI instrument functionality (e.g. which parameters can be tuned, what needs to be monitored and how) was not always developed with operations in mind (in particular where industry was contracted for the development of the readout electronics). Some of the tools used to monitor and analyse instrument data for operational purposes were developed or refined during the CPV and routine phases (e.g. "flagging" of data). In some cases, the tools duplicated – though in a simpler way – functions that were required for scientific data analysis.
Similarly, some of the tools developed by PSO for mission planning, e.g. the Survey Planning and Performance Tool, the POSH, could have been used for related purposes within the DPCs, if common requirements had been adequately identified.
It is possible that some of these "surprises" and/or duplications could have been reduced by bringing together more effectively different parts of the project in the early phases. However, overall the level of preparation for flight operations was very good.

In the end, the most important element in the success of operations is people – in this context it is noted that keeping continuity of key expertise from development to operations teams is extremely important. This needs to be planned and funded in advance.

CPV operations were delicate and required very careful planning to optimize each of the contact periods (which were longer than in routine operations). "Commissioning" activities, which were led by Project, overlapped in time with "CPV" activities, which were coordinated by PSO; even though there was close cooperation between the two, some confusion did exist. Having all the teams on site at MOC was key to achieve good communication and coordination. For diagnostic purposes, it was also very important in this phase to have access to "near-real-time" data even for a limited part of the contact window.

With regard to end-of-life operations, "good practice" with respect to end-of-life disposal trajectories from L2 was not clear at the time of Planck and, in parallel with Herschel, a number of different approaches were considered in the year before end-of-life, including "hard" disposal options like lunar impact. Hopefully a de-facto standard is now established, such that future missions at L2 can avoid performing this trade-off. More low-level aspects of the end-of-life activities, both passivation-related (e.g. preventing FDIR reactivation of RF and propellant tank emptying) and calibration-related (e.g. primary mirror temperature experiment) were finalised finalised very late, in the last months of the mission.
Regarding instrument operations, end-of-life activities were only thought of near the end of the mission. Although not extremely critical, these activities should have been thought through more carefully at an earlier stage.

Lessons learned:

- Try to ensure that instruments are designed to the extent possible with in-flight operations in mind.
- Ensure that the teams developing the instruments and operating them during ground testing participate fully in the preparation and execution of flight operations.

- Maximize commonality of planning/monitoring tools among different parts of the SGS, i.e. instrument operations, mission planning and data processing.
- Ensure that leadership is clear in all operational phases of the mission (if possible avoid time overlap between phases).
- Co-location of all operational teams at MOC during CPV phases is essential to optimize operations in this critical phase.
- Ensure the availability of near-real-time data even for limited periods.
- Plan in advance for end-of-life operations.

Software development approach of DPCs

Together with the plans for common infrastructure, the original concept was to develop DPC software along engineering standards (the "waterfall" approach). Considering that most of the individual scientists writing software were not trained nor really interested in using such an approach, it is not surprising that this was not the way things actually worked. At the same time, it has to be recognized that Planck is the kind of experiment which requires a lot of learning on the job, i.e. understanding the properties of the data and handling it most effectively evolves to a large extent by trial-and-error and through many iterations. Even the most basic assumptions on the data need to be questionable, and may lead to a complete re-writing of existing processing pipelines. In this context, incremental fast-turnaround development and continuous testing are the most appropriate approach, starting with the "difficult" parts and going on to the subtler ones as the understanding increases. At the same time, the generation of "official" data products has to be efficient, verifiable, and traceable.
To allow this, the infrastructure has to be able to accommodate both "creative"/personal data handling, and "official"/controlled pipelines, and the means for features to be absorbed from the former into the latter. Bringing engineers and scientists together in this process from an early stage is very important. A important element is the early establishment of a basic set of tests that must be applied to the data repeatedly, which will evolve as processing advances, and which provides a synthetic view of the current status of the data quality.

Lessons learned:

- The software development approach for DPC-like entities has to be able to accommodate both "creative"/personal data handling, and "official"/controlled pipelines, and the means for features to be absorbed from the former into the latter. Engineers and scientists should work closely together to create and use this infrastructure.

Generation of key data products

The originally proposed concept for data processing was based on the principle that scientific data processing leading to generation of the public data products should be carried out by the same teams who are exploiting it scientifically during the proprietary period. The scientific success of Planck shows that this is definitely the best approach. It would have been in practice extremely difficult and certainly much more costly to achieve acceptable results with an ESA-based SOC.

At the same time one has to recognize that the LFI and HFI Consortia contained many more scientists than actually contributed to the final quality of the data – many of these were interested in specific secondary science which had very minor impact on the actual understanding and therefore the quality of the data products (and it is entirely reasonable that they should do their own science as "reward" for their contributions in e.g. hardware components). Nonetheless, a number of these groups generated specific data products released to a wider community. Which part of "secondary science" will turn out to be interesting is difficult to predict in the early phases of a mission, so the approach to accrete many interested scientists at an early stage is reasonable. Nonetheless, at a later stage the overhead of carrying along a very large collaboration become very significant, and so it is important to be able to separate out clearly which are the core activities and which are secondary ones. Note that "core activities" are certainly scientific ones: most of the actual effort goes into the "creative" side of the development rather than the generation of the data products; this implies that a significant level of computing resources has to be made available to the "creative" side of the core activities.
Planck had from the start a very broad scientific scope and it was difficult to make this separation, although an attempt was made relatively early at Consortium level ("Core Teams"). This type of division may cause some friction, but is essential to distinguish more clearly what is critical activities and software from nice-to-have activities and software, as well as to understand who are the individuals working towards the former and target the infrastructure to their needs.

Lessons learned:

- Generation of data products in a mission which relies critically on understanding instrumental and other systematic effect, is most effectively done where the instrumental knowledge is, and where the data is being scientifically exploited.
- At all stages of a mission, it is important to distinguish which are core activities (i.e. leading to products) and which are nice-to-have ones.

Simulations

Many kinds of simulations are needed for a project such as Planck, of two generic types: (a) very detailed and realistic simulations of specific aspects of a mission (e.g. to understand and quantify systematic effects), which require very specific expertise and can most likely be done in small numbers only; (b) less detailed simulations which are needed in massive numbers (e.g. Monte-Carlo realisations to propagate uncertainties in parameter estimation) and require supercomputing facilities. Both types are crucial in different ways, and the number of variants needed of each is quite large. To be effective, the simulations (code and inputs) must evolve significantly together with the understanding of the real data: they support the data analysis, but must also learn from the results of the analysis.

Simulations are required at all stages of the mission development, even at the earliest stages, to support the instrument development and to trade off instrument vs DPC complexities. Many of the simulations that are now understood to be needed by Planck were anticipated but not enough effort was put into developing them at an early stage. Clearly the simulations of the subtlest instrument systematic effects cannot be anticipated, but they eventually become crucial as they support the most difficult stages of data cleaning.

Planck (with its ~70 detectors) is – by today's standards – quite a modest instrument, and yet the amount of supercomputing resources that it has used is huge. Even so, they are barely enough to accommodate Planck's needs, and they constitute a limitation on the quality of some of the scientific results. Not only are the resources needed large, but the management effort to utilise them effectively is also very important. State-of-the-art instruments with many thousands of pixels will require much more in this sense than Planck, and this is something that needs dedicated planning and effort.

Lessons learned:

- Scope simulations realistically at an early stage, and dedicate adequate resources and management to this effort. Allow significant resources to the evolution of needs along the lifetime of the mission.

Mission extensions

Planck's routine phase was initially planned to carry out two full sky surveys over one year. In the end, due to very good and stable performance of the payload in orbit, Planck acquired data for HFI over almost 5 surveys, and for LFI for more than 8 surveys. Planck science benefited hugely from these extended phases of the mission. The redundancy over many surveys and several years allowed multiple data consistency tests, and was crucial to understand the systematic effects on the data to the level where it could be calibrated and cleaned down to the detector noise level. It seems fairly safe to say that without the additional redundancy, the quality of the data and of the scientific results would not be as good.

It is clear that although the notion of redundancy was built from the start into the Planck design, the scientific importance of the one-year timescale had initially not been recognized, in spite of the fact that it is the only one which provides equality of observing conditions. At the same time, the complexity of the cryogenic chain implied that the inherent uncertainties on the total lifetime were large, and making strong requirement on long operational timescales could have increased the total cost beyond feasibility. On the other hand, the fairly short observing baseline implied harsh requirements on avoiding observing gaps which were costly to implement (e.g. the "small gap recovery" which became a mission driver for the SGS). A system-level tradeoff on lifetime taking into account such considerations was not carried out on Planck.

However, once again, the excellent performance of the payload in orbit worked very much to Planck's benefit, and ESA's system of incremental extension requests/approvals allowed the necessary flexibility for Planck to benefit from this. Nonetheless, it would be useful for similar survey missions (especially in a context of diminishing resources to fund competing mission extensions) to recognize from the start the benefits to be gained from mission extensions beyond a minimum baseline, to trade this off against other system-level constraints, and in any case to prepare for longer lifetimes by including "goals" in addition to "requirements".

Lessons learned:

- Survey missions which are susceptible to subtle systematic effects whose understanding benefits from long observing timescales should evaluate at system level the costs and benefits of minimum vs longer timescales.

- In any case, the SGS should plan for the longest timescales of benefit, even if only as "goals".

Data products distribution

The mechanism for distribution of Planck data products to the public is the Planck Legacy Archive (PLA), developed by ESA's Archive team (SAT) under the direction of PSO. Although the PLA was initially foreseen only to provide data to the public, following the reduction in scope of the common DPC infrastructure, it was requested by the instrument teams that it be used earlier for distribution of internal "exchange" products within the Planck Collaboration.
In fact it was never widely used for internal distribution, probably because within each DPC there were internal means of distribution which were more conveniently accessible – however it was used for distribution to the Consortia people not present in Core Teams or DPCs. The early development also provided early practice with internal delivery and ingestion of data products, which was quite useful for the development of the public PLA.
The development of the PLA started quite early (well before launch) and was operational in 2009. Some tension was present between DPCs and PSO regarding the level of detailed definition of product formats – this is understandable from the point of view of the DPCs who did not know exactly what the products would contain, but it posed problems for the developers of the PLA who needed to adapt their code to frequent – and occasionally last-minute - changes. Overall it seems that a reasonable compromise was found since – in spite of a number of non-compliances – interface issues never became a cause for delays in any release – however, this was at least partly due to higher-level delays on the product delivery
The validation level of the release products on both sides of the interface also seems to have been satisfactory: only a few faulty products slipped through the release process and had to be corrected. Devising a 100% guaranteed validation system would certainly have been much too costly.

Following the recommendation of a review board, a User Group was formed in 2009 to advise the development of the Archive. However, because of confidentiality rules within the Planck Consortia, the members of the Group could not be selected from the community at large (i.e. the eventual users of the PLA) but instead were internal to the Consortia. The amount of influence of the User Group on the development of PLA was very limited, to a large extent because resource and infrastructure limitations gave it very little room to maneuver. The User Group stopped working – mostly it seems from lack of interest on all sides - shortly after the release of the ERCSC. Following the first major release, there was an attempt to form an external User Group, but it foundered because of lack of a suitable Chair. In summary, a User Group could have been very useful for Planck (e.g. see below on the choice of interface technology), but the effort was not given enough importance and died by inertia. It is likely that if the Group had been made up of external members of the community, it would have had a higher profile and exerted more influence.

The choice of technology for the online interface to the PLA (java applet) was the one being used at the time to develop several ESA archives selected, and had been selected by SAT based on a wide variety of reasons. This choice went against the recommendation of the PLA User Group, which stated preference for a much simpler web-browser ftplike system (à la “Lambda”), which however could not have met all the requirements set on the PLA. In addition, the SAT needed to maintain commonality with other archives and stuck to its choice. A-posteriori, it appears that the User Group was correct in its basic advice against the java applet, as the community clearly disliked using it. On the other hand, a simple “Lambda-like” system would not have been able to cope with the huge quantity and variety of Planck products. In the end, the PSO and SAT managed – after the first release but in time for the second one – to implement a new browser-based technology which suited the community much better; this new technology was selected not only for PLA but following a detailed “market study” covering ESA archives in general. It should be noted that this change had not been foreseen and required significant additional resources.

The Explanatory Supplement (ES) is a DPC deliverable accompanying the data products. Assessments of the quality of the ES vary from “adequate” to “awful” (with weight on the former). The perceived deficiencies are probably related to the fact that each of the Planck data releases was accompanied by a set of scientific papers – some of these papers describe how the data products were generated and characterize them, and some of them describe science results. The contents of the former type of papers overlap heavily (but not entirely) with the contents of a “classical” ES, and the authors were mostly the same. This overlap did cause confusion considering that the contents of the ES were never defined in detail. Papers naturally received high priority within the Planck Collaboration, but the ES got very little attention until the release was imminent, at which point there was a mad rush to finish it off. At that time in the project, the pressure to go ahead with a release is very high, leading to delivery and acceptance of a sub-optimal ES. After the release, there is little incentive for the DPCs to improve the already-public ES. As a consequence, it is not too surprising that the ES contains barely the minimum required information, which the user must supplement with consultation of various papers. Overall, the information needed by most users of Planck data is available but dispersed and not always easy to find.

The ES is implemented as a wiki, which differs from the traditional document-based solution. This choice was made considering that there were at least three sources contributing to the ES (two DPCs and the PSO), in many cases to the same product; and that it would help to keep the ES as a “living” document. Furthermore, the web nature of a wiki means that it is easy to embed cross-linking to other sources of information (e.g. the papers). The adoption of this solution did imply that PSO should be the overall coordinator of the contents, which confuses to some extent who delivers and who receives this product. The wiki solution has worked reasonably well, both in development and as a delivery mechanism, though the presentation is in some areas deficient (this depends on the specific wiki engine selected), and the cross-linking is not as extensive as originally envisaged. The question whether this is a good long-term solution, i.e. well after the last Planck release has taken place, remains open.

The above discussion highlights the fact that the Planck data releases depended heavily on the readiness of scientific papers. Even though the data may have been ready for release, it was very difficult to carry out the release if the papers were not ready in time. The main reasons were that the ES relied in part on the papers; and that the Consortia naturally tried to preserve the high profile of their scientific results. Release dates are established in the Science Management Plan many years in advance, and further constraints come from the wish to make a "splash" in the media, and even from self-organized conferences. All these constraints contribute to very hard deadlines, and pressure within the collaboration mounts to very high levels as these critical times approach. In general, it is not good to cumulate constraints, and even better to have backup plans in case deadlines cannot be met. For example, the possibility to release data even if not all scientific papers are ready needs to be discussed and planned for well in advance of any release deadline.

Lessons learned:

- If there is more than one way to distribute data products, people will always choose the one "closest to home". Avoid duplication of distribution systems.
- It is very useful to start the development of the Archive very soon, including the definition of interfaces (formats). However, enough flexibility has to be kept (by both deliverers and recipients of data products) in the system to allow some of the contents to change, even at the last minute.
- Ensure that product validation procedures are present on both sides of the delivery interface, and that each side is aware of what these procedures are. The validation does not need to be 100% fail-safe.
- It is important to establish a User Group to advise the development of an Archive such as PLA. However, it needs to represent the user community unquestionably (and it would be best if it reports to an independent body), it has to be supported adequately, and it has to be given the ability to have a real influence in the development.
- The choice of interface technology is critical in terms of acceptance by the users of the Archive, who generally speaking prefer the simplest possible solution (in contrast to the software engineer's optimal "flavour of the day"). Since such technologies are in continuous evolution, it is important that the Archive maintains the ability to evolve and respond to what the users want.
- It is important to make a clear difference between scientific papers and the Explanatory Supplement. The latter is a deliverable whose contents should be defined in the same timeframe and in as much detail as the products it accompanies. The data products should not be considered delivered until the accompanying ES is also completed and delivered.
- Provision to the public of an ES via a web-based mechanism e.g. a wiki, is feasible and can work well if planned and coordinated adequately. Pay attention to who coordinates the effort, and to the long-term future of the ES.
- Hard deadlines for data releases are necessary and useful for many reasons, but it is also important not to add more constraints than are strictly needed, and to discuss well in advance what are the options if deadlines cannot be met.

8 MANAGEMENT ASPECTS

Planck was an ESA medium-sized science mission and was managed in the "normal" ESA way. There are many things which could be said about how it was managed, but most of these are due to features peculiar to Planck or to individual management styles, and which do not lead to useful "lessons learned". Nonetheless, some aspects are listed below which might be of use to other similar experiments.

One aspect which was almost unique at the time of Planck's selection was the large size of the two Consortia put together to develop and deliver to ESA the two Planck instruments and generate the data products. They each included dozens of institutes funded by many different national agencies, each with their own management approach, technical and funding constraints, etc. This heterogeneity caused many difficult problems that needed resolution at an almost day-to-day level. The Herschel/Planck Project Manager (PM) instituted regular meetings of the Consortia PIs and PMs together with representatives of the major funding agencies, which reviewed the development status at frequent intervals and resolved inter-agency problems on the spot. This is considered to have worked very well as a problem-solving mechanism during the development phase.

The Planck Science Team, which included representatives of the two instrument consortia and from the reflector provider consortium, acted as a similar problem- and conflict-resolution entity for all matters related to science and the SGS. Its effectiveness probably derives from the fact that it was led by ESA, which is considered a "neutral" party.

The general lesson learned is that it is important to institute a forum where all involved parties are represented, can voice their opinion, and where conflicts can be resolved by an unbiased party.

As for all large projects, documentation is a heavy burden, and the status of the project is regularly reviewed via documentation packages. A very large number of documents are generated over the lifetime of a project like Planck, and it is very likely that many of them have not been reviewed in detail or even read in some cases; this seems to be "the way it is" and it is difficult to think of ways to control the flow adequately. Conversely, it does happen that useful information is hidden in the dense forest of documents, and this situation could be improved. A common repository of documents, to be used by all parts of a project (at ESA, industry and the scientific Consortia), would be an immensely useful tool. It is understood that for good reasons, individual elements of a project do not make everything visible to all other parties; but it is certainly feasible to achieve many levels of privacy within a single repository. If a single repository is not feasible, the next best thing would be to maintain a single comprehensive document tree/index, where any member of the project could at least locate the existence of potentially useful documents and request them from other parties.
For a centralized repository or index to work, it is essential that "rules" are put in place that allow any document to be found and traced; such a system has been partly implemented within the Planck Consortia using the Livelink software, with rather limited success – probably because lack of metadata has resulted in some (many ?) documents promptly being "lost", and because sub-groups have tended to use "their own" isolated storage systems. For real success, it seems essential to (a) implement a tool that everyone in the project is required to use; and (b) nominate a central "librarian" whose only job is to keep order in such a system, and more generally make it work.

Tracking people is similarly important in as large a project as Planck. A central database of people, accessible online, has been used for many years within the Planck Collaboration, and is used to manage their basic information (including e.g. data access rights, authorship rights, etc). Such a tool has been found to be very helpful.

Communicating information within a large Project is important and not easy, and the two Consortia have both instituted wikis accessible online to their own Consortia, where individuals can informally post any kind of information, from "ideas" to "reports", and the rest of the group can read and comment them. This has been found to be a very effective way to diffuse complex information in a fairly painless way. It is highly recommended to other projects. Once again, transparency of such systems to all members of a project is highly recommendable.

Wikis do not substitute for direct discussion, and in a project such as Planck it is extremely important to make everyone aware of the status of key issues. Each of the two Consortia have held for many years a ~2-day meeting of a largish group of "core" people, typically at intervals of two months, where all critical issues are discussed, and major common decisions are made and/or communicated. This has been found to be an effective way to keep the Consortia well informed, and to keep them involved the main decision-making processes.

A key fact in the life of Planck has been the data release milestones, which in effect constitute hard deadlines that the entire project has to meet. In order to accommodate these deadlines, it has been very important to agree a detailed schedule of events leading up to them. It is key for everyone in the project to be aware of this detailed schedule, which in Planck is discussed and agreed at highest (Science Team) level. A particularly important date is the time at which a data set becomes "frozen", i.e. it becomes the official basis for the scientific analysis which is described in the papers accompanying the. As discussed elsewhere in this document, the papers must be ready for the release to take place, and there is a minimum amount of time necessary to actually complete the analysis release (in Planck this is typically ~6 months). It is extremely important that the freeze date is considered adequate and reliable by the Consortia, i.e. it provides enough time for the papers to be ready by the release, and the data set is not changed in the meantime. Although the data will continue to evolve in parallel to the science analysis, it must always be clear which is the data set which should be the basis of the next release. In Planck there has been an occasional tendency to let, the data set drift, causing uncertainty, distress and in some cases inability to meet deadlines. This is a rather risky practice which should have been avoided.