January 31, 2002

Charting the Winds that Change the Universe, II

The Single Aperture Far Infrared Observatory (*SAFIR*)

G. H. Rieke, D. J. Benford, P. M. Harvey, C. R. Lawrence, D. T. Leisawitz, D. F. Lester, J. C. Mather, G. J. Stacey, M. W. Werner, and H. W. Yorke

Abstract: "*SAFIR* will study the birth and evolution of stars and planetary systems so young that they are invisible to optical and near-infrared telescopes such as NGST. Not only does the far-infrared radiation penetrate the obscuring dust clouds that surround these systems, but the protoplanetary disks also emit much of their radiation in the far infrared. Furthermore, the dust reprocesses much of the optical emission from the newly forming stars into this wavelength band. Similarly, the obscured central regions of galaxies, which harbor massive black holes and huge bursts of star formation, can be seen and analyzed in the far infrared. ***SAFIR*** will have the sensitivity to see the first dusty galaxies in the universe. For studies of both star-forming regions in our galaxy and dusty galaxies at high redshifts, ***SAFIR*** will be essential in tying together information that NGST will obtain on these systems at shorter wavelengths and that ALMA will obtain at longer wavelengths." – page 110, Astronomy and Astrophysics in the New Millenium, National Research Council - AANM, National Academy Press, 2001.

1. The Role of the Far IR/Submm

Winds and flows in the interstellar medium convert a potentially static scene into our mysterious and fascinating Universe. A supermassive black hole lurks unseen until gas collects into a central accretion disk and spirals in, causing an active galactic nucleus (AGN) to blaze up. Galaxy collisions spray stars in intriguing patterns, but the fundamental consequences arise from the ability of the interstellar medium (ISM) to lose angular momentum and collapse to fuel nuclear starbursts. Stellar populations everywhere are established and renewed by the formation of new stars in molecular clouds. The heavy elements that shape stellar evolution and make life possible are transported by interstellar material to the sites of star formation, awaiting incorporation into new stars and planets.

Regardless of the original emission process, cosmic energy sources glow in the far infrared and submm. The continuum emission is due to the incredible efficiency of interstellar dust in absorbing visible and ultraviolet photons and reemitting their energy. The appearance of the early Universe, of AGNs and starbursting galaxies, and of star forming regions is transformed through suppression of the visible and ultraviolet and augmentation of the far infrared and submm. Low-lying far infrared fine structure lines are the major coolants for interstellar gas. Molecular transitions in this spectral range carry the signature of conditions in warm and dense interstellar clouds where stars form. Thus, we must look in the far infrared and submm for clues to the underlying processes shaping the origin, structure, and evolution of our Universe.

Accessible advances in technology can produce huge advances in our capabilities for far infrared and submm astronomy. Consequently, the consensus-based Astronomy and Astrophysics Survey Committee recommended ***SAFIR***, an 8-m class, space-borne far infrared telescope (operating from ~ 20μm to ~ 1mm), as a high priority to be started in this decade (page 10, AANM). The scientific importance of this mission spans both the Origins and Structure and Evolution of the Universe science themes within OSS. The FAIR concept for a large FIR/submm space telescope was recommended in the last

Origins roadmap with the same goals and community endorsement as ***SAFIR***. As an observatory-class instrument, ***SAFIR*** (and FAIR) would serve a broad range of science needs.

With a start near the end of the decade, ***SAFIR*** follows several missions that will lay its scientific and technological foundation. SIRTF, whose mission will be complete by the time ***SAFIR*** would go to Phase B definition, will leave a Great Observatory legacy of mid and far infrared observations along with the technical demonstration to make a large cold telescope a compelling next generation facility. SOFIA will have operated for a number of years, providing both a testbed for detector technology and a high level of flexibility and accessibility in its explorations of the far infrared sky. Herschel will be completing its examination of the submm range, both for continuum sources and with powerful spectroscopic instrumentation. However, none of these missions will approach the potential for this spectral range, SIRTF because of its small aperture (85cm) and SOFIA and Herschel because of the thermal background due to their relatively high operating temperatures (~ 220K and ~80K respectively). Yet ***SAFIR*** is well in reach technically, given the modest requirements on its optics and pointing, the infrastructure under development within NASA for large telescopes and large satellites in general, and the potential for dramatic advances in detector technology at modest cost. ***SAFIR*** will lay the foundation for the next steps in exploring this spectral region, such as the SPECS space-borne interferometer (see "Probing the Missing Universe" by D. Leisawitz et al.).

NGST and ALMA will begin operation toward the end of the decade. The timely ability of ***SAFIR*** to bridge the spectral gap between them will complement well the capabilities of both facilities, maximizing the productivity of two other major investments in astronomical capability.

2. *SAFIR* and the Formation and Evolution of AGNs

It appears that central supermassive black holes are a universal component of galactic bulges. Do the central black holes form first and serve as condensates for galaxies? Or do they build up as galaxies grow and merge? The low lying H_2 lines at 17 and 28.2μm are one of the few conceivable ways to study warm molecular gas condensations prior to the formation of metals, for example molecular gas around primordial massive black holes. A number of processes, such as formation of a small number of stars, can heat molecular clouds above the ~100K threshold for high visibility of these lines. The lines are undetectable together from the ground until $z > 35$ (both must be detected to confirm the identification). ***SAFIR*** will be well suited to searching for them. Line widths and profiles would indicate whether the central mass is highly compact (suggesting a black hole), or if the molecular cloud is just in a mild state of turbulence (as expected if it is self gravitating without a central black hole).

At the current epoch, galaxy mergers produce huge far infrared fluxes through a combination of violent starbursts and of AGNs associated with their cental black holes. “Distinguishing starbursts from supermassive black holes is complicated by the fact that AGNs are often shrouded in dust, so that much of the direct emission is hidden from view. Long wavelengths penetrate the dust more readily, so .. ***SAFIR*** and NGST with an extension into the thermal infrared are .. suitable for separating the two phenomena" (page 85, AANM).

What happens during the much more common mergers that build galaxies in the early Universe? COBE showed that the far infrared/submm energy density in the early

Universe is comparable to that in the visible/near infrared. What are the relative roles of dust embedded AGNs and starbursts in producing this luminosity? Do AGNs at high redshift differ in basic properties from nearby ones? Models of the cosmic X-ray background indicate that the great majority of AGNs at high redshift are heavily absorbed (Gilli et al.; Comastri et al.). Thus, these answers must be sought in the far infrared where optical depths are low (ISM optical depths are similar at 20μm and 6kev and rapidly decrease at longer infrared wavelengths and higher X-ray energies).

The fine structure lines of NeII (12.8μm), NeIII (15.6μm) and NeV (14.3μm) are the best tool to distinguish unambiguously whether the ISM of a dusty galaxy is ionized by a starburst or by an AGN. Figure 1, based on work by Voit and Spinoglio and Malkan, is a demonstration. Not only are the line ratios very well separated, but their extinction is reduced by more than a factor of thirty compared with the visible. At the epoch of peak quasar activity, these lines will be redshifted to the 45 to 55μm range. A 8-m far infrared telescope would have both the necessary resolution and sensitivity to use this tool to determine the relative roles of star formation and nuclear activity in the early Universe.

The full suite of infrared fine structure lines probes a very wide range of excitation energy, allowing ***SAFIR*** to constrain the UV spectra of AGNs and extending work with the Infrared Space Observatory (ISO) on a few nearby Seyfert galaxies to large lookback times. In addition, many of these lines have relatively high critical densities (up to ~ 10^{10} cm^{-3}), so they have a unique ability to probe the density of the gas around AGNs.

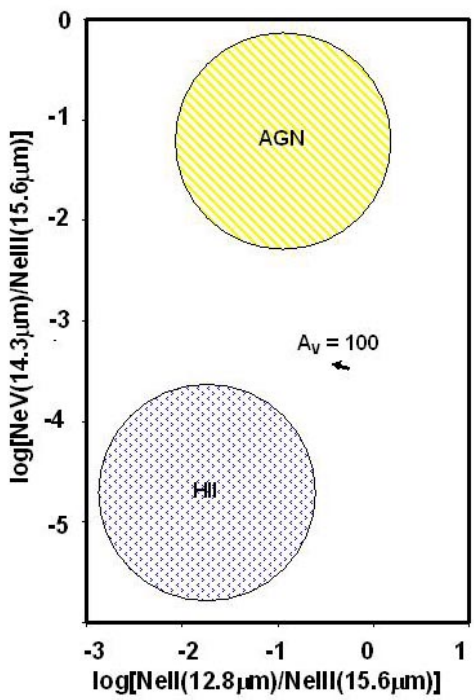

Figure 1. Ne V is not excited appreciably by hot stellar spectra, but is produced by the hard UV spectra of AGNs. Hence, the line ratios plotted distinguish the two types of ionizing source unambiguously. The plotted ratios are virtually extinction independent; the short arrow shows the effect of 100 magnitudes of visible extinction on the line ratios. Even this level of extinction has little effect compared with the separation of excitation mechanisms.

The angular resolution of ***SAFIR*** is a critical contribution to these studies. Figure 2 shows the Hubble Deep Field and the X-ray sources discovered in a deep Chandra exposure. A portion of the HDF is degraded to 1" resolution, the beam diameter of a 8-m telescope operating at 40μm. Individual galaxies are well isolated for detailed study.

3. *SAFIR* and the Emergence of Stars and Galaxies

The history of star formation determines the evolution of galaxies and the generation rate for heavy elements. It has been traced by deep Hubble Space Telescope (HST) imaging followed up with large groundbased telescopes. However, even at modest

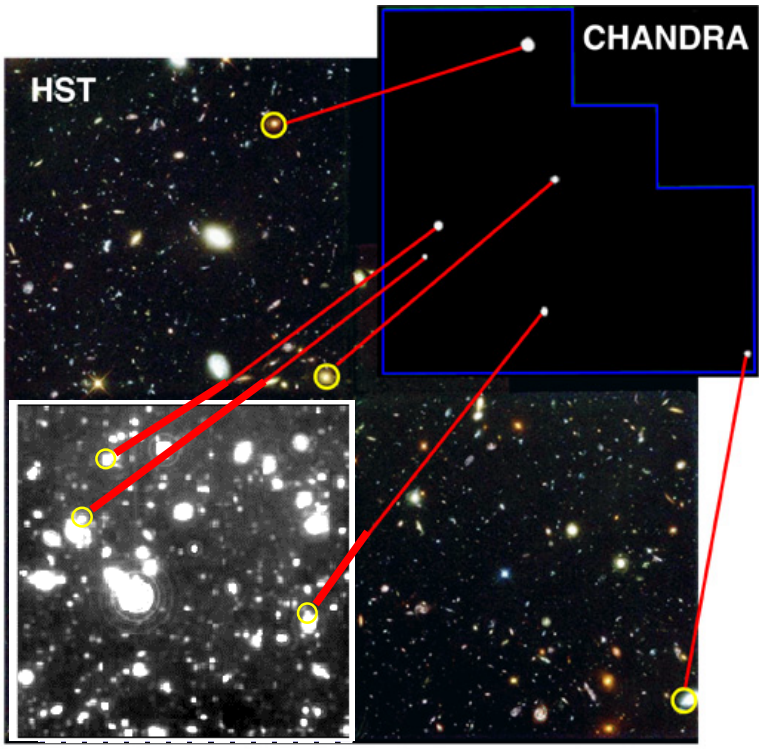

Figure 2. Deep Chandra detections of X-ray sources in the HDF. In the degraded portion, the resolution has been reduced to the diffraction limit of a 8m telescope operating at 40μm. The individual galaxies, including the X-ray identifications, are readily isolated at this resolution

redshifts, these techniques only probe the rest frame ultraviolet. Interstellar dust can absorb nearly all the UV in star forming galaxies. In the best-studied starburst galaxies such as M82, a debate raged for more than a decade regarding how to correct even the near infrared emission for interstellar extinction. Such corrections are poorly determined for galaxies at high redshift. Consequently, there are significant uncertainties in the star forming rate for $z > 1$.

These uncertainties can only be removed by measuring the far infrared emission emitted by dust heated by young stars in these galaxies. The importance of this approach is underlined by the large cosmic far infrared/submm energy density discovered by COBE. This background has been partially resolved by ISO in the very far infrared and is thought to arise from starburst galaxies at z up to 3. A 8-m telescope would resolve most of this high redshift background into individual galaxies, thus showing the dominant phases of dust embedded star formation and nuclear activity throughout the Universe.

Ultradeep optical images (e.g., Hubble Deep Field) reveal many galaxies too faint to contribute significantly to the submm diffuse background. A full understanding of star formation in the early Universe requires that we extend far FIR/submm measurements to these small systems and possible galaxy fragments. In this luminosity range and over $1 < z \leq 5$, ALMA and other groundbased submm telescopes are mostly sensitive to infrared cirrus emission and the output of cold dust that are not necessarily heated by recent star formation. The rate of star formation in modest galaxies for $1 < z \leq 5$ can best be determined through high sensitivity imaging from 20 to 200μm. Combining ***SAFIR*** and ALMA measurements of SEDs can give photometric redshift estimates, and ***SAFIR*** spectroscopy can measure redshifts using the strong PAH features in galaxy spectra near 8μm. The sensitivity of ***SAFIR*** will allow us to measure ordinary galaxy luminosities to below L_* , even out to $z \sim 5$. Equally important is the resolution afforded by an 8-m aperture; previous space far infrared missions like SIRTF will be limited in this science objective by confusion noise.

4. *SAFIR* and the Dynamical and Chemical Evolution of Galaxies and Stars

How do the first gas clouds form? What chemical processes occur within them and how do their characteristics change as the first traces of metals are injected into them by stellar processing? Once even traces of metals have formed, the C^+ line at 158μm becomes very bright. Its luminosity in nearby spiral galaxies is typically a few tenths of a percent of the entire bolometric luminosity of the galaxy. Although this line is partially accessible in the poor atmospheric windows between 300 and 700μm, it will be routinely observed from the ground only at $z \geq 4$, when beyond 800μm. N^+ lines at 122 and 205μm also play important roles in cloud cooling. Study of the molecular hydrogen and these emission lines in the early Universe and as a function of redshift promises to reveal many of the processes occurring in the gas clouds that build early galaxies. Space-borne observations in the FIR/submm must be a major component of this study.

The far infrared fine structure lines also control the cooling of molecular clouds in the Milky Way. Understanding this process and related ones revealed by far infrared spectroscopy is a key to advancing our knowledge of how these clouds begin their collapse into stars and planets (see below).

5. *SAFIR* and the Birth of Stars and Planetary Systems

Stars are born in cold interstellar cloud cores that are so optically thick they are undetectable even in the mid infrared. In about 100,000 years, a young star emerges, ejecting material along powerful jets and still surrounded by a circumstellar disk. The subsequent evolution is increasingly well studied, but the birth of the star has occurred hidden from view. How does the cloud core collapse? How does subfragmentation occur to produce binary stars? What are the conditions within protoplanetary disks? When, where, and how frequently do these disks form planets?

The birth of stars and planets can be probed thoroughly at FIR/Submm wavelengths. Imaging with the resolution provided by a far infrared 8-m telescope ($\leq$ 100 AU at 40μm for the nearest star forming regions) can probe the density, dynamics, and temperature structure of these ~1000 AU collapsing cores on critical physical scales. In addition, 100 AU resolution would reveal the steps toward binary formation. Far infrared polarimetry is a powerful probe of magnetic field geometries, both for studying core collapse and mapping the fields that must play an important role in accelerating and collimating jets.

As it collapses, the gas in the core is warmed until its primary transitions lie in the FIR/Submm. Spectroscopy in molecular lines such as H_2O and the J>6 high series lines of CO, as well as in FIR atomic lines, can probe the physical and dynamical conditions in the collapse. The spectrum predicted for a collapsing cloud core is shown in Figure 3. The [OI] lines have narrow components from the infalling envelope and broad ones from outflow shocks. They are the main coolant of the gas in the intermediate regions of the cloud. Bright H_2O lines between 25 and 180μm are the dominant coolant in the inner cloud, where a broad component is expected from the accretion shock and a narrow one from the disk. The CO lines from 170 to 520μm are the main coolant for the outer cloud; warmer CO from within the cloud can also be studied because of velocity shifts due to the collapse. This suite of lines therefore would allow us to probe the process of star birth thoroughly.

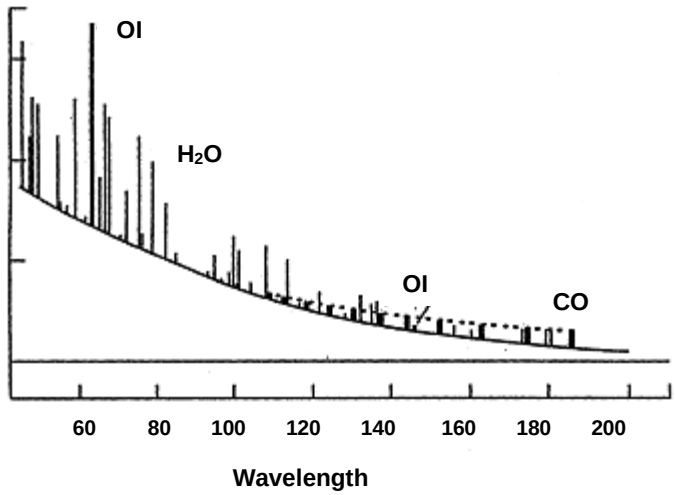

Figure 3. Predicted far infrared spectrum of a collapsing cold cloud core, from Ceccarelli, Hollenbach, and Tielens. The spectrum is dominated by [OI] and complexes of CO and H_2O.

6. *SAFIR* and the Evolution of Planetary Systems and the Origin of Life

What were the conditions in the early solar nebula, as the protoplanetary disk formed and planets and small bodies accreted out of it? All the bodies in the inner solar system have been so heavily processed that they no longer reflect clearly the conditions at their formation. The discovery of many small bodies in the Kuiper Belt outside the orbit of Neptune gives access to objects where accretion proceeded slowly and its products should be primitive and still reflect conditions in the early solar nebula.

There is a large population of Kuiper Belt Objects (KBOs), including objects of size rivaling the largest asteroids. They have a broad variety of surface characteristics. To interpret the clues they provide for evolution of the solar system requires that we understand how this variety of surface chemistry has come about. Two very important parameters are: 1.) the albedoes of the surfaces (important to help identify the substances that cover them); and 2.) surface temperatures (both to help understand what chemical reactions can occur and to determine the escape rates for different molecules). Both of these parameters can be determined in the far infrared, through measurements of the thermal emission. It is for this reason that the 1998 National Academy of Sciences study on "Exploring the Trans-Neptunian Solar System" placed a very high priority both on large, far infrared telescopes and on development of high performance far infrared detector arrays.

The Kuiper Belt is thought to be the source of short period comets and hence has a central role in the comet impacts that brought water to the earth and made life possible here. However, most traces of this process have been erased by time. How can we understand the conditions that regulated the early formation and evolution of the KB and its release of comets toward the inner solar system? The Infrared Astronomy Satellite (IRAS) discovered debris disks around Vega, β Pic, and other stars, with evidence for inner voids that might have resulted from planet formation. Many more will be discovered by SIRTF. The Kuiper Belt is similar in many ways to these systems and should be interpreted as the debris disk of the solar system. Taking an example, β Pic is thought to be only about 20 million years old. Transient and variable absorptions by the CaII H&K lines in its spectrum have been interpreted as the infall of small bodies from the debris system (e.g., Beust et al.). This system contains fine grains that heat sufficiently to be detected in the mid infrared and scatter enough light to be seen at shorter wavelengths. Because it should be drawn into the star quickly, this fine dust may have been produced in

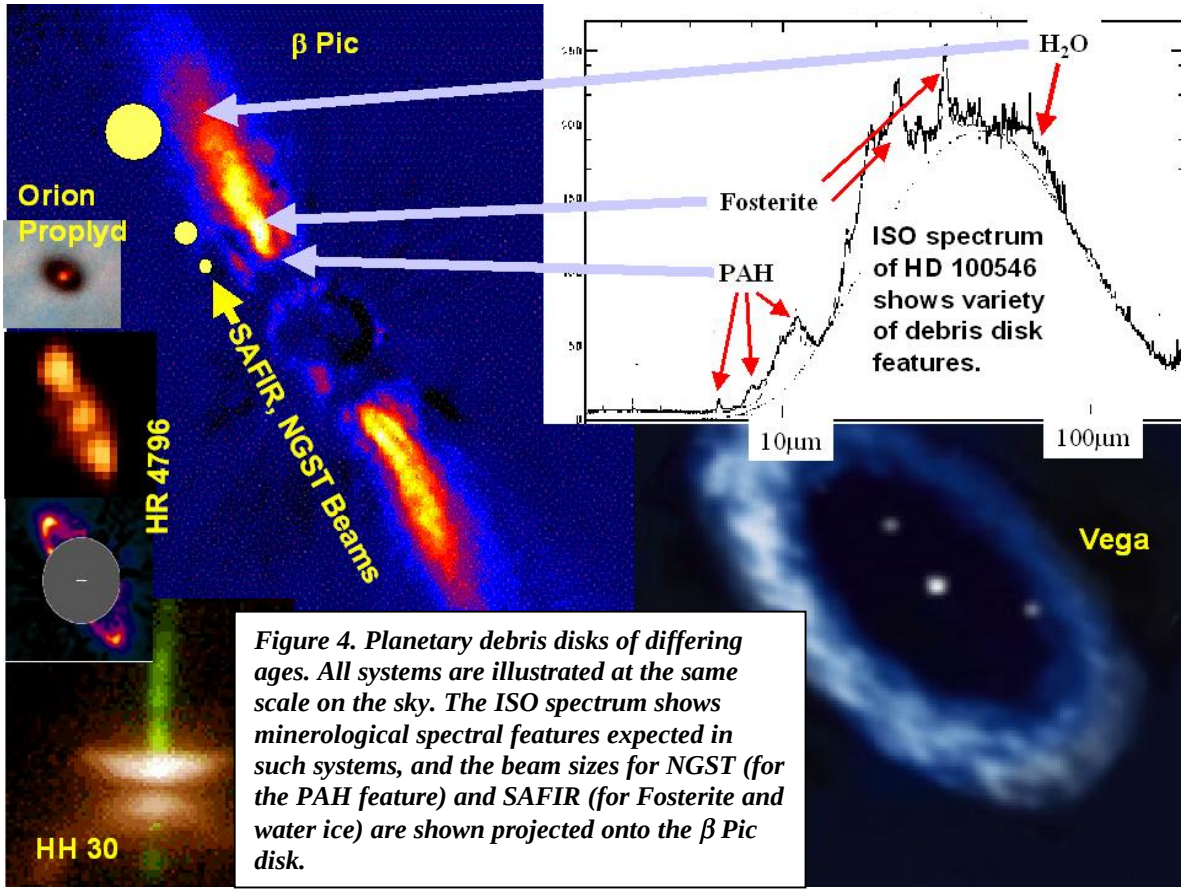

Figure 4. Planetary debris disks of differing ages. All systems are illustrated at the same scale on the sky. The ISO spectrum shows minerological spectral features expected in such systems, and the beam sizes for NGST (for the PAH feature) and SAFIR (for Fosterite and water ice) are shown projected onto the β Pic disk.

recent collisions between planetesimals. Thus, this system and others like it demonstrate the potential of examining the early, violent evolution of debris disks and the infall of comets.

Debris disks are bright in the far infrared, where they can be imaged to identify bright zones due to recent planetesimal collisions, as well as voids. The radial zones sampled will vary with wavelength, from a few AU near 20μm to hundreds of AU in the submm. Figure 4 illustrates the potential advances with ***SAFIR***. Spatially resolved spectroscopy with such a telescope could probe the mineralogy of the debris disks in the 20 - 35μm region where the Infrared Space observatory (ISO) has found a number of features diagnostic of crystalline and amorphous silicates, and can locate ice through its 63μm emission feature. Giant planets similar to Jupiter and Saturn could be detected to compare their placement with the debris disk structure.

7. *SAFIR* and the Discovery of New Phenomena

Technological advances enable astronomical discoveries. Harwit tried to quantify this relation in "Cosmic Discovery." In the 25 years preceding publication of the book, important discoveries were made within 5 years of the development of new technology making them possible. The exceptional discovery potential in the FIR/Submm region arises because the sensors are still substantially short of fundamental performance limits and the telescopes available to date have been very modest in aperture (less than 1 meter!).

The previous decadal survey developed a parameter to describe the discovery potential of new missions, which they called astronomical capability. This parameter is proportional to the time required to obtain a given number of image elements to a given sensitivity limit. ***SAFIR*** will have astronomical capability exceeding that of past far

infrared facilities by a factor of about 10^{10}, and will still offer a gain of about 10^5 after SIRTF and Herschel have flown. A gain of 10^5 is similar to the gain from the initial use of the Hooker 100-Inch Telescope on Mt. Wilson to the Hubble Space Telescope.

8. Mission Development

8.1 Telescope

With the imminent selection of the NGST prime contractor, it is timely to begin mission concept studies for ***SAFIR***. There are two general possibilities, as indicated in Figure 5. The development of the NGST telescope may result in approaches that can be readily adapted to the far infrared, with the differing requirements of (1) colder operating temperature; (2) relaxed image quality; and (3) larger aperture (now that NGST has decreased in size to 6m). However, these three important differences may lead to unique architectures for the far infrared telescope. This basic decision must be made as soon as possible to guide further development of the mission. Although autonomous deployment is a possibility, opportunities for reduced cost and risk through in-space assembly should also be explored. Finally, there is a high level of interest in the far infrared and submm both in Europe and Japan, making it timely to consider possible international collaborations. Such cooperation has already been fruitful in the infrared and submm for Herschel, Planck, and ALMA.

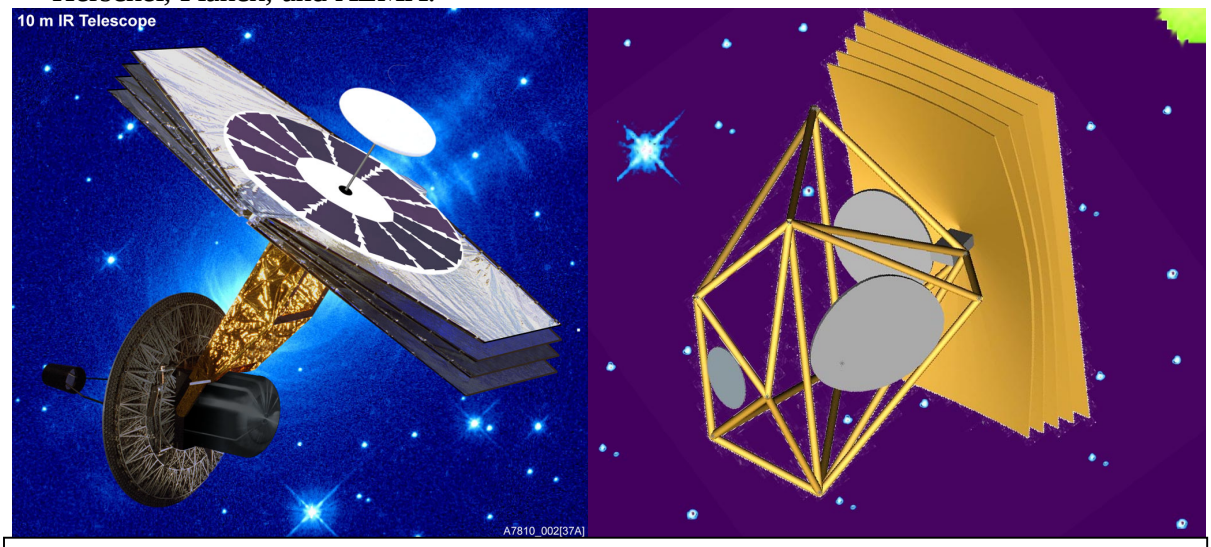

Figure 5. Two possible development paths for SAFIR. The figure to the left illustrates the potential for a telescope based on NGST developments, in this case placed at about 4AU to obtain greater radiative cooling (courtesy Ball Aerospace). The figure on the right illustrates that focused developments for the far infrared may also be promising. In this case, the telescope uses a stretched membrane approach that may offer a lower construction cost than NGST-based telescopes (courtesy M. Dragovan).

8.2 Detector Technology

The far infrared and submillimeter ranges have benefited relatively little from investments in detector technology by non-astronomical pursuits. In this regard, they differ dramatically from the visible, near and mid-infrared, and radio regions. Detectors in those spectral regions closely approach theoretical performance limits. For example, in the visible, CCDs have quantum efficiencies greater than 90%, read noises of about two electrons, and formats including many millions of pixels. In the far infrared, the much smaller prior investment has left the possibility for orders of magnitude further progress

toward fundamental limits. NASA missions are the best customers for this technology, and an augmented NASA investment will return substantial benefits to ***SAFIR*** and other far infrared and submillimeter missions such as Explorers. These investments will guarantee our nation's leadership in this important technology.

Figure 6 illustrates the three major detector technologies. Each has current strengths and weaknesses. Far infrared photoconductors are the most advanced in array construction, as shown by the space qualified SIRTF array in the figure, and require relatively modest cooling. However, they fall somewhat short of theoretical limits in potential performance and respond only up to the excitation energy. Development should address larger arrays, at least 128x128. Bolometers have broad spectral response and are the most advanced submm continuum detectors. They require extremely low operating temperatures. Development needs to emphasize improved array technology, such as SQUID-based multiplexing, and superconducting-thermometer bolometers that interface well to SQUID electronics. Hot electron bolometer mixers provide the best heterodyne operation above the superconducting gap frequency of NbTiN, around 1200 GHz. They can have large advantages for spectroscopy over photoconductors and bolometers. Development needs to address reducing noise temperatures and developing support electronics to allow large scale spatial arrays.

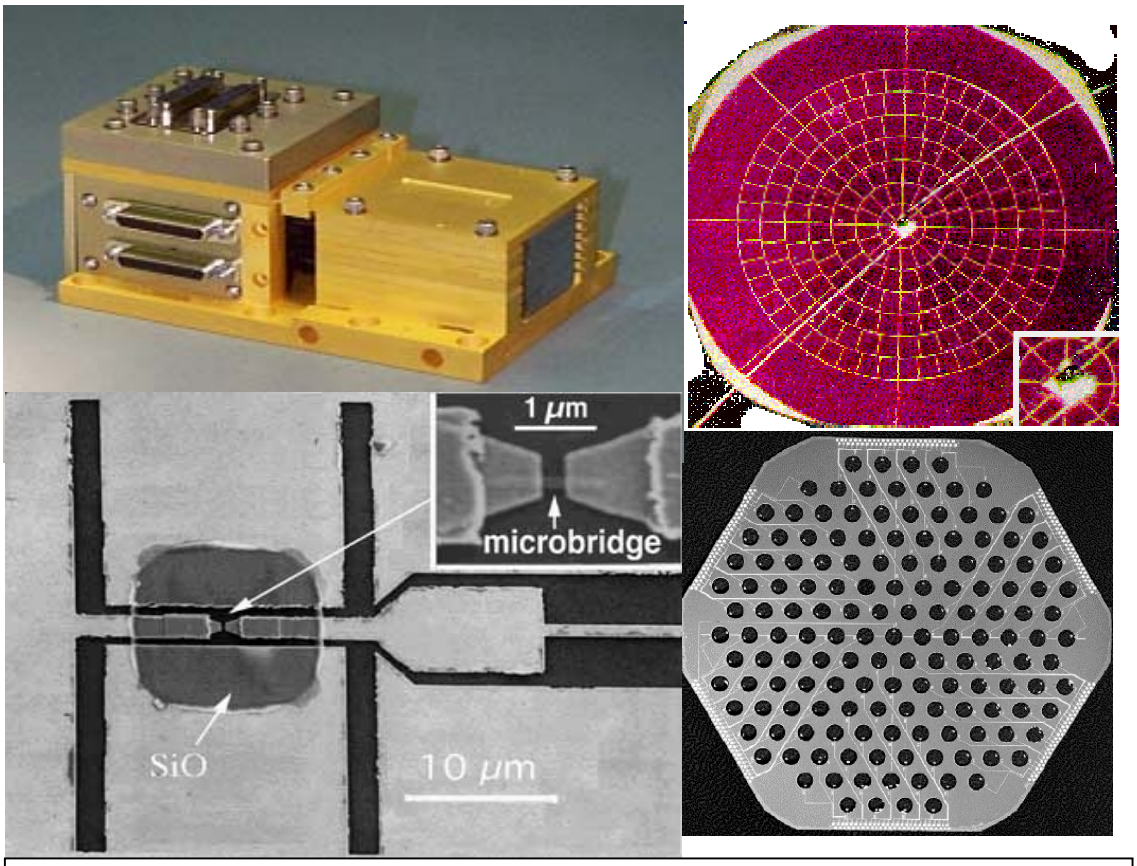

Figure 6. Far Infrared and Submillimeter Detector Approaches. Clockwise from upper left: (1) the SIRTF 32x32 Ge:Ga far infrared photoconductor array; (2) a spiderweb bolometer element; (3) an array of spiderweb bolometers; and (4) a hot electron bolometer mixer.

8.3 Budget

Goddard Space Flight Center carried out an estimate of the budget for ***SAFIR*** for the UVOIR panel of the decadal survey. They drew on their experience estimating the cost of NGST, so the comparison of the two missions is also pertinent. Their results are in Table 1. They assumed that no additional development would be required beyond that for NGST, although the report indicated that this was probably not entirely correct. We allow here for a significant development program, perhaps even departing significantly from the NGST telescope architecture. In the spirit of the above estimate, we take this program to be half that for NGST, or an additional $125M, for a total cost of $620M. For comparison, the estimate of the UVOIR panel for NGST is $1114M.

Construction (assumes an ESA instrument)	$310M
Launch (using a new mid-sized EELV)	$85M
Science and Mission Operations (5 yrs)	$100M
Total	**$495M**

Table 1. Cost estimate for SAFIR from the decadal UVOIR panel report.

Augmentation to budget for far infrared detector arrays	$10M
Very low temperature refrigerators for space environments	$50M
Large, lightweight optics	$80M

Table 2. Technology investments relevant to SAFIR and other far infrared/submillimeter projects.

The decadal survey committee also recommended a budget over the decade for the technology development that would support ***SAFIR*** and other projects in the far infrared and submillimeter, as shown in Table 2.

9. Summary

"***SAFIR***…will study the relatively unexplored region of the spectrum between 30 and 300μm. It will investigate the earliest stage of star formation and galaxy formation by revealing regions too shrouded by dust to be studied by NGST, and too warm to be studied effectively with ALMA…. It will be more than 100 times as sensitive as SIRTF or the European [Herschel] mission….To take the next step in exploring this important part of the spectrum, the committee recommends ***SAFIR***. The combination of its size, low temperature, and detector capability makes its astronomical capability about 100,000 times that of other missions and gives it tremendous potential to uncover new phenomena in the universe." – pages 39, 110 Astronomy and Astrophysics in the New Millenium, National Research Council, National Academy Press, 2001.

SAFIR can contribute substantially to both the Structure and Evolution of the Universe and the Origins themes of NASA space science, through realizable technology developments of a moderate scale. With clear science priorities and exciting science goals that are intellectually accessible to the greater public, SAFIR is a mission that would help engage the nation in shared exploration. It will also showcase new technology in dramatic images capturing events at the dawn of time and at the birth of stars and planets.