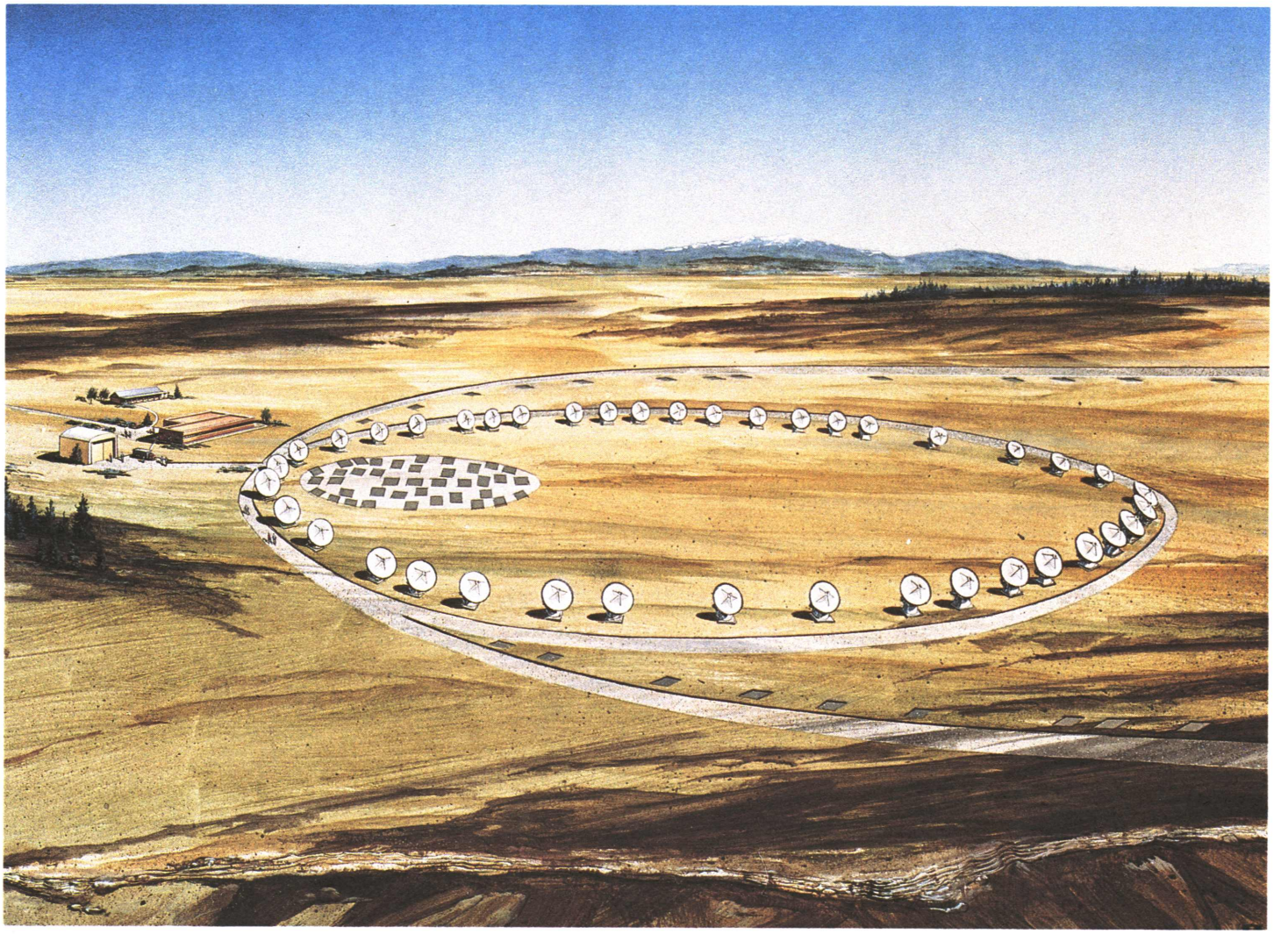
THE MILLIMETER ARRAY



NATIONAL RADIO ASTRONOMY OBSERVATORY

Operated by Associated Universities, Inc., under cooperative agreement with the National Science Foundation.

The Millimeter Array

Proposal to the National Science Foundation

Submitted by

ASSOCIATED UNIVERSITIES, INC.

July 1990

Robert E. Hughes, President
Associated Universities, Inc.

TABLE OF CONTENTS

I. INTRODUCTION

Millimeter-wave astronomy has had an important impact on virtually all areas of contemporary astronomy. It provides a testing ground for our theories of stellar evolution, galactic evolution, and the evolution of the universe itself. The chemistry and composition of the interstellar medium, the earliest stages of star formation, and the internal kinematics of luminous galaxies are uniquely revealed by observations at millimeter wavelengths. Recognizing the promise and potential of millimeter-wave astronomy, many nations and international consortia have built, or are constructing, modern telescopes and support facilities for millimeter-wave research (e.g., West Germany, France, Spain, the Netherlands, Great Britain, Sweden, Canada, and Japan). In the U.S., in a field pioneered by astronomers with but modest resources, no major national millimeter-wave instrument has ever been funded.

In 1983, the National Science Foundation's Astronomy Advisory Committee formed a subcommittee, chaired by Professor Alan Barrett, to offer guidance on the priorities for future development of millimeter- and submillimeter-wavelength astronomy. Among the subcommittee's three principal recommendations was a request for a design study for a national millimeter-wavelength synthesis array.[1] This particular recommendation was an encouragement that the NSF build on the seminal interferometric millimeter-wave work being done with the Owens Valley Radio Observatory's (OVRO) three-element interferometer and the Berkeley Radio Astronomy Laboratory's Hat Creek interferometer. These instruments have demonstrated the wealth of scientific insight that is possible with observations taken at high angular resolution. The Barrett subcommittee saw the opportunity, and pointed to the need, for a millimeter-wavelength synthesis array capable of producing high-quality images. The array, they noted, should have the following properties:

(1) One-arcsecond resolution at 115 GHz;
(2) 1000–2000 m^2 total collecting area;
(3) Good imaging capability at 1 mm.

The design of an array intended to fulfill these needs began at the NRAO in 1984. The resulting conceptual design was presented to a scientific workshop of approximately fifty astronomers held in Green Bank, West Virginia, in September 1985. The strong scientific endorsement given for the Millimeter Array (MMA) at this workshop, and at subsequent MMA design meetings held in 1987 and 1989,[2] placed particular emphasis on the array's imaging capability. The MMA should be a true imaging array: It should combine high brightness

[1] A. H. Barrett, C. J. Lada, P. E. Palmer, L. E. Snyder, and W. J. Welch (1983), *Report of the Subcommittee on Millimeter- and Submillimeter-Wavelength Astronomy*, National Science Foundation Astronomy Advisory Committee.

[2] A complete list of the participants at MMA scientific and technical design workshops is given in Appendix C.

Table I-1. The MMA — Summary of Instrumental Parameters

Array —	
Number of Antennas:	40
Total Collecting Area:	2010 m^2
Angular Resolution:	$0''.07\lambda_{(mm)}$
Antennas —	
Diameter:	8.0 m
Precision:	$\lambda/40$ at 1 mm
Pointing:	1/20 beamwidth
Transportable	
Configurations —	
Compact:	70 m
Intermediate (2):	250 m, 900 m
High Resolution:	3 km
Frequencies —	
Emphasis on:	195–366 GHz
Broadband Coverage:	30–50 GHz
	68–115 GHz
	130–183 GHz
Versatility:	Simultaneous multi-band operation
Site —	
High-Altitude—suitable for precision imaging at 1 mm	

sensitivity, equivalent to that of a large single antenna, with the angular resolution of an array telescope. Frequently the objects to be observed are many times larger than the telescope beamwidth; this means that many overlapping regions must be observed and the individual images merged into a single mosaic image. Construction of such mosaics is practical only if the array is fast and has good 'snapshot' imaging capability (i.e., good instantaneous u-v plane coverage). Many antennas are required in order to achieve this capability. Finally, the MMA study groups have noted that new science—not possible with any existing instrument—will follow if the MMA is situated on an excellent high-altitude site and if it comprises a total collecting area at least as large as 2000 m^2, a sensitive high-frequency capability, and baselines that allow sub-arcsecond resolution.

This set of requirements defines a versatile imaging instrument with the following capabilities:

- Sub-arcsecond imaging at 115 GHz and higher frequencies;
- Wide-field imaging, mosaicing;
- Rapid imaging, 'snapshots' of high fidelity;
- Sensitive imaging at high frequency (to $\sim$ 350 GHz);
- Simultaneous broadband, multi-band operation;
- Comprehensive 'single-dish' capability.

High sensitivity implies that the total collecting area of the array should be made as large as possible, while fast imaging is achieved by distributing the

collecting area over many elements. Sub-arcsecond imaging constrains the array dimensions. Sensitive imaging at high frequency demands that the MMA be situated on a high-altitude site with excellent atmospheric transparency.

Translation of these general requirements into the specific parameters of the array design is described in the sections that follow. Considerations which led to their definition, summarized here in Table I-1, include the following: The total collecting area of the MMA should exceed that of any existing millimeter-wave single antenna. The MMA should be reconfigurable, as is the VLA, so as to permit study of both extended and compact emission regions with proper u-v sampling. Four scaled configurations, of dimensions 70 m, 250 m, 900 m, and 3000 m, are planned. The maximum extent of the array, 3 km, corresponds to $0\farcs1$ angular resolution at 230 GHz. Finally, all antennas will be equipped with total-power systems so that, in several modes, the full collecting area of the array can be used as though the MMA were a single antenna, for those observations requiring this capability.

The Millimeter Array combines the sensitivity provided by the collecting area of a telescope fifty meters in diameter with angular resolution, $< 0\farcs1$, equaling that of the Hubble Space Telescope (HST) and operates at frequencies at which thermal emission processes illuminate the sky. This unprecedented combination of sensitivity and angular resolution at short wavelengths will make available for astronomical investigation a wealth of unique opportunities and new science. Scientists using the MMA will:

- Image the redshifted dust continuum emission from evolving galaxies at epochs of formation as early as $z = 10$;
- Reveal the kinematics of optically obscured galactic nuclei and QSOs on spatial scales smaller than 100 pc;
- Assess the influence that chemical and isotopic gradients in galactic disks have on star formation and spiral structure;
- Image heavily obscured regions containing protostars, and protostellar and pre-planetary disks in nearby molecular clouds, with a spatial resolution of 10 AU and kinematic resolution < 1 km s^{-1}.
- Detect the photospheric emission from hundreds of nearby stars in every part of the Hertzsprung–Russell (H–R) diagram.
- Reveal the crucial isotopic and chemical gradients within circumstellar shells that reflect the chronology of stellar nuclear processing and envelope convection;
- Resolve the dust-formation region and probe the structure of the magnetic field in stellar winds;
- Establish the relative distributions of the large number of complex molecular species in regions of star formation, relating them to shock fronts, grain disruption, and energetic outflows—information which is essential to the understanding of astrochemistry;
- Obtain unobscured sub-arcsecond images of cometary nuclei, hundreds of asteroids, planetary atmospheres and surfaces, and solar regions of active particle acceleration.

Chapter II gives an overview of the MMA project. The MMA scientific program is described in Chapter III. The instrument, its various technical components, and operations and construction costs are detailed in Chapters IV through X.

II. OVERVIEW OF THE MILLIMETER ARRAY

1. The Array Concept

The Millimeter Array (MMA) is a unique instrument. It is the only synthesis telescope conceived and designed as a complete imaging instrument capable of measuring accurately all spatial frequency components, from zero to the maximum array baseline. Partially, the motivation for such a complete instrument comes from the spatial complexity of the sky at millimeter wavelengths and the interrelationship of astrophysical phenomena on a wide range of spatial scales. (For example, do sites of nascent star formation—cold cores—within extended GMCs evolve from, or initiate, specific local changes in the cloud chemistry?) And partially, the motivation derives from the experience of imaging with existing synthesis-array telescopes, all of which have identifiable limitations. The design of the MMA incorporates remedies for the deficiencies recognized in existing synthesis arrays. Thus the Millimeter Array provides those capabilities that will be essential for continuing progress in millimeter-wavelength astronomy during the next decade and beyond.

The Millimeter Array is a fast-imaging telescope situated on a high-altitude site and optimized so as to provide high-fidelity images in the λ 1 mm (200–360 GHz) spectral region. The MMA features flexible spectral-line and continuum capabilities in the atmospheric windows between 9 mm and 0.9 mm. The expected MMA sensitivity results from a combined collecting area of 2000 m^2 achieved with forty antennas of 8-m diameter. The antennas are transportable, and the MMA is reconfigurable into any of four configurations, whose dimensions are set by the need to match the Array's angular resolution to a wide range of astrophysical investigations. Imaging resolution at λ 1 mm will range from $0''.07$, for the largest 3-km configuration, to $3''$ for the compact, 70-meter configuration (with the latter configuration providing essentially complete u-v sampling).

As an instrument specifically designed to recover information on all angular scales, the MMA necessarily includes a wide variety of capabilities which find application in specialized circumstances. Among these is the capability for astronomers to use the array antennas as individual elements to make total-power measurements, either independently or by summing incoherently all the total-power signals. This capability is unique to the MMA: each of the MMA antennas is complete as a 'single-dish' instrument—by virtue of its total-power measurement capability—and may be used as an independent telescope, as described in Appendix B.

Mosaiced images of large objects will be constructed using both total-power and interferometric observations of sub-fields. With a sensitivity of 1 mJy for a one-minute integration in continuum and 30 mJy per spectral-line channel in spectroscopy, large fields can be observed with only minutes devoted to each

pointing, or small sources can be observed with a combination of high resolution and sensitivity much greater than has been possible before. Moreover, the electronics and the correlator will have frequency-agile capabilities allowing observation of different spectral regions simultaneously, as well as tunability over wide bandwidths in each atmospheric window. Such versatility, sensitivity, and angular discrimination will provide astronomers with the proper tool to extend the pioneering millimeter-wave research of the last quarter-century to an unprecedented level of sophistication.

Here we summarize the properties of the MMA, its various sub-components, and its imaging capabilities. These and other aspects of the Array are discussed in more detail in subsequent chapters.

2. Configurations, Resolution, and Imaging Properties of the Array

The specification of forty antennas, each eight meters in diameter, and the choice of array sizes determine many of the imaging properties of the MMA. Those that depend upon the spatial distribution of antennas in each configuration will be discussed in Chapter III. Briefly, three principal criteria drive the specification of the Array dimensions; details of the imaging procedure are discussed in Chapter IV.

(1) The need to mosaic large fields rapidly and recover the information in each field at all relevant spatial frequencies leads to the requirement for a tightly packed configuration. One possible such configuration is shown in Figure II-1: here the antennas are arranged in three dense rings of maximum diameter 70 m. Such a configuration provides an *instantaneous* beam which is nearly Gaussian and which has very low sidelobes, owing to the large filling factor (greater than fifty percent for the array shown here).

(2) Since the Millimeter Array will be used by scientists in the study of astrophysical problems that honor no wavelength boundaries, the angular resolution of the MMA must be competitive with that of the premier instruments with which it will be in contemporaneous operation. The Hubble Space Telescope sets the standard: $0''.1$ angular resolution. To achieve $0''.1$ at the 'workhorse' frequency of the MMA, 230 GHz, requires an array configuration with maximum dimension 3 km. Suitable imaging large arrays are discussed in Chapter IV.

(3) The CO lines will provide the principal tools for a wide range of kinematic and morphological studies. Indeed, the CO lines are to the study of molecular gas as the 21-cm line is to the astrophysics of atomic gas. Further, unlike the single 21-cm line, observations of the multiple CO rotational transitions available to the MMA will enable one to separate abundance effects from effects resulting from excitation or optical depth. Such studies depend upon a comparison of images of the various CO transitions—CO ($J = 1$–0), CO ($J = 2$–1), and CO ($J = 3$–2)—which are made more precise

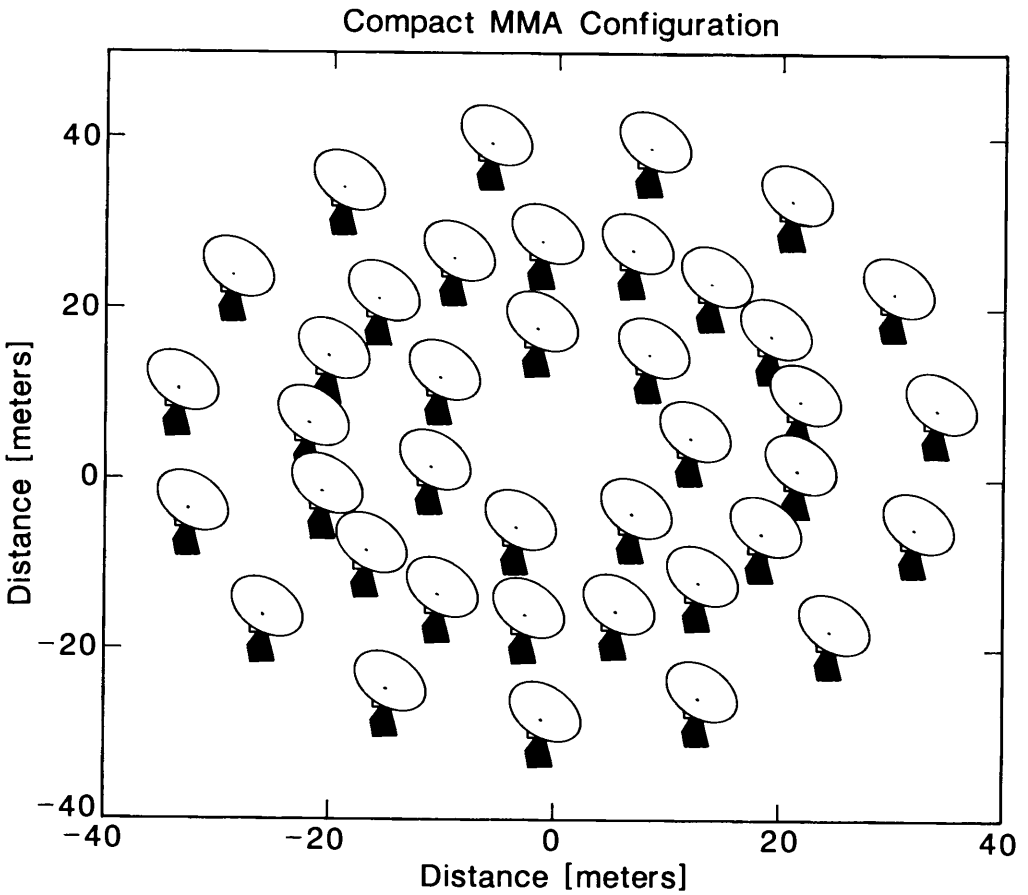


Figure II-1. A possible arrangement for the antennas of the MMA in its most compact configuration.

when images made with about the same resolution are compared. MMA configurations intermediate between the compact 70-m and extended 3000-m configurations are needed. Three scaled-array configurations are planned in addition to the compact array. Having established the maximum array dimension, 3 km, two smaller configurations of approximately 250-m and 900-m extents will provide the range of spacings necessary for CO excitation studies and similar comparative imaging research. The 250- and 900-meter configurations will be circles or ellipses, the u-v coverage from which will yield a beam shape very close to a $J_1(x)/x$ function.

The approximate field of view of each antenna, as given by the half-power beam width, is $\theta_{\mathrm{HPBW(ant)}} = 31'' \lambda_{\mathrm{(mm)}}$. The typical 'size' of an image, in pixels, is set by a minimum of twice the ratio of $\theta_{\mathrm{HPBW\ (ant)}}$ divided by the image beam width, $\theta_{\mathrm{HPBW}} = (0\farcs25/B_{\mathrm{(km)}})\lambda_{\mathrm{(mm)}}$, and is independent of wavelength. In Table II-1 we summarize the main image characteristics of each configuration of the MMA in terms of resolution, angular field of view, minimum image size (in pixels), and the computational image size used in computing images. The latter,

together with the numbers of frequency channels, polarizations, and mosaiced sub-fields, sets the magnitude of the computing problems for the data for each array configuration.

Table II-1. Image Properties of Each of the MMA Configurations

Array Size B	Image Resolution ($\theta_{\rm HPBW}$)	Minimum Image Size	Computational Image Size
70 m	$3\farcs5\lambda_{\rm (mm)}$	18	32
250 m	$1\farcs0\lambda_{\rm (mm)}$	62	64
900 m	$0\farcs28\lambda_{\rm (mm)}$	225	256
3 km	$0\farcs08\lambda_{\rm (mm)}$	750	1024

The parameters of Table II-1 correspond to an observation of a single field $(0\farcs25/B_{\rm (km)})\lambda_{\rm (mm)}$ in diameter. If the diameter of the object to be imaged is N times the size of a single field, the image must be computed using data from observations of roughly $4N^2$ fields or pointings. The requirements for both mosaic imaging and the general MMA imaging problems are discussed in Chapter IV.

Finally, the Millimeter Array may be used, in several modes, as a total-power instrument. In order to realize the goal of making the MMA a complete imaging instrument capable of measuring all spatial frequencies from zero to the maximum baseline, each of the MMA antennas will be equipped with a beam switcher (either a nutating subreflector or a tertiary mirror) for differencing observations taken at different sky positions. Frequency switching or load switching with a chopper-wheel would also be options. The MMA total-power capability, unique to the design of the MMA and not found on any existing millimeter-wavelength or longer wavelength interferometer, makes it possible for the astronomer to use the Array in its compact configuration as though it were a total-power single antenna. The MMA *single-dish* capability can be thought of in three principal modes; these modes are summarized below and amplified in Appendix B.

First, one may accumulate incoherently the sum of the signals from all forty antennas. For measurements of the brightness of an extended object (line or continuum) where the angular resolution of an 8-meter antenna is no limitation, the incoherent sum from the output of all forty MMA antennas will provide a sensitivity greater by a factor of $\sqrt{40}$ (≈ 6.3) greater than the brightness sensitivity of a single antenna (of any size!). In addition to being the most sensitive high-resolution telescope in the world, the MMA, equipped with total-power instrumentation, also is the world's most sensitive low-resolution telescope. The brightness of very extended objects could be imaged (for MMA 'preview' purposes, for example) either by rapidly rastering the entire array to a large number of positions or by pointing each antenna to a different position and recording the results individually. This is the mode in which 'zero-spacing' information would

be obtained, and it is also a mode appropriate for detection programs involving objects of low surface brightness.

Second, the forty MMA IF outputs can be summed coherently into one output signal, producing an adding interferometer. The Array would have the point-source sensitivity of the entire collecting area (2000 square meters) and also the angular resolution corresponding to the maximum interferometer baseline. However, it would be sensitive to emission coming from only one synthesized beam area in the sky, rather than from a solid angle the size of the primary beam of an 8-meter antenna. Thus, just as for the corresponding mode of the VLA, this mode would be good for VLBI and for very-high-time-resolution work, but not good for imaging. Its principal advantage is that it minimizes the computing needed to obtain the brightness in the central pixel; this is probably not a decided advantage for the MMA (as it is for the VLA).

The two 'single-dish' modes described above are, in fact, limiting cases of a multi-dimensional continuum of capabilities that a synthesis array such as the MMA provides, and provides simultaneously. The fixed relationship between resolution and antenna diameter (at a particular frequency) that applies to single-antenna observations becomes a variable for a synthesis array: one can tailor the resolution to the astronomical requirements over a very wide range. The two 'single-dish' modes described above are examples of extrema within this range. In the specific case of the Millimeter Array in its compact configuration (Fig. II-1) there is a third 'single-dish' mode. The instantaneous (one 10-second sample) u-v coverage of the MMA is complete: there are no holes. Thus the effective equivalent illumination pattern is that of an entire 70-m circular aperture. Observations in the compact MMA array configuration, therefore, will provide images entirely equivalent to those obtained by a single antenna with an aperture of 70-m diameter and an effective collecting area which may be thought of as corresponding to an aperture efficiency of approximately 35%. The MMA provides such a 'single-dish' measurement not at one position in the sky, but at $18^2 = 324$ positions (*cf.* Table II-1) simultaneously.

Thus these three MMA modes cover all the common modes of operation of a single antenna but do so with more flexibility. For most problems at any resolution the MMA should outperform any existing or planned single-antenna instrument (see Appendix B).

3. Sensitivity of the Array

The observing time needed to achieve a particular scientific goal is determined by a combination of the inherent sensitivity of the receivers, the effect of the atmosphere (emission and extinction), the number of fields to be observed, and the integration time per field needed to attain a particular signal-to-noise ratio. The desired sensitivity can be described as sensitivity to detection of a point source or sensitivity to particular size-scales of surface brightness. In this section we summarize the sensitivity properties of the MMA, assuming the antennas and array configurations that have already been described and certain capabilities of the receiving systems.

The receiver-noise temperature $T_{\rm rcvr}$ and the bandwidth $\Delta\nu$ for a broad continuum channel or a spectral-line channel are the most important parameters for array sensitivity, once the number and size of antennas are fixed. Incremental improvements to the noise figures of current generation SIS receivers suggest that it is reasonable to expect that the MMA can be instrumented with receivers at the three principal CO lines, 115, 230, and 345 GHz, having single-sideband noise temperatures of 50, 100, and 150 K, respectively, where the receiver noise refers to a measurement made in the laboratory at the input to the dewar. Adding to this the expected loss in the quasi-optical devices ahead of the window, which is frequency-dependent and contributes something like $9(\nu_{\rm (GHz)}/115)^{0.75}$ K, we may parametrize the receiver-noise temperature as

$$T_{\rm rcvr}\left(\nu_{\rm (GHz)}\right) = 0.44\nu_{\rm (GHz)} + 9\left(\frac{\nu_{\rm (GHz)}}{115}\right)^{0.75}\ {\rm K}\,, \tag{1}$$

which applies across the range 68 to 360 GHz. Note that this estimate is still more than ten times the fundamental quantum limit, so there is considerable room for future improvement. We adopt 2 GHz as the bandwidth to be analyzed at the correlator—that is, two orthogonal polarizations of 1-GHz bandwidth each (see Chapter VIII).

There are a number of other parameters that affect antenna and array sensitivity. As an illustration, assume that all antenna pairs for forty antennas are correlated, yielding 780 independent baselines. Let $\epsilon_{\rm a} = 0.7$ be the aperture efficiency at 230 GHz, and let $\epsilon_{\rm q} = 0.82$ be the correlator quantization efficiency, assuming three-level correlation. The sensitivity, or rms noise, for visibilities measured by a single antenna pair is then

$$\sigma_{\rm c} = \frac{4\sqrt{2}kT_{\rm sys}}{\epsilon_{\rm a}\epsilon_{\rm q}\pi D^2\sqrt{\Delta\nu\Delta t}} \times 10^{23}\ {\rm Jy} = \frac{0.039(T_{\rm sys}/200)}{\sqrt{\frac{1}{2}\Delta\nu_{\rm (GHz)}\Delta t_{\rm (min)}}}\ {\rm Jy}\,, \tag{2}$$

where k is the Boltzmann constant, D is the antenna diameter in meters, $\Delta\nu_{\rm (GHz)}$ is the bandwidth, $\Delta t_{\rm (min)}$ is the integration time, and $T_{\rm sys}$ is the equivalent system temperature measured above the earth's atmosphere (see Eq. 5).

This formula leads to simple scaling equations for point-source and surface-brightness sensitivities for the forty-antenna MMA. Point-source sensitivity is given by

$$\Delta S = \frac{0.99(T_{\rm sys}/200)}{\gamma\sqrt{\frac{1}{2}\Delta\nu_{\rm (GHz)}\Delta t_{\rm (min)}}}\ {\rm mJy}\,, \tag{3}$$

while surface-brightness sensitivity for an array of diameter $B_{\rm (km)}$ is

$$\Delta T_{\rm b} = \frac{0.36B^2_{\rm (km)}(T_{\rm sys}/200)}{\gamma\Gamma\sqrt{\frac{1}{2}\Delta\nu_{\rm (GHz)}\Delta t_{\rm (min)}}}\ {\rm K}\,, \tag{4}$$

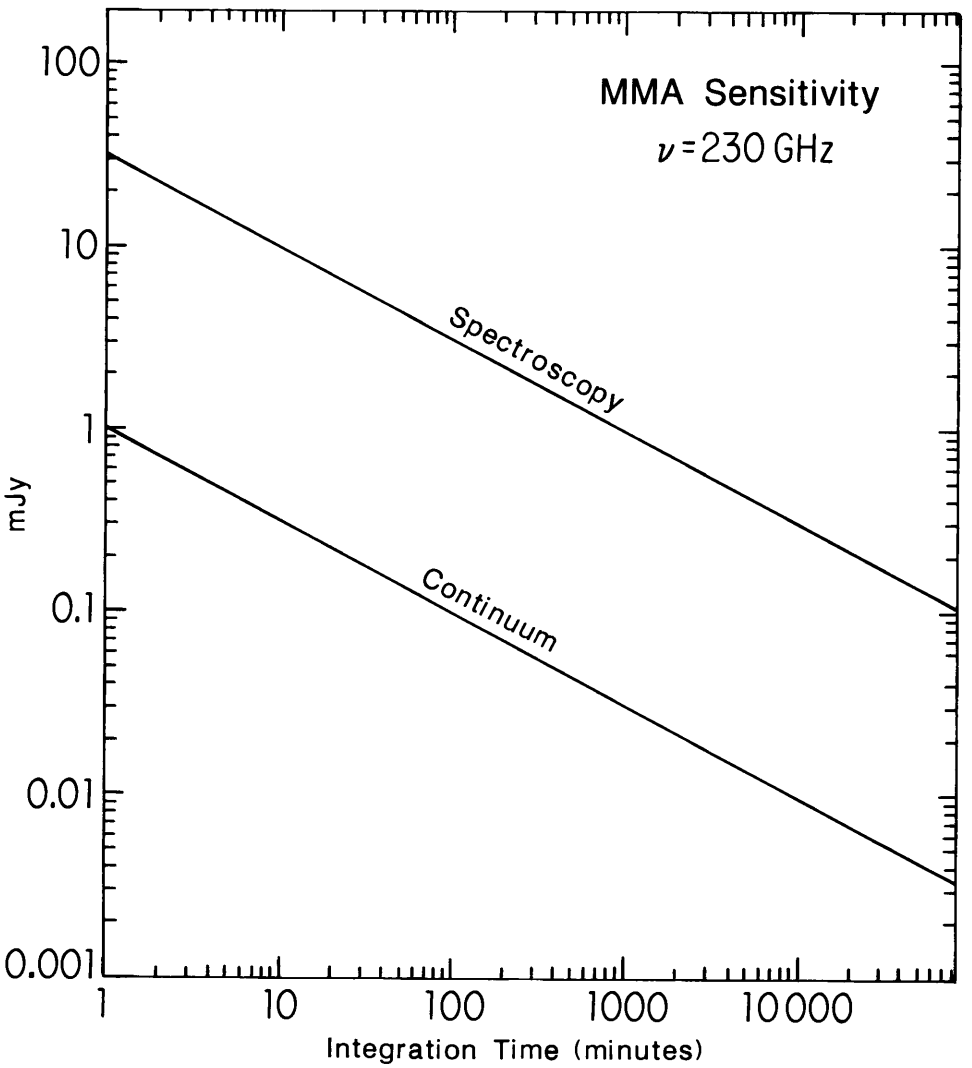


Figure II-2. Point-source sensitivity of the MMA, as a function of integration time, at an observing frequency of 230 GHz (assuming dual-polarization observations and a 2-GHz bandwidth).

where γ and Γ are constants of the order of unity that are dependent upon the spatial distribution of antennas in each array.[2,3]

In Figure II-2 we plot sensitivity (ΔS), as a function of integration time, for point-source detections at 230 GHz assuming $T_{\rm sys}$ = 200 K, for both spectroscopic and continuum MMA observations. The former corresponds to extragalactic spectroscopy with 2 MHz per frequency channel, and the latter cor-

[2] *Cf.* Thompson, A. R., Moran, J. M., and Swenson, G. W., Jr. (1986), *Interferometry and Synthesis in Radio Astronomy*, John Wiley, New York.

[3] Hjellming, R. M. (1985), "Sensitivity Criteria for Aperture Synthesis Arrays", NRAO Millimeter Array Memo No. 29.

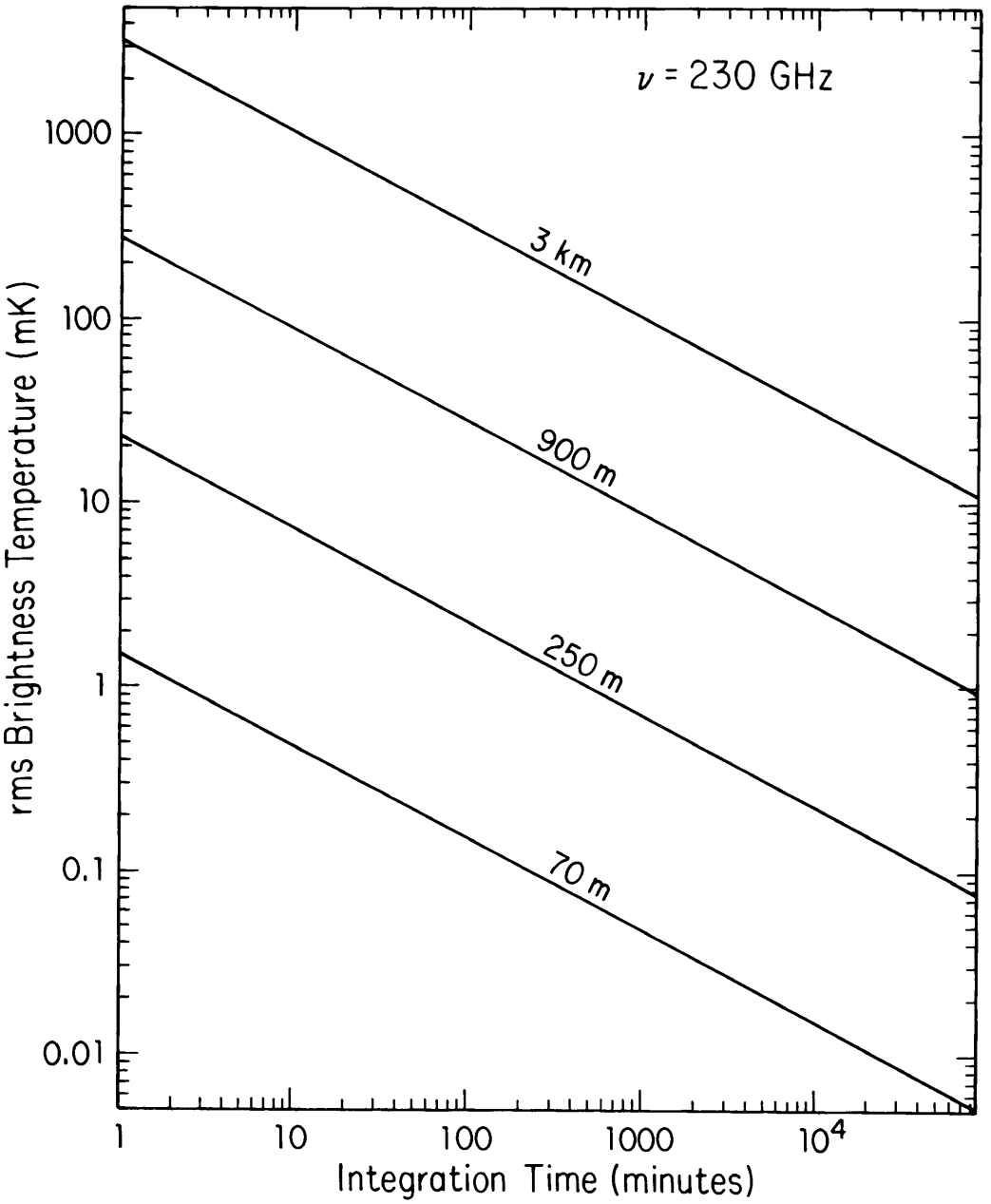


Figure II-3. Surface-brightness sensitivity of the MMA at 230 GHz (assuming dual-polarization observations and a 2-GHz bandwidth).

responds to continuum observations with a bandwidth of 2 GHz. A plot of surface-brightness sensitivity at 230 GHz as a function of integration time for continuum observing with 2-GHz bandwidth is shown in Figure II-3 for the 70-, 250-, 900-, and 3000-meter configurations. For narrow-bandwidth spectroscopy, multiply these curves by the reciprocal square-root of the ratio of spectroscopic bandwidth to 2 GHz.

The true relationship between receiver temperature and system temperature is extremely complicated because of its dependence upon atmospheric transparency, zenith angle, and a number of other instrumental and environmental parameters. Let us adopt a convention whereby radiation temperatures are denoted by primes, that is, $T' = (h\nu/k)/(\exp(h\nu/kT) - 1)$, where h and k are the Planck and Boltzmann constants. We can then express the total system temperature, referred to outside the atmosphere, as

$$T_{\rm sys} = T_{\rm rcvr} e^{\tau_0 A} + \epsilon_l T'_{\rm atm}(e^{\tau_0 A} - 1) + T'_{\rm bg} + (1 - \epsilon_l) T'_{\rm sbr} e^{\tau_0 A}\,, \qquad (5)$$

where τ_0 is the atmospheric transparency at the zenith, A is the air mass, $T'_{\rm atm}$ is the ambient atmospheric temperature (≈ 280 K), $T'_{\rm bg}$ is the cosmic background temperature ($= 2.7$ K), $T'_{\rm sbr}$ is received temperature for rear spillover, blockage, and ohmic losses (≈ 280 K), and ϵ_l is the 'warm-spillover' efficiency of an antenna (≈ 0.85). [This system temperature definition differs from the commonly used $T^*_{\rm sys}$ such that $T_{\rm sys} = \epsilon_l \epsilon_{\rm fss} T^*_{\rm sys}$ where $\epsilon_{\rm fss}$ is the 'cold-spillover' efficiency of an antenna (≈ 0.85).]

In Figure II-4 the system temperature (Eq. 5) is computed by including all the effects discussed previously, including those due to the atmosphere. We have assumed a 2-mm column of precipitable water above the Array (see Chapter VI). The upper panel is the atmospheric transmission shown as a function of frequency, while the lower panel is $T_{\rm sys}$, shown to reflect, and indeed be dominated by, the atmospheric contribution. The three curves that are shown, which correspond to observations taken through $A = 1$, 2, and 3 air masses, respectively, emphasize the importance of siting the MMA at high altitude. They also illustrate the need for fast imaging: since A is (roughly) equal to the secant of the zenith angle, observations taken near the time of a source's meridian transit, through the least amount of atmosphere, have—by a significant amount—the best sensitivity.

4. Components of the MMA

In this section we briefly summarize the characteristics of the major components of the MMA. More detailed discussion can be found in subsequent chapters of this proposal.

4.1. Antennas and Antenna Optics. The antennas will have paraboloidal reflectors eight meters in diameter, with carbon-fiber support structure members used to obtain the thermal stability needed for efficient operation at frequencies as high as 360 GHz. Each reflector will have a surface accuracy of 25 μm or better and must be pointed with an accuracy of $\sim 1'' \lambda_{\rm (mm)}$. Each antenna will be transportable. The preliminary optical design of the antenna uses coudé optics located directly beneath the antenna surface. Receivers will be permanently located at fixed positions in the coudé room and will be selected by rotation of a central mirror.

4.2. Antenna Transport, Stations, and Signal and Power Distribution. The antenna transporter and its coupling to antennas while they are

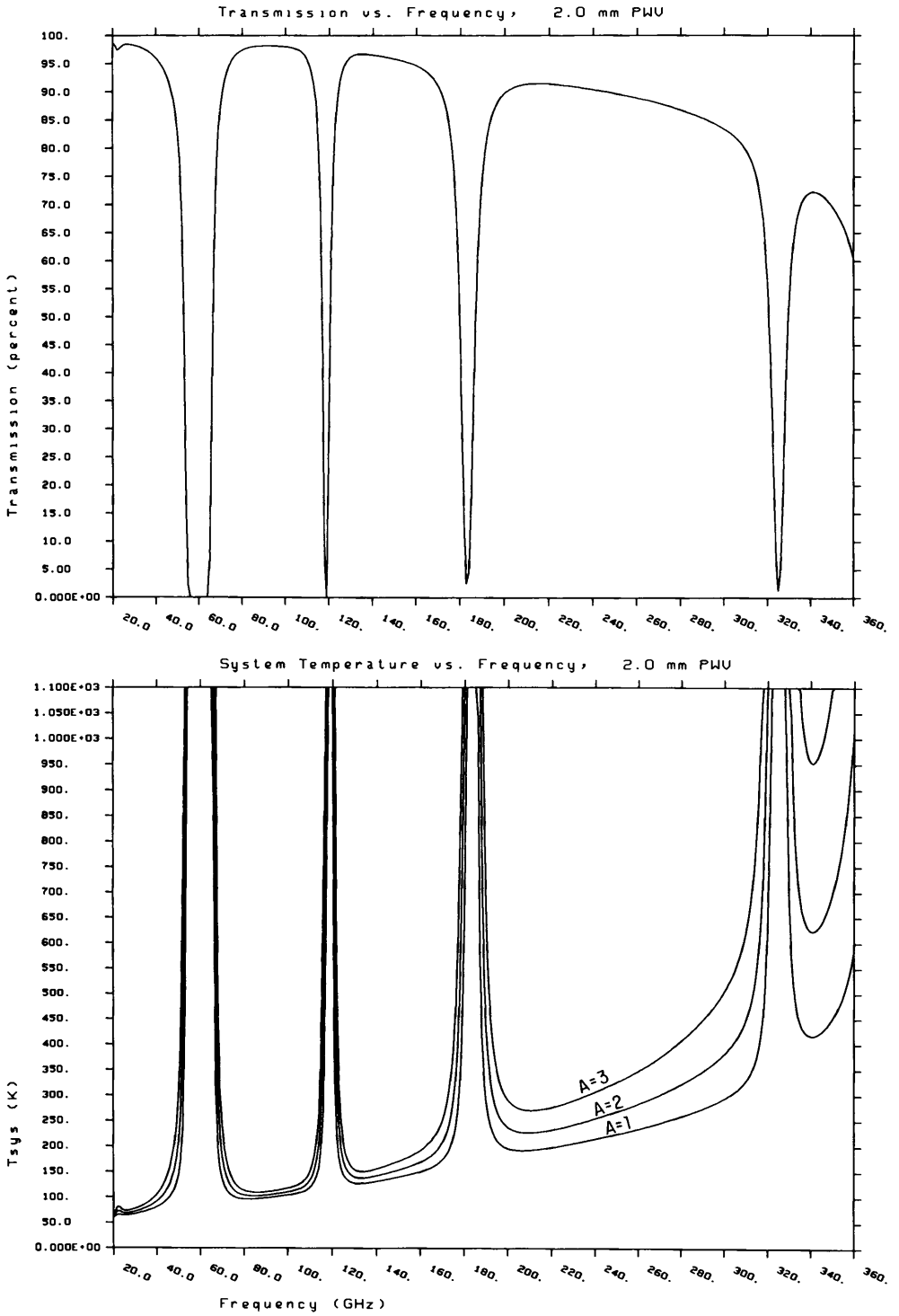


Figure II-4. Transmission of the atmosphere, as a function of frequency, modeled from radiosonde observations and the MPM atmospheric propagation model [H. J. Liebe (1989), *Int. J. Infrared & Millimeter Waves*, **10**, 631] *(upper panel)*. Data correspond to a 10,500-ft site with 2 mm of precipitable water along the zenith-path. The lower panel shows the expected MMA system temperature for the same atmosphere as above. The three curves correspond to observations made through 1, 2, and 3 air masses, respectively.

in transit have not been designed in detail, but modification of commercially available equipment for our application is feasible. To minimize the number of transport motors and the hydraulic complexity, the transporter will carry each antenna as a dead load. The transporter will also provide temporary power for the receivers during transit between stations. As many as 150 stations may be required for the four configurations, and a paved road system will connect the stations. The stations will be equipped with power and optical fiber connections for power and signal distribution. Optical fibers will be used for transmission of signals from a central control building and will also be used for transmission of astronomical and equipment-performance data from each antenna. Two-way signal distribution will operate continuously.

4.3. Electronics and Correlator. Orthogonal-polarization receivers operating in the 350-, 230-, 115-, and 35-GHz atmospheric windows are planned for each MMA antenna. The design goal is to have each receiver tunable over as much of the associated atmospheric window as possible, with continuum operation at bandwidths of 1 GHz per polarization and spectral-line operation with channel widths ranging from 12 kHz to 2 MHz. The correlator will be capable of handling two or more simultaneous continuum channels across the 2-GHz analyzed bandwidth. For spectral-line work it may function as eight or more separate spectrometers, all of which may be tuned to different central frequencies from the same or different receivers.

All electronic components will be designed as replaceable modules, so as to minimize the on-duty operational staff. A complete stock of replacement modules will be maintained on-site. Modules will be serviced at the central support facility.

4.4. Array Control, and Acquisition of Astronomical and Monitor Data. A central computer system, with redundant backup capability, will be used to control the antennas and instrumentation. Data will be sent continuously via the optical fiber communications system. Return timing data, monitor data on the performance of all systems, environmental parameters, and total-power signals for each antenna/polarization/frequency channel will be sent continuously from each antenna, with de-multiplexing and data acquisition before display, analysis, and recording by the on-line computer system. Astronomical data will be recorded on high-density media, probably optical disks, and transferred in real-time to a computer system capable of real-time calibration and imaging. On-line displays will allow the operator to monitor the Array and its equipment and to dispatch technicians to replace defective modules.

4.5. Computational Imaging. As discussed in Chapter V, the computational imaging problems of the MMA are less than for the VLA. This is because, as listed in Table II-1, the typical linear dimension of an image, in number of pixels, is at least a factor of four less for each of the four MMA configurations. Mosaiced observations, likely to be a common type of observation with the MMA, exacerbate the computing situation, pushing the MMA computing needs into near parity with those of the VLA. The most difficult problem will be in providing easy-to-use software for the imaging-computer system—software

which will allow the astronomer to concentrate on scientific computing options, while hiding the complexity of the data being processed. This crucial task is one that should begin immediately, and, indeed, the first steps in this direction have been taken by those using the Owens Valley Radio Observatory (OVRO) interferometer and the Berkeley–Illinois–Maryland Array (BIMA) for imaging—these are sources of vital experience that we will want to draw upon.

4.6. Site. As discussed in Chapter VI, there are three primary MMA sites under consideration. Absolute site requirements are a site at 9000 feet or above, with topography able to accommodate configurations from 70 to 3000 meters in size, allowing transportation between configurations, and with good off-site access. In addition, a low degree of cloud cover over non-summer months, and excellent atmospheric transparency, are necessary. The three sites that meet the absolute criteria and currently are under consideration are: a site in the Magdalena Mountains near Socorro, New Mexico; a site in the Seven Springs area on a plateau south of Springerville, Arizona; and another smaller, but higher plateau site located near Alpine, Arizona. Each of these sites is discussed in Chapter VI. Virtually all operational support for the Magdalena Mountains site could come from the AOC in Socorro. For the Arizona sites, the site operations and maintenance would be supported from a location in Springerville, with the data-processing and institutional support coming from the AOC or from other NRAO facilities.

5. Role of the MMA in Modern Multi-Wavelength Astronomy

The MMA brings to astronomy a combination of high sensitivity and high resolution that has never been achieved in the atmospheric windows from 9 mm to 0.9 mm. In Figure II-5 we plot the sensitivity attainable, as a function of wavelength, for the MMA and other major ground- and space-based instruments. The MMA is necessary to fill in the gap at millimeter wavelengths between the high sensitivity and resolution of the VLA at long wavelengths, and the high sensitivity and resolution of SIRTF and HST at infrared and optical wavelengths. Together these instruments provide astronomers with the research tools necessary in the next decade, and beyond, to unveil the mysteries of the origin of structure in the universe, the birth and death of stars, and perhaps the course of chemistry that leads to life.

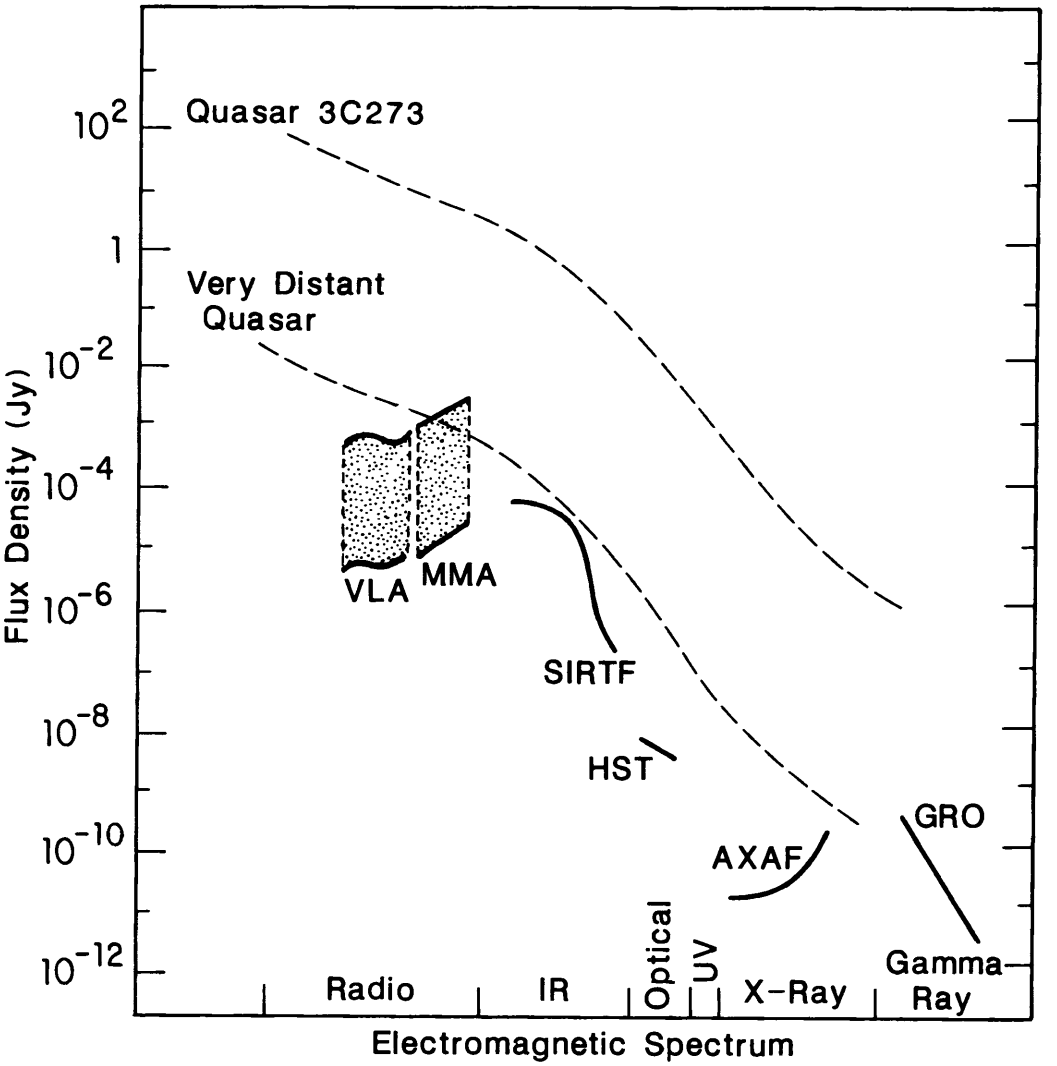


Figure II-5. Sensitivity of the MMA in comparison with that of other major instruments. The regions shown shaded for the VLA and MMA illustrate how the sensitivity improves with integration time. The sensitivities shown here correspond to integration times ranging from one minute to eight hours.

III. THE SCIENTIFIC PROGRAM

1. Observations of the Distant Universe

The Millimeter Array will:

- Image thermal dust emission in evolving galaxies at epochs as early as $z = 10$.
- Yield images of the dust emission in AGNs and QSOs, resolving details as fine as a few tens of parsecs.
- Detect CO emission lines from galaxies and QSOs. A line luminosity of $10^7 L_\odot$ at $z = \frac{1}{2}$ will be detectable in under ten minutes.
- Image the microwave decrement in galaxy clusters; together with AXAF observations, this will provide an accurate determination of H_0.
- Resolve regions of particle acceleration in the jets and lobes of radio galaxies.
- With the world's largest aperture at λ 1 mm, serve as the anchor of millimeter-wavelength VLBI. Accretion disks as small as 10^{15} cm in galaxies as distant as the Virgo Cluster will be resolved.

1.1. The Sunyaev–Zel'dovich Effect. Dense clusters of galaxies are powerful sources of X-radiation, mainly due to bremsstrahlung emission by hot, ionized intergalactic gas. The same hot electrons that provide the bremsstrahlung emissivity also scatter via inverse-Compton collisions the cosmic microwave background photons, and the gas cools. The interaction distorts the spectrum of the background radiation, an effect which is visible in the direction to the cluster. Through this interaction, the photon number density is conserved, but the scattered photons emerge with increased energy and are thus shifted to higher frequency. This results in a decrement in the intensity of the radiation on the Rayleigh–Jeans side of the spectrum, with an increment on the Wien side; the null effect occurs at a frequency of about 218 GHz, slightly higher than that of the peak of the undistorted blackbody spectrum ($\sim$ 162 GHz).

The magnitude of the distortion is directly proportional to the integral of the electron pressure along the line of sight through the intracluster plasma and therefore provides a linear probe of the cluster gas. In the Rayleigh–Jeans limit,

$$\frac{\Delta T}{T} = \frac{\Delta F}{F} = \frac{2kT_e}{m_e c^2} \int n_e \sigma_T \, d\ell \,, \tag{1}$$

where T is the radiation temperature, F is the spectral energy density of the radiation, k is Boltzmann's constant, T_e is the electron gas temperature and n_e is its density, m_e is the electron mass, c is the speed of light, and σ_T is the cross-section for Thompson scattering. The integral, which is taken along the line of sight through the cluster, gives τ_e, the optical depth for the Sunyaev–Zel'dovich effect.

The Sunyaev–Zel'dovich effect has been detected in a few clusters (0016+16, Abell 665, 1558+41, Abell 2218) at $\lambda = 1.5$ cm at a level of about -0.5 mK. These results imply central decrements up to -1.0 mK, depending on the clumping of the gas and the run of temperature in the gas distribution. Thus we can expect that at $\lambda = 9$ mm and 3 mm, distant dense clusters of galaxies will appear as diffuse negative sources with peak decrements of up to -1.0 mK and angular sizes of a few arcminutes. The measurement of such weak and probably smoothly varying signals will be made with the MMA in its most compact configuration and will require mosaiced observations. Detection of the Sunyaev–Zel'dovich effect in clusters at the highest frequencies, ~ 350 GHz, will provide an important confirmation of the predicted frequency-signature of the effect.

Besides providing an important probe for the study of the hot gas in clusters of galaxies, the Sunyaev–Zel'dovich effect is an important cosmological tool. The combination of X-ray (AXAF) and MMA images of the same clusters allows a direct determination of cluster distances, avoiding the stepladder approach to cosmological distances.

Schematically, the problem is solved as follows: The X-ray surface-brightness of a cluster is given by

$$S_x = C_x \int f(T_e) n_e^2 \, d\ell \,, \tag{2}$$

where C_x is a constant of proportionality that depends on basic parameters. Spatially resolved X-ray spectra will provide the run of temperature in the cluster as well as an image of the X-ray emission. Since Equations 1 and 2 depend on different powers of n_e and T_e, one can solve for n_e using T_e from the X-ray observations or one can combine the two equations, eliminating n_e, which allows one to write an equation for the distance as a function of the X-ray luminosity and the Sunyaev–Zel'dovich decrement. Since both AXAF and the MMA will image the cluster emission with about the same angular resolution, and since this analysis should apply to each independent path through the cluster, the distance solution should be heavily constrained by the spatial information and by the integral properties of the observed emission. Additionally, this will constrain the clumpiness of the gas as well as the run of temperature in the cluster.

To first order, the fractional uncertainty in the determined distance is proportional to the combination in quadrature of the fractional uncertainties in the X-ray flux, the X-ray core size, twice the X-ray temperature, and twice the fractional microwave decrement. Currently available data allow estimates of distances to a precision of about 30%. Distance estimates obtained by combining MMA and AXAF observations are expected to be accurate to better than 10%; of that total error the MMA measurement should contribute less than 3%.

1.2. Primordial Galaxies and the Early Evolution of Galaxies. Cosmological models for the large-scale structure of the universe lead to specific predictions for the structure and evolution of primeval galaxies. The ultimate assessment of cosmology can be made by observations of galaxies in the formative stages, but, as yet, there are few candidate protogalaxies known. The paucity

of identifiable primeval galaxies, despite thorough and systematic searches, has led to the conclusion that most galaxies form stars at high redshift, or they are much less luminous than expected, or they emit their radiation at frequencies outside the optical/IR bands that have been searched.[1] The latter possibility appears most likely, and it provides one of the great opportunities for the MMA.

Contemporary cosmological models differ in their assumptions about the dominant forms of matter and energy; this, in turn, is reflected by differences in their predictions concerning the hierarchical ordering of structures present in the observable universe. An inflationary universe dominated by cold dark matter will give rise to a scale-invariant spectrum of initial fluctuations which will, with time, grow in magnitude. As the small scales are the first to go nonlinear ($M \approx 10^6 M_\odot$), galaxies and all larger structures form later from the interaction and gravitational accretion of small objects. This leads to an important observational test noted below. Alternatively, if the universe is dominated by the energy of neutrinos or other weakly interacting particles that stream freely over comoving scales of megaparsecs, then primordial galactic-scale fluctuations will be suppressed and only the fluctuations with cluster masses ($M \approx 10^{15} M_\odot$) will survive. Individual galaxies arise from fragmentation at quite late epochs. Finally, a speculative class of cosmological models invokes some sort of primeval seeds, such as cosmic strings, which act as local sources of gravity around which a hierarchy of structures may be built up.[2]

These models make different and testable predictions for the early evolution of galaxies. If we concentrate on the evolutionary phase during which most of a galaxy's stars form, then we can anticipate that such a protogalaxy can be described as a collection of gas clouds gravitationally interacting. Collisions of the clouds and the subsequent violent compression or merger will lead to a vigorous burst of star formation that is the defining characteristic of a protogalaxy. While the first generation of stars will be deficient in heavy elements, the short-lived luminous stars will enrich both the composition of the gas clouds, and subsequent generations of stars, and they will provide a reservoir of dust to aid the star-formation process. The stellar cycle can proceed through several generations in the time scale for gravitational collapse of the protogalaxy. Thus, a galaxy in its earliest evolutionary phase will appear as a highly inhomogeneous structure dominated by vigorous bursts of star formation at the sites of cloud collisions distributed over a volume far larger than that occupied by the galaxy in its eventual equilibrium configuration. An increasingly larger fraction of the ultraviolet and optical radiation emitted by successive episodes of star formation will be absorbed by the dust in the protogalactic gas and re-emitted in the far-infrared.[3] This radiation, redshifted to millimeter wavelengths, will be imaged by the MMA.

---

[1] Koo, D. (1986), in *Spectral Evolution of Galaxies*, C. Chiosi and A. Renzini, Eds., (Reidel: Dordrecht), p. 419.

[2] Silk, J. (1987), in *High Redshift and Primeval Galaxies*, J. Bergeron, D. Kunth, B. Rocca-Volmerange, and J. Tran Thanh Van, Eds., (Editions Frontières: Gif sur Yvette), p. 301.

[3] Baron, E. and White, S. D. M. (1987), *Astrophys. J.*, **322**, 585.

Starburst galaxies are the closest analogue we have to the processes likely to dominate the early evolution of galaxies. Here the defining characteristic of a 'starburst' is a transient event in a galaxy in which the star-formation rate is much higher than normal. The necessary condition leading to a starburst is the rapid increase in gas surface density,[4] to perhaps $\sim 1000 M_\odot$ pc^{-2}. The accumulation of the gas must occur rapidly, on the time scale of the starburst itself, $\sim 10^7$ years. Therefore, fueling a starburst is also a violent event which we observe locally to be driven by the tidal interaction of colliding galaxies or by the gravitational disturbance of a nuclear bar. In the early evolution of galaxies, violent dynamical disturbances—and the consequent starbursts—were certainly more frequent, owing both to the higher average gas surface density in protogalaxies and to the increased tidal torques resulting from the collapse of irregular structures. This, together with the bimodality of the initial-mass function which we expect, but need the MMA to verify, in starbursts and protogalaxies leads us to describe protogalaxies as luminous, inhomogeneous, and dusty objects.[5]

Observationally, the thermal dust emission spectrum from starburst galaxies has the form

$$I_\nu = \text{constant} \times \nu^n B_\nu(T)\,, \tag{3}$$

which is the Planck function $B_\nu(T)$, modified by the dust emissivity of index n. Locally, n varies between 1 and 2.[6,7,8] At frequencies well below the peak of the Planck function, $B_\nu(T) \propto \nu^2$, so that $I_\nu \propto \nu^3$ to ν^4.

As we observe protogalaxies at earlier epochs, a given object observed at a particular frequency in the rest frame will be fainter, owing to the greater distance associated with early epochs. But the very steep and favorable K-correction along the Rayleigh–Jeans portion of the spectrum more than compensates: the flux density observed at a fixed frequency of observation will vary little with redshift until the K-correction carries the frequency observed over the peak of the Planck function.

As an example, consider a galaxy at redshift z with a molecular hydrogen mass-to-IR luminosity consistent with the correlation noted by Sanders *et al.* for starburst galaxies.[9] The flux density observed at a frequency $\nu_{\rm obs}$ (GHz) along the Rayleigh–Jeans portion of the spectrum will be

$$S_{\rm obs} \approx \frac{1.7 \times 10^{-10} \nu_{\rm obs}^3 (1+z)^3 \left(L_{\rm IR}/10^{12} L_\odot\right)}{\left(q_0 z + (q_0 - 1)\left(\sqrt{1+2q_0 z} - 1\right)\right)^2} \ \text{mJy}\,, \tag{4}$$

[4]Larson, R. B. (1987), in *Starbursts and Galaxy Evolution*, T. X. Thuan, T. Montmerle, and J. Tran Thanh Van (Editions Frontières: Gif sur Yvette), p. 467.
[5]Silk, J. (1987), *op. cit.*
[6]Draine, B. T. and Lee, H. M. (1984), *Astrophys. J.*, **285**, 89.
[7]Emerson, J. P. (1988), in *Formation and Evolution of Low Mass Stars*, A. K. Dupree and M. Lago, Eds., (Kluwer: Dordrecht), p. 21.
[8]Helou, G. (1989), in *Interstellar Dust*, IAU Symp. 135, (Reidel: Dordrecht), in press.
[9]Sanders, D. B., Soifer, B. T., Elias, J. H., Madore, B. F., Matthews, K., Neugebauer, G., and Scoville, N. Z. (1988), *Astrophys. J.*, **325**, 74.

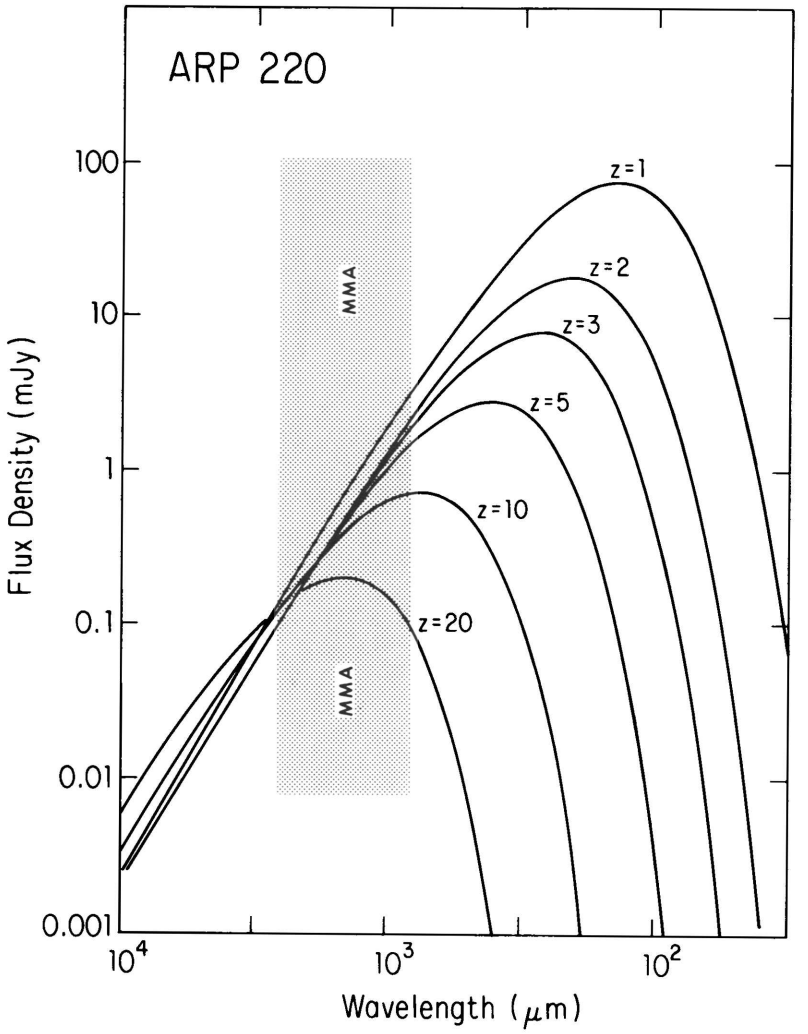


Figure III-1. The expected spectrum of the starburst galaxy Arp 220 as the object is moved to higher redshift. The shaded area indicates the frequency range of the MMA.

where we adopt $q_0 = \frac{1}{2}$, $H_0 = 75$ km s^{-1} Mpc^{-1}, $n = 1$ as the dust emissivity index (see Eq. 3), and $T = 60$ K as the dust temperature. Here $L_{\rm IR}$ is the bolometric luminosity at all wavelengths $\lambda > 2$ μm.

Figure III-1 illustrates the variation in the observed flux density of Arp 220—a local example of a starburst galaxy—as we imagine the object being moved out to higher redshift. Note that while this object is extremely bright in the far infrared at $z = 1$, by the time it recedes to $z = 5$ the 100-μm flux density has diminished by a factor $\sim 10^5$. Not so at millimeter wavelengths: Arp 220 at $z = 1$ is detectable by the MMA at 1-mm wavelength in a few minutes' integration time. Precisely the same object at $z = 10$ observed at λ 1 mm still has a flux density on the order of 1 mJy and may be imaged quickly by the MMA. To the extent that protogalaxies resemble starburst galaxies, the MMA will image local protogalaxies as well as those at epochs of formation as early as $z = 10$ with equal ease.

How many protogalaxies could there be? We can place a crude limit on this number by assuming that every galaxy during the formation process exhibits starburst properties n times and that each starburst episode lasts 10^7 years. We then expect approximately $0.1n$ starburst galaxies per square arcminute. As n could easily be of order 10, meaning that any given galaxy spends 1% of a Hubble time in starburst phase, and again using Arp 220 as a representative protogalaxy, meaning that the expected flux density per protogalaxy is $\sim$ 1 mJy at λ 1 mm (see Fig. III-1), one or more protogalaxies may appear as a confusing source in every MMA beam at a level $\approx$ 1 mJy. Indeed, this estimate is consistent with the inferences drawn by Tyson and his collaborators from their deep optical imaging of L^* galaxies to $z \approx 7$: they find $\sim$ 150 such galaxies per square arcminute. (Of course, we do not know how many of these are as luminous, or have as much dust, as Arp 220.) In any case, if our models of cosmology and galactic evolution are correct, objects difficult to detect now at any wavelength—protogalaxies—will be to the MMA a nuisance!

1.3. Spectroscopy of Evolving Galaxies. Figure III-2 is a montage showing both the optical image of the $z = 0.16$ QSO Mk 1014 and the CO ($J = 1$–0) emission-line spectrum of this object. There are several points to be made with this illustration: (1) the angular extent of this gas-rich QSO is small, 20″, yet there is significant structure on 1″ scales; (2) the optical image is distorted, showing signs of a tidal interaction; and (3) the flux density of the CO emission line is quite substantial. Mk 1014 has all the characteristics of an evolving starburst galaxy discussed above: the gas surface-brightness exceeds $1000 M_\odot$ pc^{-2} in a small region; the starburst-fueling process is evident; and the stellar optical and ultraviolet light is re-radiated in the far-infrared. To understand such rapidly evolving galaxies we need a spectroscopic tool that will allow us to determine masses and internal kinematics throughout optically opaque sources. MMA observations of CO and its isotopes are such a tool that can be employed to image Mk 1014 (in much less than an hour's observing), and this technique can be extended to numerous similar galaxies and QSOs at significantly higher redshifts.

If we let $L_{\rm CO}$ (ergs s^{-1}) be the total luminosity of the galaxy in the CO rotational transition that is emitted at frequency ν_0, then the flux density that we will observe in this line is related to the luminosity diminished by two factors of $(1 + z)$. One power of $(1 + z)$ reflects the diminution of photon energy with redshift, and the second accounts for the reduced rate at which photons arrive. The observed flux density is

$$S_{\rm CO} = \frac{10.6\left(L_{\rm CO}/10^6 L_\odot\right) H_0^2 q_0^4}{(1+z)^2 \nu_0 \Delta v} \left(q_0 z + (q_0 - 1)\left(\sqrt{1 + 2q_0 z} - 1\right)\right)^{-2} \ {\rm mJy}\,, \quad (5)$$

where Δv (km s^{-1}) is the full width of the line, at frequency ν_0 (GHz). Objects such as Mk 1014 and IRAS 07598+6508 have CO luminosities $\sim 10^6 L_\odot$ to $10^7 L_\odot$, meaning that they can be imaged by the MMA out to redshifts $\sim$ 1.5 in less than a day's integration time. Such an MMA image is more than a picture of the gas distribution: The MMA kinematic resolution, $\Delta v < 1$ km s^{-1} (finer

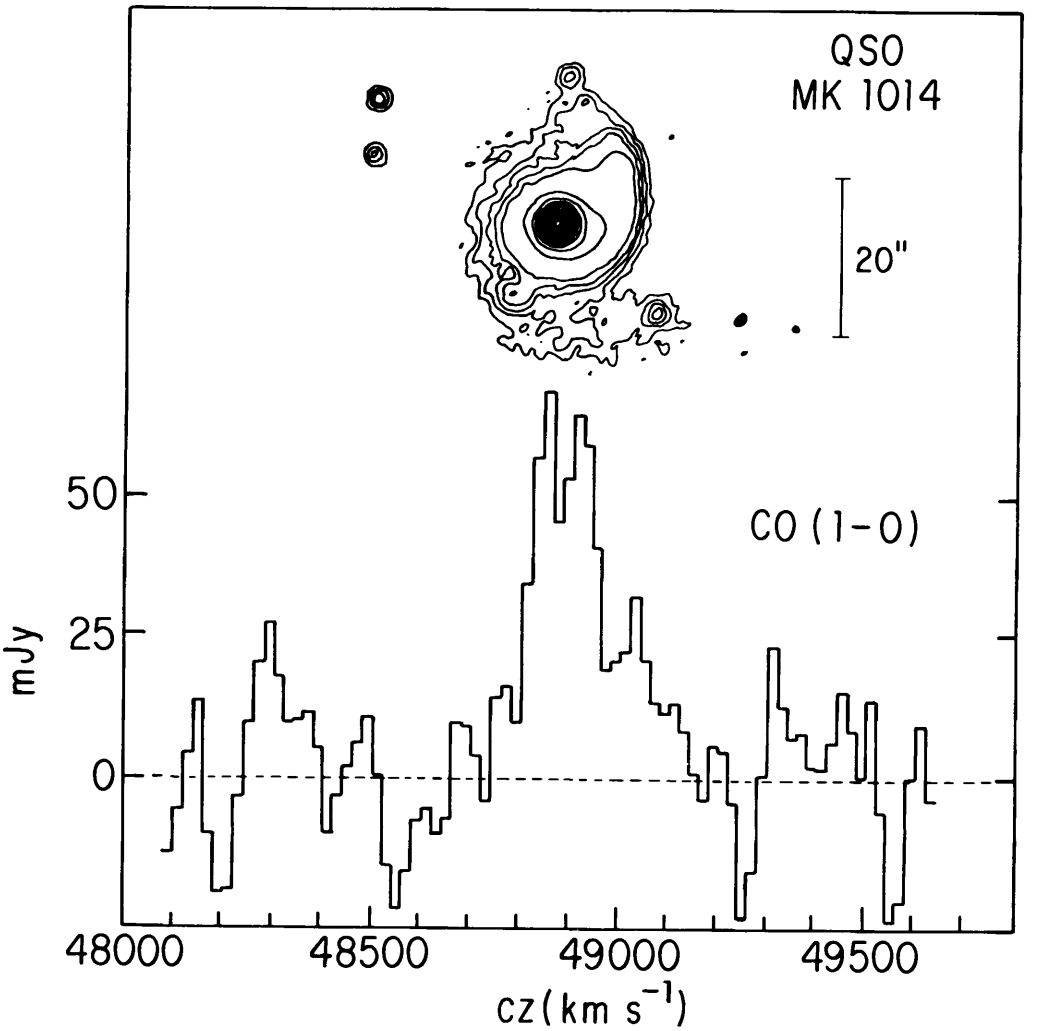


Figure III-2. A combination of an Hα photograph of the QSO Mk 1014 *(upper)* and the CO (J = 1–0) spectrum obtained on the NRAO 12-Meter Telescope *(lower)*.

than one would need for extragalactic applications), is sufficient to trace the gas trajectories within the system and provide clues as to the origin and fate of the interstellar material.

The best observational approach for detecting high-redshift CO will be determined through experience. For a wide range of redshifts one could image more than one CO rotational transition. Which will be easiest to detect? At $z > 1.3$, CO (J = 1–0) in the 9-mm window may be particularly favorable. In addition, the Green Bank Telescope will be helpful to identify, but of course not to image, higher-redshift CO emission in evolving galaxies.

At a redshift of 0.4, the $^3P_0$–$^3P_1$ fine-structure transition of neutral atomic carbon will be redshifted into the 350-GHz atmospheric window accessible to the MMA. Although we have no experience on which to base an estimate of the luminosity of this line in protogalaxies, nevertheless, since the [C I] line brightness

in Galactic photodissociation regions is comparable to that of CO, and since the ionization in a starburst is very similar to that in a photodissociation region, it is reasonable to speculate that $L_{\rm C\,I} \approx L_{\rm CO}$. If so, the remarks above regarding the detectability of CO at high redshift apply equally to [C I]; Equation 5 applies without modification.

1.4. Active Galaxies: Extended Non-Thermal Emission. The MMA sensitivity, angular resolution, and frequency range will allow the study of particle acceleration in radio jets and lobes to reach a new maturity. Filaments in the radio lobes of the prototypical radio galaxies Cygnus A and M87 have flatter spectra than the more diffuse components in which they are embedded. No evidence for a high-frequency turnover in the spectra is evident at centimeter wavelengths. If this trend continues to higher frequencies, millimeter-wave images of the lobes of such sources may consist *only* of filaments. Are processes in the filaments responsible for producing most of the high-energy particles in radio lobes?

The combination of the centimeter-wave and millimeter-wave images will allow us to learn more about the time evolution of radio sources by separating the young/high-field regions from the older/low-field parts of the source. It will give us a new and perhaps clearer window through which to view the problem.

The MMA will also allow a second, wholly novel constraint to be placed on the physics of radio galaxies. The Sunyaev–Zel'dovich effect should be observable from high-pressure thermal gas in and surrounding the extended lobes of radio galaxies. This effect yields a line integral of the thermal pressure through the scattering region. Thus, if, as current numerical models assume, radio lobes are overpressurized regions in the intergalactic medium due to high Mach-number flows, and if they are dominated by thermal gas, then the MMA will detect and image the Sunyaev–Zel'dovich decrement from these regions. This measurement will yield the thermal pressure confining the lobes, and combining this with X-ray or optical data will give us the Mach number and velocity of the flow. If the effect is not seen, one can safely conclude that the confining pressure is provided by magnetic fields or very relativistic particles.

1.5. Active Galaxies: The Nuclear Regions. The MMA will, in two important ways, facilitate studies of active galactic nuclei (AGNs): First, the dust emission and CO emission used for the study of high-redshift objects will also allow us to image the processes responsible for AGN activity at millimeter wavelengths. At $z \approx 0.1$, the $0\farcs1$ angular resolution corresponds to about 200 pc. Kiloparsec-sized nuclear disks will be resolved. The interaction kinematics between galaxies, perhaps responsible for the AGNs, will be studied using CO and other tracers. Second, the MMA will serve as the anchor for millimeter-wavelength VLBI observations. The addition of the MMA to existing telescopes will increase the number of sources observable at λ 1 mm from just a few to well over one hundred.

As a millimeter-wave VLBI station, the MMA will facilitate the highest-resolution investigation achievable from an Earth-based telescope of the cores of powerful radio galaxies, and it might allow us to see the outer parts of the

accretion disks around the black hole or whatever the power source in the nucleus may be. An Earth-diameter baseline at λ 1 mm corresponds to 20-μarcsecond resolution or 5×10^{15} cm at the distance of M87, about the scale of the accretion disks many models predict. Also, unlike VLBI at centimeter wavelengths—even those VLBI observations of equal resolution involving an orbiting element—the millimeter-wave images are made at frequencies at which the sources are transparent, so that one can, indeed, see into the nuclear core.

2. Observations of the Universe Nearby

The Millimeter Array will:

- Reveal the masses and kinematics of optically obscured nuclei of galaxies, with a resolution superior to that of the Hubble Space Telescope.
- Image the distribution of deuterium and the distributions of C, N, O, S, and Si, and their isotopes, in galactic disks.
- Observe giant molecular clouds (GMCs) in galaxies as distant as 100 Mpc, including all galaxies in the Shapley–Ames catalogue.
- Study the dependence of spiral structure and galaxy evolution on properties of the interstellar medium (ISM).
- Provide unobscured kinematic images of starburst galaxies.
- Detect spectral lines of H_2O and O_2 in distant galaxies, uncontaminated by telluric emission.

2.1. Galactic Evolution and the Interstellar Medium. Understanding the circulation of matter within a galaxy in the gas phase, and also the circulation of matter into and out of stars, is fundamental to an understanding of how galaxies evolve. Since star formation takes place exclusively in molecular clouds, observations of the dust concentration and molecular distribution are crucial. While Galactic studies have provided information about the physics of star formation in molecular clouds in some detail, extragalactic studies are necessary to obtain information on molecular clouds and star formation under a wide variety of astrophysical conditions.

For example, galaxies of different Hubble types are thought to have different histories of star formation, based on studies of galaxy colors. The ability to probe the molecular-gas distributions in galaxies over a range of epochs may determine directly the physical origins of such different evolutionary histories. These studies require high-resolution imaging at high sensitivity.

2.2. Elliptical Galaxies. Recently, CO has been detected in a number of elliptical galaxies. Yet unanswered is the old question of where this material originates and where the gas which is expelled by the supernovae and planetary nebulae goes. What supports the continuing star formation in otherwise dead stellar systems? Is the accumulation of gas that we see in elliptical galaxies a repository of gas ejected by aging stars, or is it the reservoir for new star formation? The new observations of the molecular gas and the somewhat older detections of atomic gas can begin to address these questions; MMA images will facilitate detailed studies in a vastly larger sample of elliptical galaxies. Furthermore, the supposition that the central galaxies in clusters grow by devouring

smaller galaxies can also be probed by observing the distribution and kinematics of residual gas in systems suspected of having undergone recent mergers.

2.3. Galactic Structure and Star Formation. In a series of papers, van der Kruit and Searle showed how it is possible to probe the dark-matter halo of an edge-on galaxy by mapping the scale-height distribution of H I. Using this information, together with the velocity dispersion measurements of the gas distribution in face-on systems and optical photometry of the disk, one may determine the run of surface density in the disk. The rotation curve measures the total mass of the galaxy as a function of distance, and the difference between these two determinations is then the mass that must be attributed to the dark halo. The method as presently applied to H I is limited by the resolution and surface density obtainable for galaxies observed with the VLA. The MMA will increase the resolution attainable by more than an order of magnitude and the surface-brightness sensitivity per H-nucleon by a factor of about five. This allows the acquisition of a much larger data set and permits the *weighing* of a disk all the way into the galactic nucleus. Knowledge of the density distributions in dark-matter halos bears importantly on our understanding not only of the structures of galaxies, but of galactic evolution as well.

SO galaxies also have now been detected in CO, some with surprisingly large emission-line flux densities. Since SO galaxies do not have spiral arms, how is the molecular gas distributed in these galaxies? Why is it that the high CO contents do not give rise to even a relatively low level of star formation? A related question is, to what degree does spiral structure affect the global star formation in a galaxy? How does the distribution of molecular gas in a flocculent galaxy differ from that in a grand-design spiral, if at all? The answers to these questions require the sensitivity and angular resolution of the MMA.

Figure III-3 shows a 6″ map of CO ($J = 1$–0) emission from M51, made with the OVRO interferometer, superimposed on an optical photograph. A clear correlation of CO and dust in the spiral arms is evident. The CO observations provide quantitative measures of temperature, velocity field, velocity dispersion, and the gas column density of the dense molecular phase of the ISM. The MMA has sufficient sensitivity and resolution to image the molecular constituents of galactic disks in galaxies nearby, as well as those at cosmological distances. The sensitivity of the MMA will permit an order-of-magnitude greater spatial resolution than is currently possible for a galaxy such as M51. Indeed, the MMA image of M51 will show considerably more detail than the optical photograph.

In the Milky Way a large fraction of the high-surface-brightness CO is confined to GMCs. If GMCs are commonly the repository of molecular material, the MMA will be able to image most of the molecular emission from each of the galaxies in the Shapley–Ames catalogue. It will be possible to determine the relationship between molecular clouds and spiral structure, how molecular gas is distributed in detail, and what relationship the molecular gas has to star formation and the general interstellar medium in galaxies of all morphological types in a large variety of environments. Because the integrated emission from galaxies can be detected to great distances, it will be possible to investigate how the molecular content affects the evolution of galaxies. The determination of

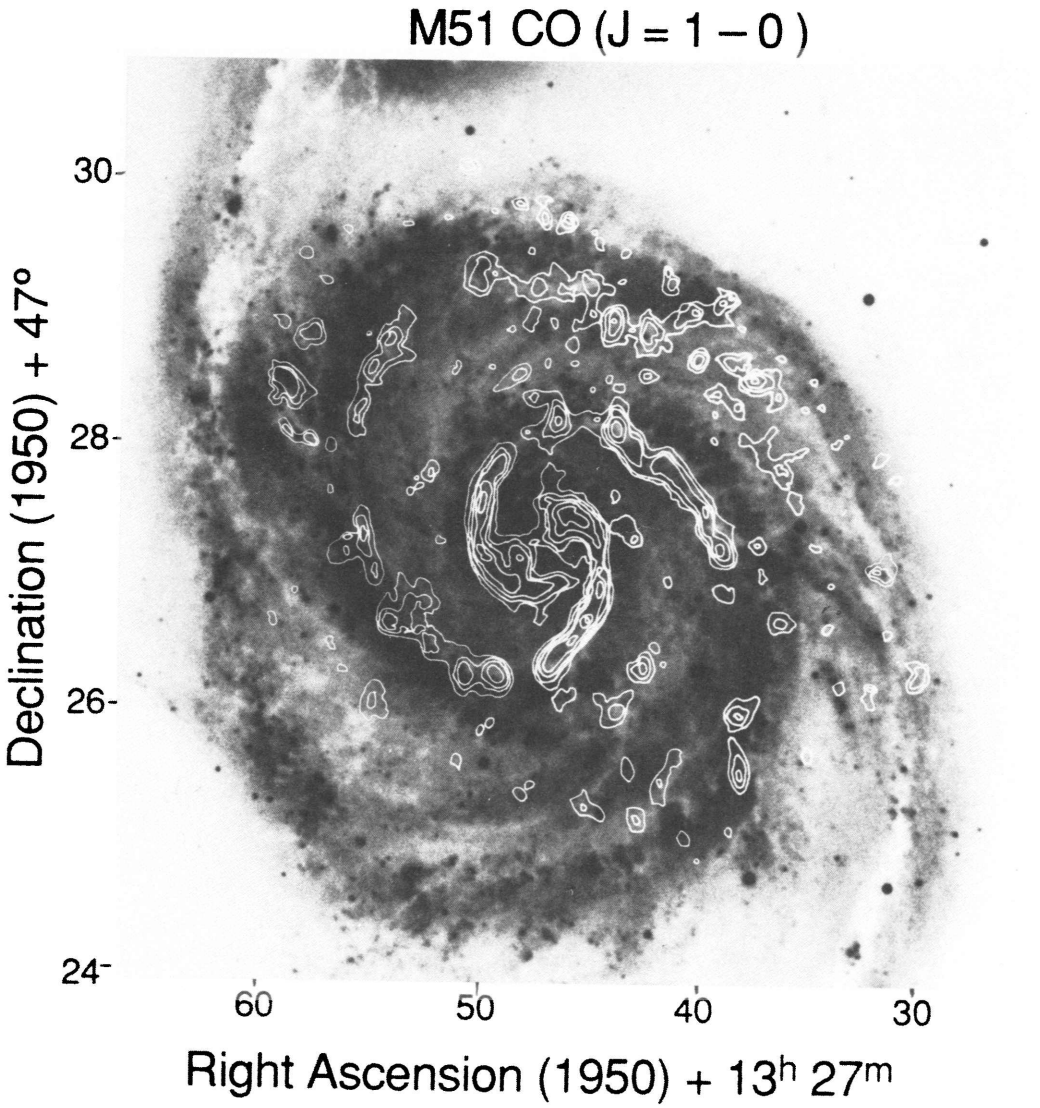


Figure III-3. CO ($J = 1$–0) emission contours superimposed upon an optical photograph of M51. [From R. Rand and S. Kulkarni (1990), preprint.]

H_0, using tidally limited GMC diameters as objects of known physical size, may be possible with the MMA.

2.4. Barred Spiral Galaxies. High-resolution observations of CO emission from the nuclear regions of a number of spiral galaxies have revealed the presence of kiloparsec-scale molecular bars. Such unexpected distributions have led to the recognition that in nuclear regions, oval or bar-like stellar distributions may be very common and that the gas distribution is strongly affected by the underlying non-axisymmetric mass distribution. Figure III-4 shows a 2″ map of the CO distribution in the central one arcminute of a nearby spiral galaxy IC 342, made with the Nobeyama (NRO) millimeter-wave array. Both the kinematic and spatial distribution of the gas are very well accounted for as the gas response to

Figure III-4. CO (J = 1–0) emission from the nuclear region of the spiral galaxy IC 342. The 2″ resolution of this image corresponds to a spatial resolution of 45 pc. [From S. Ishizuki, R. Kawabe, S. Okumura, K.-I. Morita, and M. Ishiguro (1990), private communication.]

an oval potential, although such a feature is not readily apparent on photographs taken at optical wavelengths. This illustrates the importance of high-resolution observations of obscured regions—a wide variety of unusual phenomena occur in galactic nuclei. Observations of the central regions of galaxies with the MMA can provide a very powerful probe of the gravitational potentials in galactic nuclei.

2.5. Active Galactic Nuclei. A fundamental and fascinating problem in contemporary astronomy is to understand the nature and evolution of active galactic nuclei (AGNs), ultraluminous galaxies, radio galaxies, starburst nuclei, and their relationship to the quasars. The gas from which the stars in starburst nuclei form can be studied at sub-arcsecond resolution with the MMA in galaxies several hundred megaparsecs distant. Further, since the millimeter-wavelength observations are completely unaffected by interstellar absorption, the MMA will make it possible for us to probe fully the behavior of the neutral gas in optically obscured regions.

Energetic galactic nuclei encompass a variety of phenomena, ranging from starburst activity such as that found in M82 and NGC 1068, to Seyfert nuclei like NGC 3227 and NGC 4151, to possible quasars such as Mk 231. Different manifestations of these phenomena can frequently be found in a single galaxy, suggesting an evolutionary link between them. Moreover, there is increasing evidence that these processes are the result of interactions and mergers between galaxies. An example of an interacting pair, the 'Antennae' is shown in Fig-

ure III-5, an image from OVRO. Carbon monoxide uniquely probes the gas kinematics in the system; strong concentrations are seen at the individual nuclei and in the region of overlap between the two galaxies. Furthermore, the distribution of the CO compared to the smoother H I distribution permits one to investigate the gravitational potential in far greater detail than is possible using observations of the atomic gas or the stars. With the high spatial and velocity resolution of the MMA, it will be possible to untangle the history of such interactions and to elucidate the dynamics of the merger process. Studies of many pairs will make it possible to determine the physics leading from a gravitational interaction and merger to a starburst.

An example of the type of data needed for these investigations is illustrated in Figure III-6. These are images of M82 made in the radiation from HCO^+, HCN, and ^{13}CO with the Berkeley–Illinois–Maryland Array (BIMA) at Hat Creek. The molecular-gas distributions differ considerably from the optical appearance of the galaxy, illustrating the ability of millimeter-wavelength interferometric observations to penetrate optically opaque regions. Furthermore, note that rare and exotic molecular species can be mapped easily in starburst galaxies. Indeed, the majority of molecules detected in the interstellar medium of the Milky Way can probably be detected in starburst nuclei with the MMA. This, in turn, will make it possible to obtain a detailed knowledge of the densities, temperatures, pressures, and chemistry in the star-forming gas, at a resolution of ~ 2 pc for the nearest starburst galaxies. Indeed, the MMA will obtain images of extragalactic molecular-gas emission at the resolution of the Hubble Space Telescope and probe the gas properties deep in the active nuclei of Seyfert galaxies. There are increasing indications that a reservoir of molecular gas feeds the 'monster' in each AGN. MMA observations should be able to decide this issue once and for all. Since it is possible to obtain a resolution of ~ 5 pc for the nearest AGNs, we can determine whether there are accretion tori circumscribing their nuclei and whether there exists dynamical evidence for massive black holes.

Figure III-7 shows the CO emission from the nucleus of the Seyfert 1 galaxy NGC 3227. The emission is elongated in the direction of the interacting neighbor NGC 3226. The figure also shows the central knots of CO, which are embedded in the narrow-line-emitting gas. These knots may be the limb-brightened edges of a circumnuclear torus; we need higher angular resolution in order to know for sure.

Finally, the MMA will provide the opportunity to observe dust complexes in galactic nuclei out to cosmologically interesting distances. Consider the infrared-luminous galaxies discovered by IRAS. Millimeter-wave interferometry has shown, for several of these objects, that huge amounts of molecular gas lie in the innermost kiloparsecs of the galaxies. The dust which gives rise to the IRAS emission accompanies this gas, and it too will be measured by the MMA. For thermal dust emission from a galaxy, one can write, roughly

$$M_D = 7 \times 10^4 D^2_{(\mathrm{Mpc})} S_{1.2\mathrm{mm}} M_\odot \,, \tag{6}$$

where M_D is the galaxy dust mass, $D_{(\mathrm{Mpc})}$ is its distance in megaparsecs, and $S_{1.2\mathrm{mm}}$ is its flux density at 1.2 mm. Using this equation, one could estimate that

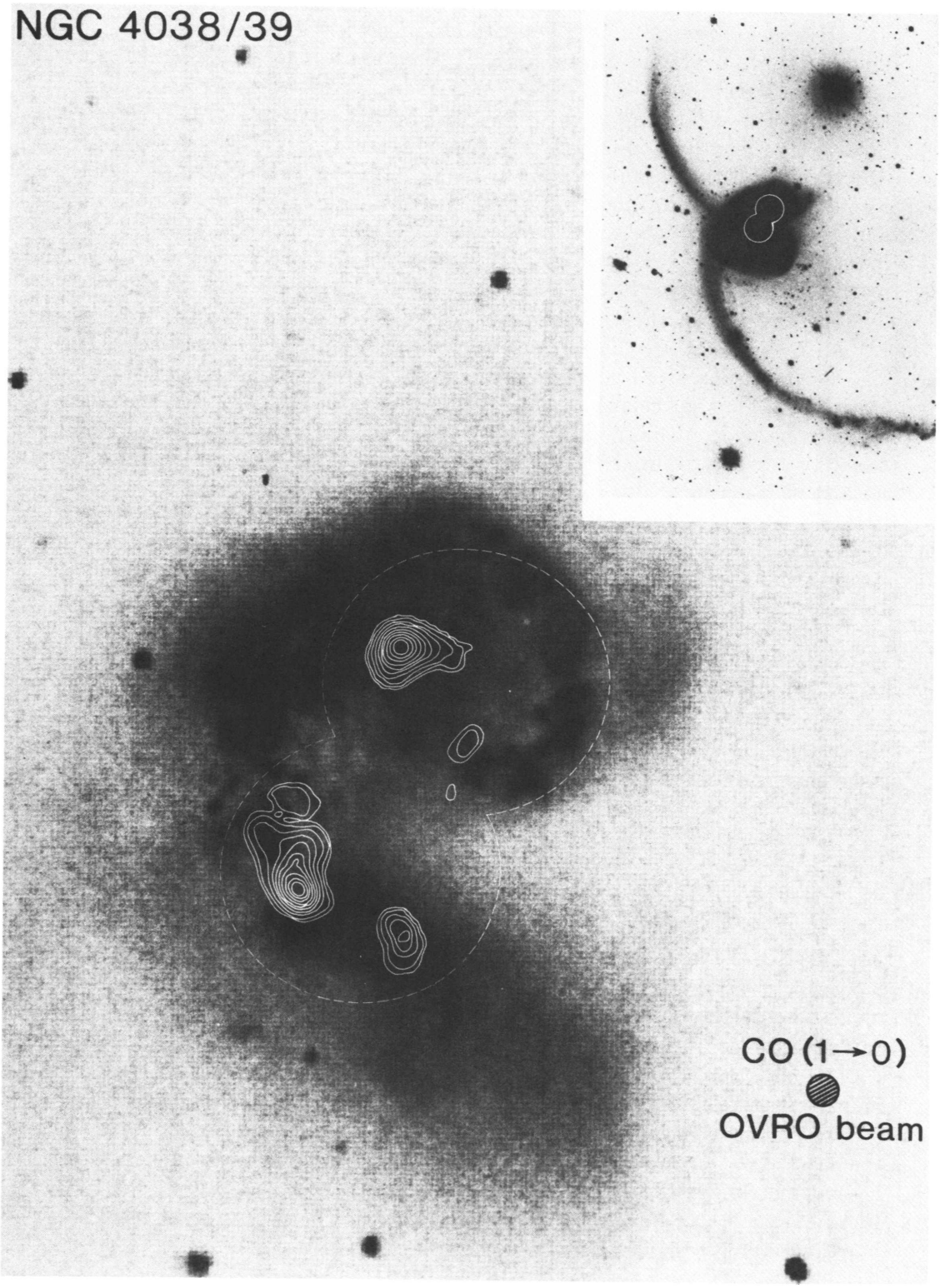


Figure III-5. The 'Antennae' galaxies, NGC 4038, are shown in an optical photograph upon which integrated-intensity CO ($J = 1$–0) contours have been superimposed. [From A. Stanford, A. Sargent, D. Sanders, and N. Scoville (1990), preprint.]

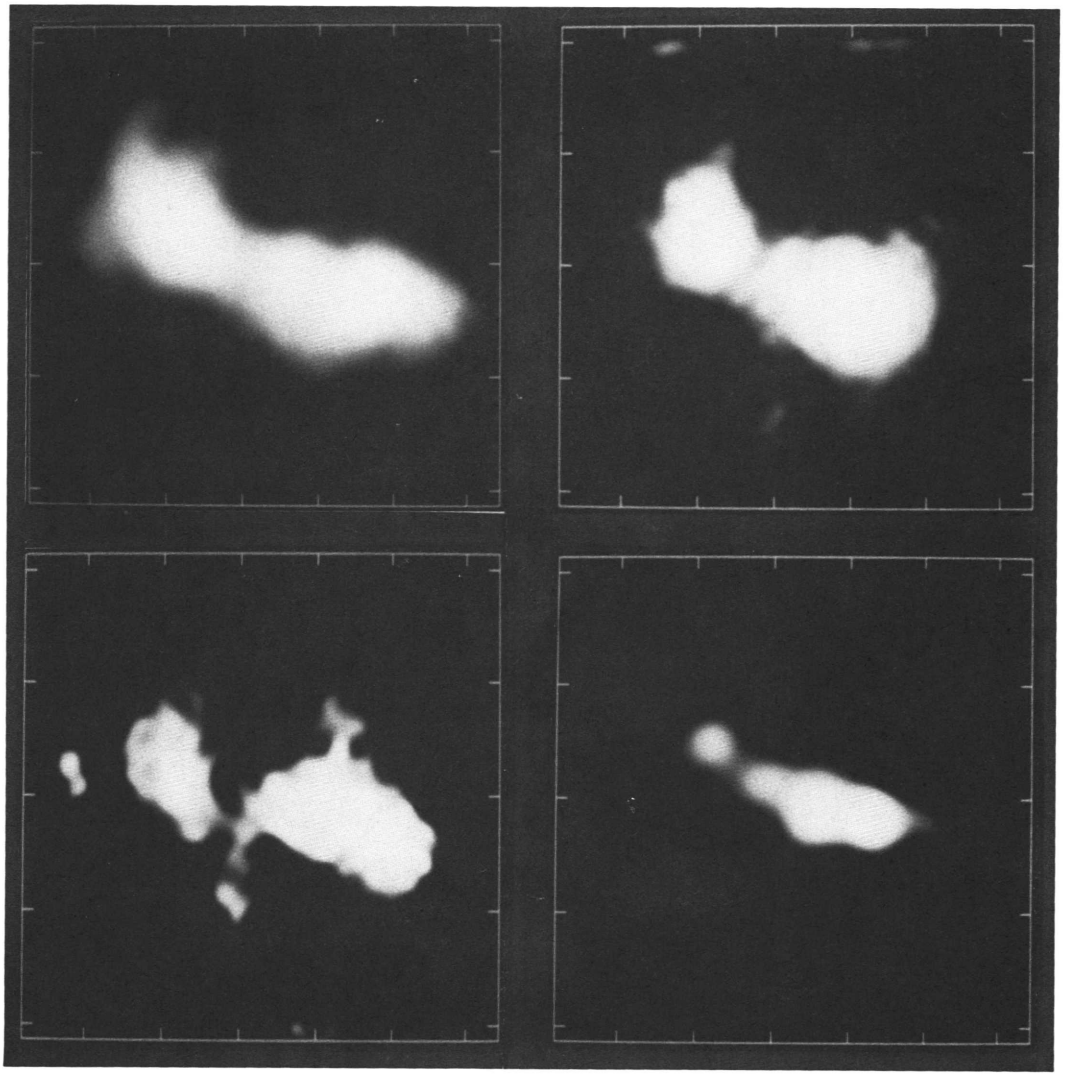


Figure III-6. The nuclear region of the starburst galaxy M82. The upper two images, left to right, are (1) integrated CO ($J = 1$–0) intensity and (2) integrated HCO^+ intensity. The lower two images, left to right, are (3) HCN ($J = 1$–0) integrated intensity and (4) the 3.3-mm continuum emission, which is mostly free–free radiation. [From J. E. Carlstrom (1988), in *Galactic and Extragalactic Star Formation*, R. E. Pudritz and M. Fich, Eds., (Reidel: Dordrecht), p. 571.]

the value of $S_{1.2\mathrm{mm}}$ for 650-Mpc distant Mk 1014, for example, should be 15 mJy, assuming a CO-derived gas mass[10] and a gas-to-dust mass ratio of 100 : 1. Such a flux density is difficult to measure with current instrumentation,[11] but a source with this flux density could be imaged with a signal-to-noise ratio of 15 : 1 in just one minute with the MMA.

[10] Sanders, D. B., Scoville, N. Z., and Soifer, B. T. (1988), *Astrophys. J. (Letters)*, **335**, L1.
[11] *Cf.* Keene, J., Carico, D. P., Neugebauer, G., and Soifer, B. T. (1990), in *Submillimetre Astronomy*, G. Watt and A. Webster, Eds., (Kluwer: Dordrecht), p. 209.

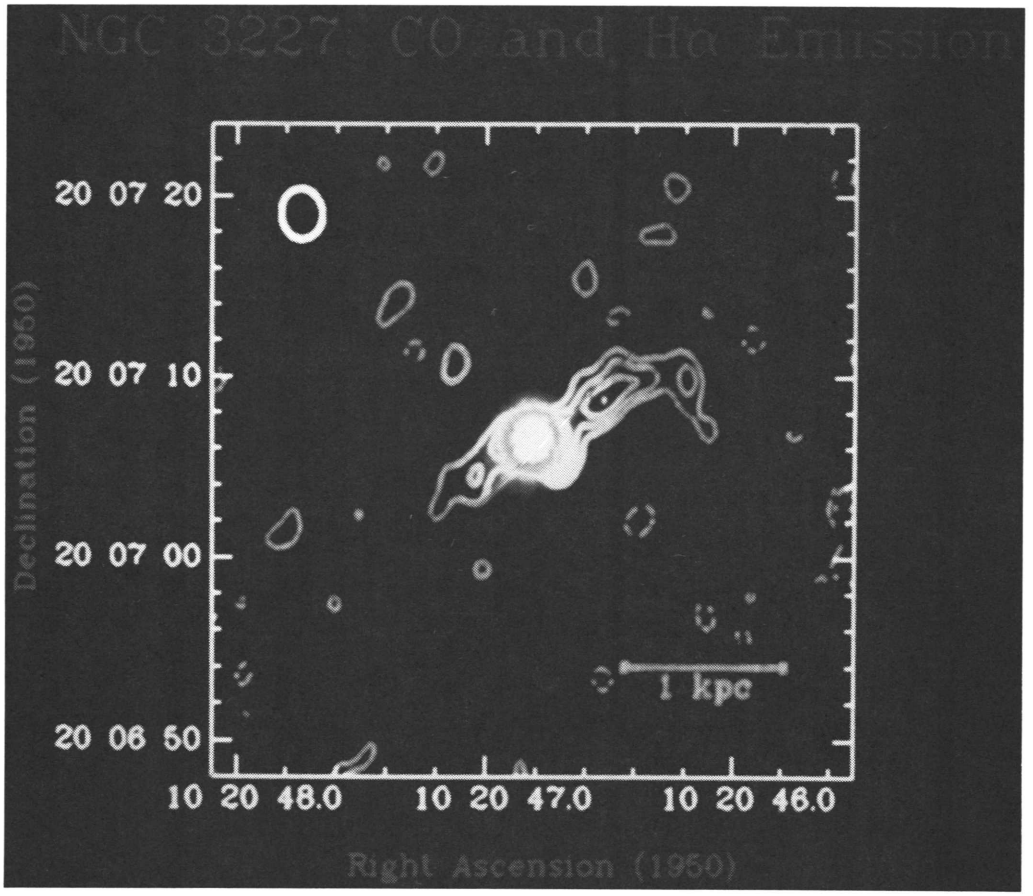


Figure III-7. Integrated-intensity CO (J = 1–0) map of NGC 3227 overlaid on an Hα photograph. The companion galaxy NGC 3226 is 2′ to the northwest. [From M. Meixner, R. Puchalsky, L. Blitz, M. Wright, and T. Heckman (1990), *Astrophys. J.*, **354**, 158.]

2.6. The Chemistry of Galactic Disks. Within the operating bands of the MMA there are numerous molecular lines which have been identified in galactic disks: higher excitation lines of HCN, CS, and HCO^+ are routinely detected in single-antenna studies and imaged with interferometers. Among the more important species known is HCO^+, a probe of the cosmic-ray ionization rate. Imaging HCO^+ in a galactic disk permits an assessment to be made of how the cosmic-ray energy density depends on different galactic environments and how it varies throughout galactic disks. Emission from all three of the lines noted above is less extended than that from the ground-state CO lines, and these trace constituents serve to delineate the denser parts of the gas, perhaps the sites of current star formation. By comparing different CO transitions and molecular lines of varying excitation, many of which can be studied simultaneously with the MMA, the physical conditions of the molecular gas may be determined. Moreover, the sensitivity of the MMA will permit many more lines to be imaged successfully and to greater distances than currently possible, yielding unique

insights and detailed information on the morphology, size scale, and energetics of star formation in galactic disks.

MMA images of molecular gas, when combined with H I, Hα, and continuum images, will provide a fairly complete picture of the global properties of the interstellar medium, and the effects of spiral structure, H II regions, and supernovae in a large number of galaxies. The kinematic information obtained from the MMA will be superior to all of the other tracers of the dense ISM. Extinction, which so hampers Hα studies, will not be a limitation to the MMA observations; the spatial resolution at CO will be more than an order of magnitude greater than what is possible in H I at the VLA. The high spatial resolution is of critical importance in at least two areas: determining the variation of gas scale-height with radius in edge-on spirals, and the mapping of central bars. The former is an important datum in determining the structure of possible non-luminous (dark-matter) halos, the latter in testing models of barred galaxies.

2.7. Continuum Emission as a Probe of Galactic Disks. The millimeter-wavelength region of the electromagnetic spectrum is a unique window, through which three important mechanisms of continuum emission operate. Synchrotron, free–free, and thermal-dust emission contribute in proportions that are functions of frequency and source type. In the disks of spiral galaxies, for example, each of the three mechanisms dominates in its own specific region of the millimeter-wavelength spectrum. Figure III-8 shows a continuum spectrum of the integrated emission of the spiral galaxy NGC 3310. Note that synchrotron emission dominates for frequencies less than about 30 GHz, owing to its relatively steep spectrum. Optically thin free–free emission dominates in the 30–100 GHz range, because of its flat spectrum. At higher frequencies, thermal-dust emission overwhelms the other contributors. Note that the 30–100 GHz capability of the MMA is ideal for study of galactic free–free emission uncontaminated by a significant non-thermal contribution. As these MMA images accurately trace 10,000-K photoionized gas, they will be excellent indicators of active star-forming regions. Free–free emission is not susceptible to excitation and absorption effects at these wavelengths; it can yield accurate estimates of the galactic bremsstrahlung emissivity even in edge-on systems. This, in turn, can be used to determine star-formation rates and to calibrate our understanding of the similarities, and differences, in the initial mass function in galaxies of differing Hubble types.

High-frequency MMA images—images made at 100 GHz and higher—will reflect the distributions of cool dust in the disks of spiral galaxies; if one may assume that the distribution of hot-dust emission follows the distribution of H II, then the cool-dust/hot-dust ratio can be determined as a function of position in a galaxy. It will be particularly instructive to compare the relative distributions and flux densities of the two components in normal versus starburst galaxies.

A comparison of free–free and CO morphologies of galaxies will be used to explore the global relationship between H II and CO and the nature of triggers of global star-formation. Individual star-forming regions in nearby galaxies like M33 ($0\farcs1 = 0.3$ pc) can be studied in detail to gain better understanding of the physics of the processes that initiate and sustain star formation.

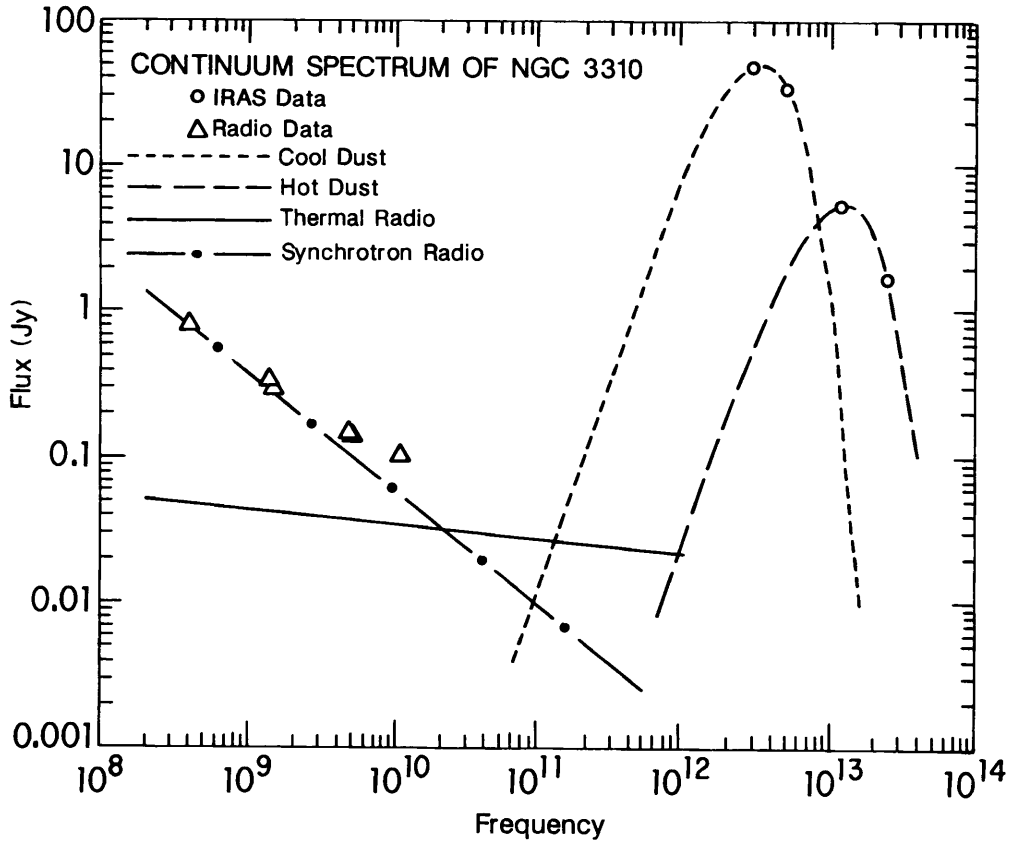


Figure III-8. The continuum spectrum of the spiral galaxy NGC 3310.

3. The Sun and the Stars

The Millimeter Array will provide images of high spatial, kinematic, and temporal resolution, enabling astronomers to confront important areas of solar and stellar science. The MMA will:

- Observe stars in every part of the H–R diagram.
- Detect photospheric emission from over 600 stars, in under ten minutes each; determine temperature gradients; and measure positions to astrometric accuracy.
- Detect stellar winds from hundreds of evolved stars and novae.
- Study flare phenomena on the Sun and other stars, and investigate the energy sources in cataclysmic variables and X-ray binaries.
- Reveal physics of particle acceleration in solar flares by combining snapshot MMA images with AXAF (X-ray) and GRO (γ-ray) images.
- Investigate causes of the solar thermal bifurcation by imaging thermal gradients.
- Provide metrology of solar oscillatory modes which will be complementary to GONG and SOHO studies.

3.1. Main-Sequence Stars. With the sensitivity of the MMA (e.g., $\sigma \approx$ 1 mJy for a one-minute integration), 5-σ detections of the thermal radiation from the photospheres or low chromospheres of more than six-hundred stars will be possible, with integration times of only a few minutes each. A few of these stars are main sequence—but most are giants, and they will be detectable simply because the MMA, for the first time, will provide the needed sensitivity and angular discrimination at short wavelengths.

Figure III-9 is a Hertzsprung–Russell diagram showing the stars that we expect to have flux densities in excess of 1 mJy, together with the limiting distances at which stars of various spectral and luminosity classes could be detected with the Millimeter Array. It also shows the approximate numbers of stars expected to have flux densities of various levels. These numbers are probably lower limits because any effect that increases the effective temperatures or stellar radii (e.g., dense transition regions, coronae, or winds) would allow detection of even more stars. Figure III-9 was derived from the 9110 stars of the Yale Bright Star Catalogue; catalogued values of spectral class, luminosity class, apparent magnitude, and color index were used to derive bolometric magnitudes, effective temperatures, distances, and stellar radii. From these, the 1.3-mm flux densities were derived. Note that the MMA may be used to observe stars in every part of the H–R diagram.

Because $S \propto T_{\rm eff} R_\star^2$, millimeter-wavelength observations can check on the canonical values (mostly inferred indirectly, a few measured) of $R_\star$, or with less sensitivity, $T_{\rm eff}$. One should expect surprises, with radio sizes or effective temperatures being different from the values anticipated; such discrepancies will be particularly interesting and important.

The high sensitivity of the MMA enormously broadens the scope of stellar millimeter-wave research. The number of detectable stars is proportional to $S_{\rm lim}^{-1.5}$, where $S_{\rm lim}$ is the limiting sensitivity. With less than an hour of integration time $S_{\rm lim}$ will be less than 20 μJy, and more than 10,000 stars may be observed. This just refers to detection: the precision of positions, sizes, etc., is proportional to the signal-to-noise ratio. Hence there is an overriding need for high $A_{\rm eff}$, wide bandwidth ($>$ 10 GHz), low system temperature, and self-calibration of weak sources. The forty antennas will facilitate accurate self-calibration.

Stellar research is one area that would greatly benefit from MMA baselines significantly longer than 3 km (a possible future enhancement of the MMA). Briefly, if the MMA angular resolution could be improved to $0\farcs01$, then: (1) the stellar disks of sixty or more nearby stars could be resolved and their photospheric radii determined; (2) starspots and active regions fixed to the stellar surfaces would provide the stellar rotation periods; (3) subtle stellar pulsations which modulate the size and flux density could be investigated; and (4) accurate metrology of star positions, over periods longer than one year, with milliarcsecond resolution, would reveal the binary nature of some stars and perhaps demonstrate the existence of low-mass (planetary) companions.

3.2. Winds of Hot Stars. Observations at 1.3 mm and 3 mm are of fundamental importance to studies of winds from hot stars. Current problems with

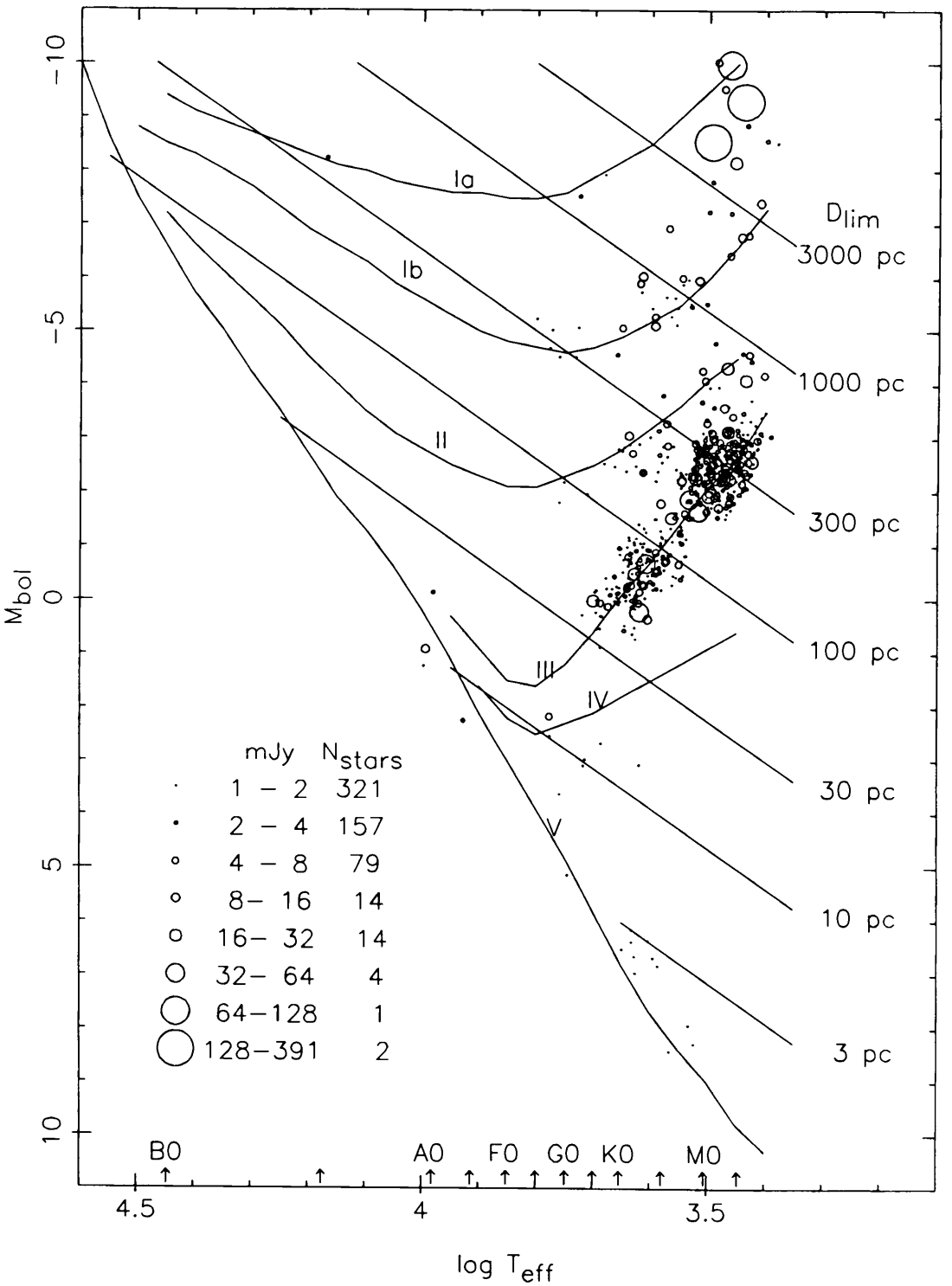


Figure III-9. Hertzsprung–Russell diagram showing stars that, with a flux density of 1 mJy or greater, will be detectable with the MMA at λ 1.3 mm. Also shown are the limiting distances to which stars of various spectral and luminosity classes could be seen with a flux density of at least 1 mJy. [See Sec. 3.1 for further discussion.]

wind flux densities at centimeter wavelengths (e.g., variability, spectral oddities, and failure to fit expected visibility curves with model temperatures) argue that millimeter-wavelength data will be critical in sorting thermal from nonthermal radiation.

The simple theory of radio emission from winds predicts:

$$S_\nu \propto \nu^{0.6} \qquad \text{and} \qquad R_\star \propto \nu^{-0.6}\,; \tag{7}$$

and typical values at 2 cm are: $S_\nu \approx 100$ mJy and $R_\star \approx 0\farcs1$. As shown below, this rising spectrum allows many wind sources to be detected at millimeter wavelengths. Effects of nonthermal electron populations are much smaller at millimeter wavelengths than at centimeter wavelengths, so that 'pure' wind radiation should be more common, and uncertainties in derived mass-loss rates should be smaller.

Because of the $\nu^{0.6}$ spectrum, the flux density S_ν of wind sources is roughly ten times larger at 1.3 mm than at 6 cm. The mass-loss rate $\dot{M}$ goes as $\dot{M} \propto S_\nu^{3/4}$. The MMA sensitivity at 1.3 mm, about the same as currently available on the VLA at 6 cm, implies that the MMA can detect mass-loss rates six times smaller than can be detected by the VLA. Alternately, sources with the $\dot{M}$ now detectable at 6 cm with the VLA would be observable with the MMA out to a distance about three times farther than is now possible.

How many stellar-wind sources would then be detectable with the Millimeter Array?

(i) Wolf–Rayet (W–R) stars: twenty-four W–R stars have been detected with the VLA at 6 cm, roughly all such stars within 2.5 kpc. With the Millimeter Array, all known Galactic W–R stars at observable declinations will be detectable, roughly doubling the sample size, depending somewhat on atmospheric attenuation vs. declination.

(ii) O–B stars: About sixteen O–B stars thought to be wind sources have been detected with the VLA at 6 cm. This number will be dramatically increased with the Millimeter Array because of two factors: First, stars of lower $\dot{M}$ will be detectable within the present limiting distance. Given the relation $\dot{M} \propto L_\star^{1.6}$, the luminosity threshold for detection will be lowered by about a factor of three. Second, stars like those presently observable will be detectable at larger distances. Together, these two effects imply that the MMA will observe the outflow winds from more than two hundred O–B stars.

There are two main reasons why detecting more stellar winds represents an important scientific goal: First, with the relatively small number now observable there are uncertainties related to small-number statistics. More importantly, the presently known objects may be the outliers of a population, unusual objects not representative of the whole. Second, lowering the $\dot{M}$ threshold will enable both UV- and radio-analysis techniques to be applied to the same stars. Currently, almost all stars with $\dot{M}$ large enough for radio detections have UV resonance lines that cannot be analyzed because they are completely saturated.

3.3. X-Ray Binaries. In the accretion environments of X-ray binaries (XRBs) the coupling between radio, optical, ultraviolet, and X-ray emission has recently been found to be much tighter than expected, with changes in all regimes varying similarly on time scales as short as tens of minutes. This leads to the concept of a 'quasar analogy', with populations of 'radio-loud' and 'radio-quiet' X-ray binaries: the radio-quiet XRBs may have the peaks in their spectra at millimeter or sub-millimeter wavelengths (as seen for some quasars). The recent outburst of V404 Cyg has exhibited a very hard spectrum and an evolution to a more and more inverted spectrum. The well-known $M/\dot{M}$ scaling of flows around compact objects is the basis of the quasar analogy since it implies an independence of scale and seems to be applicable to the XRB, quasar, and AGN environments.

Cyg X-1, Cyg X-3, SS 433, V404 Cyg, and LSI+61°303 are binary stars with strong X-ray and γ-ray emission. Their radio spectra are now known to have variations over short time scales, sometimes with the intensity reaching its peak at wavelengths shorter than one centimeter. Systematic observations over a range of frequencies are required in order to study the emission from newly accelerated, evolving electrons close to the region of acceleration, and the 'quasar-like' transition region mentioned just above. At millimeter wavelengths the MMA sensitivity, frequency flexibility, and angular resolution make this possible.

3.4. Novae and Recurrent Novae. At millimeter wavelengths one can observe nova shells in the early stages of expansion, including a time period coincident with optical studies. The combined optical/radio data will be much more powerful than either by itself in constraining models of shell density and velocity. Outbursts such as the one that occurred on RS Ophiuchus in January 1985 will be especially revealing when they can be observed with the MMA. In this outburst there were two components, an extended shell and a compact component that had a spectrum still rising at 22 GHz and an observed brightness temperature $T_B \gtrsim 10^6$ K. Possible causes for the radiation include synchrotron emission from shock-generated fast electrons, and fast electrons trapped in some sort of stellar magnetosphere. Millimeter-wavelength observations, spectra, and resolution of source sizes and positions are essential in distinguishing among models and understanding the evolution of high-energy particles.

3.5. Outbursts from Active Stars and Close Binary Systems. Here we are mainly concerned with flare stars (dwarf stars with emission lines of spectral class dMe) and close binaries such as those of type RS Canes Venaticorum (RS CVn). Millimeter-wavelength observations offer the important advantage of allowing detection of fast electrons much closer to the acceleration site, in the same way as for the Sun and many other objects. Thus it is to be expected that the correspondence between optical/UV flare emission and millimeter-wavelength emission will be close. The centimeter/decimeter-wavelength flare emission is less discriminating because it arises in fairly large, optically thick sources located high in a stellar corona, perhaps at a distance of $\sim R_\star$ above the photosphere, whereas millimeter-wave sources should be only a small fraction of a stellar radius in size and located much lower. Hence flares

can be associated better with centers of activity (starspots or active longitudes) than has been possible at centimeter wavelengths.

3.6. Solar Gamma-Ray/Millimeter-Wave Flares. One of the most important solar problems that will be addressed with the MMA is that of γ-ray/millimeter-wave flares. Recent evidence (mostly from the Solar Maximum Mission spacecraft and millimeter-wave radiometers) has demonstrated that electrons and protons are accelerated almost simultaneously to very high energies. In particular, electrons attain energies of 10 to 100 MeV within one or two seconds of flare onset, and emit both millimeter-wavelength radiation and continuum gamma-rays of high intensity. This continuum radiation is accompanied by nuclear gamma-ray lines at energies less than about 10 MeV due to protons, and neutrons are sometimes detected at Earth.[12]

Presently, there is no widely accepted explanation for the very rapid acceleration. Some argue that a 'first-phase' process must be the cause because of the very short time scale, possibly a process involving electric fields in double layers. Others argue[13,14] that stochastic acceleration can act on short enough time scales.

With the firm evidence that some physical process produces relativistic electrons on the Sun so rapidly, we must accept that the same process operates at other locations where magnetic energy is available—i.e., on a wide range of stars including flare stars, interacting binaries, and X-ray binaries. Possibly, but less surely because of the uncertain role of magnetic fields, the same process may be important in active galactic nuclei.

The special characteristic of γ-ray/millimeter-wave flares is that the flux density increases with frequency into, and perhaps beyond, the millimeter-wavelength regime. Figure III-10 shows spectra of several flares, most of them recorded before the γ-ray connection was realized. A rapid succession of images at millimeter wavelengths is needed: there have been no spatially resolved studies at either millimeter wavelengths or in γ-rays. Instruments being designed may achieve arcsecond resolution in hard X-rays (a few hundred keV energy), but not in the energy range > 10 MeV of greatest importance for understanding particle acceleration. Hence the need for the MMA: simultaneous multi-frequency images in all the Stokes parameters are essential. The instantaneous MMA beam must be excellent, viz., low sidelobes.

An indication of the breadth of this research comes from recent BIMA observations of solar flares at 3 mm with three antennas, a spatial resolution of 1–5 arcseconds, a sensitivity of $\sim 10^2$ Jy, and a time resolution up to 0.3 seconds. BIMA observed burst-sources with scales of 2–5 arcseconds or smaller. Temporal variations of fringe amplitude and phase during some flares suggest simple spatial structures, and in other flares they imply more than one brightening within the flaring region.

---

[12] *Cf.* Chupp, E. L. (1984), *Ann. Rev. Astron. Astrophys.*, **22**, 359.
[13] Melrose, D. B. (1983), *Solar Phys.*, **89**, 149.
[14] Bai, T. *et al.* (1983), *Astrophys. J.*, **267**, 433.

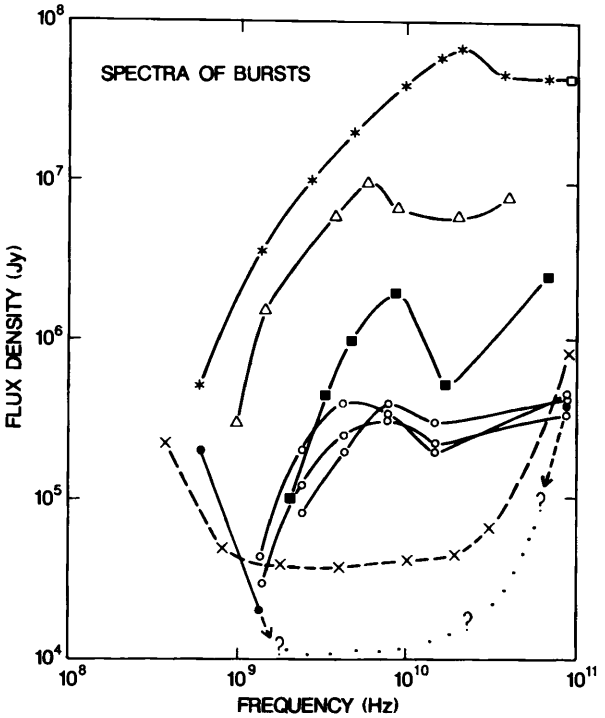


Figure III-10. The spectra of several solar flares. [Adapted from P. Kaufman *et al.* (1985), *Astron. Astrophys.*, **157**, 11.]

These first observations emphasize the importance of two-dimensional imaging with many baselines, $\lesssim 1''$ spatial resolution, and $<$ 1-second time resolution. Only then can we confirm or deny our ideas regarding the sites of energy release, the acceleration of electrons to $\sim$1 MeV within a few seconds, and the evolution of the accelerated electrons.

3.7. Penetration of Electron Beams into the Lower Solar Atmosphere. In many flares, brightenings occur in Hα, extreme ultraviolet (EUV) radiation, and even white light simultaneously with hard X-ray bursts. There is a controversy over the cause of these brightenings, whether due to electrons, to protons, or to an ion-acoustic conduction front. For the first it is uncertain whether they are able to penetrate deeply enough into the dense atmosphere; for the second it is unknown how an adequate number can be accelerated in the required one second; and for the third it is uncertain whether heat conduction can be fast enough.

Observations with the MMA will help answer these questions because the millimeter-wavelength radiation arises in the relevant region of the atmosphere, the low chromosphere, in contrast to centimeter-wavelength radiation that, during flares, originates in the corona. The millimeter-wavelength emission mechanism in most flares is probably thermal bremsstrahlung from the heated plasma, so it is straightforward to relate the brightness to the density–temperature structure in the heated regions. The relative timing of millimeter-wavelength vs. centimeter-wavelength bursts should help distinguish among the possible causes.

3.8. Imaging Solar Active Regions, Solar Filaments, and Prominences. Millimeter-wavelength emission from active regions is due to free–free bremsstrahlung, and is partially polarized because of the difference between x-mode and o-mode emissivities. It therefore gives information about the magnetic-field strengths and topology in the low chromosphere, whereas most magnetogram data apply to the photosphere, below the supposed region of magnetic dissipation in coronal heating and flares. Changes in magnetic-field topology, pre- to post-flare, should be much larger in the chromosphere than in the photosphere, and hence much more evident at millimeter wavelengths.

The variation of brightness temperature with frequency in the centimeter- and millimeter-wavelength domains provides easily interpretable information on the temperature and density structure of the transition sheaths surrounding filaments, those dense clouds of cool gas suspended high in the corona. From the temperature gradient observable at millimeter wavelengths, the thermal energy conducted into the filament from the corona can be calculated. Mini-flares, whose causes are unknown, occur in filaments: magnetic reconnections and thermal instabilities are possible causes. Comparison of high-spatial-resolution observations at millimeter and centimeter wavelengths can relate the geometries of the filament and its surrounding coronal cavity, leading to a determination of the density–temperature structure of the sheath and coronal cavity, and to reasons for their existence.

3.9. Imaging the Quiet Sun: Quiet Regions and Coronal Holes. At 36 GHz, recent results from Nobeyama[15] demonstrate that coronal holes are brighter than quiet regions, contrary to what is observed at almost all other frequencies (e.g., 10 and 98 GHz). Earlier reports of this highly unexpected effect came from U.S. and Russian observers, but seem not to have been taken seriously, perhaps because of the relatively poor spatial resolution. The cause of the anomalous brightening is unknown: it is possibly related to a lower gradient of density and temperature in the transition region of coronal holes compared with average quiet regions, or to a wider temperature plateau in the upper chromosphere. MMA images with arcsecond resolution are needed to determine whether the brightness difference is related to fine structure or widespread emission, and to compare brightness distributions at the solar limb where the emission scale height is only about one arcsecond.

3.10. Solar Observations to Challenge the MMA.

3.10.1. Coronal Heating. Mechanisms that cause heating of the outer atmospheres of the Sun and stars are not understood. For the Sun we know that deposition of mechanical energy by sound waves is inadequate. The magnetic field comes through the photosphere in small bundles with strengths greater than 1000 Gauss, and these field lines spread out by the time they get to the corona. Parker[16] and others have demonstrated that there is no equilibrium topology for helical bundles in the general case, and hence there must be dissipation of energy. Proof that this heats the corona is lacking.

[15]Kosugi, T., Ishiguro, M. and Shibasaki, K. (1986), *Publ. Astron. Soc. Japan*, **38**, 1.
[16]Parker, E. N. (1983), *Astrophys. J.*, **264**, 635.

MMA images of the brightness and polarization with better than one-arcsecond resolution may reveal the emergent magnetic-flux topology and its spreading in the chromosphere. Sites of heating in the low corona may be evident in a series of successive images taken about one second apart. If so, temperatures and densities of heated pockets could be derived.

3.10.2. Solar Recombination Lines. It has been proposed that dielectronic recombination in the solar transition region leads to significant overpopulations of high-n states of certain coronal ions such as O VI. Searches for such lines in the early 1970s using the NRAO 36-Foot Telescope yielded no detections, although the observations were drastically limited in sensitivity by the standing-wave problem common to single-antenna spectral-line observations. Even so, it is now clear that the lines are likely to be five to one-hundred times weaker than initial estimates suggested, primarily because of collisional and/or Zeeman broadening. Nevertheless, for ions with $Z \gtrsim 5$, recombination lines may be accessible to the MMA. Their detection would represent direct verification of dielectronic recombination processes in the low corona and transition region, and it would offer the possibility of measuring the magnetic-field strength in those regions through the Zeeman effect.

3.10.3. Thermal Bifurcation of the Solar Atmosphere. Imaging the solar atmosphere near the temperature-minimum region between the photosphere and the chromosphere is an extremely interesting project because there are certain theoretical and observational grounds to believe that the atmosphere is strongly thermally bifurcated due to the nature of the balance between cooling by CO and heating by the H^- ion. It has been suggested that CO cooling is highly effective in regions where the mechanical energy flux is low, so that the chromosphere, as such, may not exist. But where a large flux of mechanical energy is dissipated, CO is dissociated and is not effective as a coolant, and a dramatically higher-temperature equilibrium may exist.[17]

The details of this thermal bifurcation have not been well determined. Direct imaging over a broad range of millimeter wavelengths will reveal the thermal structure of the atmosphere in the upper and lower chromospheres, both horizontally and vertically. Changes are expected on time scales less than about one hour, the time for thermal structures to change from a high-state ($T \gtrsim 5000$ K) to a low-state ($T \lesssim 4000$ K) equilibrium. Hence flexibility in frequency selection and the complete u-v coverage of the MMA are required.

3.10.4. Oscillations and Pulsations. Finally, it may be possible to observe low-order modes of solar radial and torsional oscillation by metrology. The 3-km array configuration is approximately the right size to make such measurements and contribute to this goal, one that is of fundamental importance in solar and stellar structure, and one that will be addressed by major observing programs on the SOHO (Solar-Heliospheric Observatory) spacecraft and the GONG (Global Oscillation Network Group) project of NOAO.

[17] Ayres, T. R., Testerman, L., and Brault, J. W. (1986), *Astrophys. J.*, **304**, 542.

4. Molecular Clouds and Star Formation

Star formation is the mechanism that controls the structure and evolution of galaxies and the buildup of heavy elements in the universe with time, and it is the one responsible for the creation of the planetary environments in which life in the universe has become possible.

The MMA will be unique in its ability to discern the signature of protostellar collapse on size scales like the solar system's. We know that star formation involves gravitational collapse, but infall motions forming a new star have yet to be found. To observe unambiguous evidence for collapse, we require high spatial and velocity resolution (to map the velocity field across small structures) and high sensitivity (to take advantage of the spatial and velocity resolution). Furthermore, this must be available at a wavelength at which the collapsing object emits, and at which the surrounding material is transparent. Of current and planned instruments, only the MMA has these characteristics.

Further, the MMA will be ideal for studying the diversity of objects and physical processes involved in star formation. Its complete instantaneous u-v coverage in its most compact configuration, together with appropriate mosaicing, will allow astronomers to study the characteristics of the parent molecular clouds from which stars form. Its sensitivity, angular and velocity resolutions, and high-frequency performance will allow the study of smaller structures—including protostellar fragments, outflows, and disks.

The MMA will:

- Resolve cloud fragments as small as 10 AU and detect fragments a few Jovian masses in size.
- Identify regions of star formation in dark clouds with extensive spatial and kinematic mosaiced images.
- Image the density and velocity structures of protoplanetary disks.
- Reveal the kinematics of the earliest phases of binary-star formation.
- Unveil the physical structure of circumstellar disks and determine their role in confining material outflow from protostars.
- Determine the orientation and evolution of the magnetic field in circumstellar disks, from polarization observations of thermal dust emission.

4.1. Characterization of the Turbulent Structure of Interstellar Clouds. Understanding star formation begins with understanding the structure of a molecular cloud before stars form. The simplest scenario, in which a cloud simply collapses to form a star, does not seem realistic in general. Instead the velocity fields of molecular clouds are dominated by supersonic turbulence, the source of which is a mystery to be solved. A powerful probe of turbulence is the study of velocity fields as a function of scale size. For example, do they follow a Kolmogorov spectrum? Existing studies cover only about one order of magnitude in size. The MMA provides a unique capability to provide information over a much larger range of scales. The speed provided by the MMA will allow investigations in turbulent scales to be done on several clouds with differ-

ent properties (e.g., those forming high-mass stars, those forming low-mass stars, quiescent clouds, and Bok globules). More importantly, the present data indicate that line widths decrease with decreasing scale size, so that the transition to subsonic turbulence can be explored only with the Millimeter Array's capability for high spatial resolution. An interesting aspect of these observations is that the relevant domain for investigation of the problem is not the image plane, but rather the spatial frequency plane which is naturally provided by an array. The spectral lines that seem best suited to this project are the ^{13}CO and $C^{18}O$ $J = 2$–1 lines. For denser regions, the CS (e.g., $J = 2$–1 or 3–2) or H_2CO (e.g., 2_{12}–1_{11} or 3_{13}–2_{12}) lines may be desirable, in order to select shorter line-of-sight distances.

4.2. Identification and Characterization of 'Protostellar' Fragments. In pursuing the turbulent structure to smaller scales, one hopes to begin isolating individual turbulent elements (fragments) which could eventually become individual stars, or which might form cold, failed stars or other bodies. In particular, the distribution of angular momentum on small scales can be studied. By protostellar, we mean that no luminous condensed object (a main-sequence star or pre-main-sequence star) has yet formed. Since the gravitational energy released by contraction is effectively radiated away during the phases we wish to study here, we assume that the temperature remains low. These objects are in the natural domain of the Millimeter Array. The best way to identify these fragments is through their dust continuum emission at short wavelengths. As long as the dust temperature does not drop too low, the brightness temperature will be proportional to the dust temperature and the column density. Then, in the absence of temperature gradients induced by internal sources, the interferometer will respond primarily to enhancements in the column density which occur as a fragment contracts. Since the opacity of the dust is so low at 1 mm, it will remain a good tracer of column density up to $N = 4 \times 10^{25}$ cm^{-2}. A $1M_\odot$ protostar in spherical collapse would reach an average column density sufficient to make $\tau_{1\,\mathrm{mm}} \approx 1$ only at radii approaching 165 AU ($\approx 1''$ at the distance of the nearest clouds). Similar column densities and sizes are observed toward the densest, most compact cores, such as IRAS 16293–2422.[18] The MMA would provide detailed images of such objects; its sensitivity allows it to probe much deeper into the mass spectrum than this $\sim 1M_\odot$ object. Figure III-11 illustrates the size, other characteristic parameters, and the morphology of this typical protostellar region.

4.3. Characterization of Kinematic and Physical Properties of Protostars. Once candidate protostellar condensations are identified, the next task is to characterize their kinematic and physical properties. For kinematic studies any spectral line would serve, but one would choose CO $J = 2$–1 for maximum sensitivity, once again using the interferometer's ability to resolve out the ambient cloud. If opacity in the surrounding cloud distorts the emission from the fragment, then lines of lower opacity or higher characteristic density can be used to discriminate against the emission from the ambient cloud. As the fragment

[18] *Cf.* Mundy, L. G., Wootten, A. and Wilking, B. A. (1990), *Astrophys. J.*, **290**, in press.

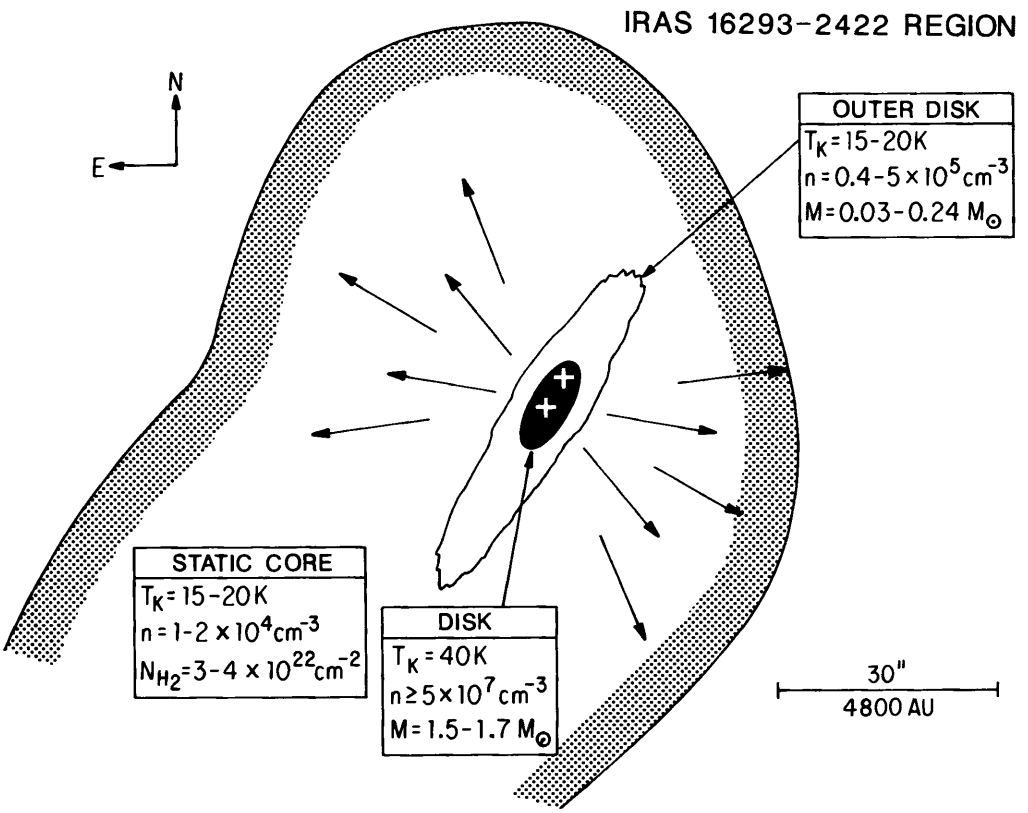


Figure III-11. A schematic illustration of IRAS 16293−2422 and its surrounding region, along with physical parameters describing its major components: static core, outer disk, and disk regions. The 'plus'-symbols within the disk region represent the centimeter-wave radio sources. The long arrows outside the outer disk represent the outflow. The hatched outer line approximates the 4.5 K km s^{-1} contour of the integrated $C^{18}O$ $J = 1$–0 emission (T^*_{rcvr}) from the static core [C. K. Walker and B. A. Wilking (1989), unpublished data]. The binary-star components located by the two centimeter-wave sources lie separated by a projected distance of $\sim$ 4 solar-system diameters.

shrinks, these lines will become comparable in brightness to the CO emission from the fragment. Convincing evidence that a fragment is truly protostellar will require the kinematic signature of collapse: infall combined with rotation. For this purpose, high spectral and spatial resolutions are required to map the line centroid and shape across the object. The ordered motions of collapse and rotation are quite small at early stages and will only emerge from the turbulent velocity field when the object is quite small. The free-fall velocity reaches 3 km s^{-1} when a $1M_\odot$ protostar reaches a radius of 140 AU (1″ at the distance of the nearest clouds). Consequently, only with the high spatial and kinematic resolutions of the MMA will we be capable of imaging the velocity fields in these objects.

Do we now have any evidence that such protostellar condensations exist? Certainly. The IRAS survey has found hundreds of sources in nearby clouds, generally embedded in dense clumps of gas, which have no optical counterparts. Those with very low 60–100 μm color temperatures are likely to be the youngest objects. Observations made by the Stratospheric Observatory For Infrared Astronomy (SOFIA) and the Space Infrared Telescope Facility (SIRTF), extending further into the infrared and deeper in sensitivity, will enlarge our catalog of unresolved cold, dense dust clumps. Notable among sources of this class is the isolated globule B335. MMA observations with arcsecond resolution of objects like B335 will reveal the structure very near the forming star, discriminating between disks and proto-binary stars.

Binary objects, even in low-mass ranges, will also produce an unambiguous kinematic signature. Abt and Levy noted that binary stars of similar mass tend to have separations less than $\sim$ 100 AU; binaries with greatly different masses tended to lie further apart. They speculated that the formation of the nearer binaries involved fission in the protostellar fragment. Only the MMA, among currently planned instruments, will provide the angular resolution on nearby clouds needed to discover the role of fission in binary-star formation.

4.4. Identification and Characterization of Disks Around Young Stars. There is growing evidence that disks are integral features of the star-formation process and that they are commonly found around young stars. For an understanding of star and planet formation we would like to know when such disks form, and we would like to study their kinematic and physical properties. A related phenomenon is that of high-velocity (often bipolar) outflows from young stars. The disk is invoked as a mechanism to channel the flow into the bipolar form that often is observed. In other models, the disk itself is the source of the outflow. Magnetic fields may play an important role in driving or channeling the outflow. In turn, the outflow may have a major impact on subsequent star formation, by disrupting surrounding material, sweeping up dense shells, compressing pre-existing clumps, or by regenerating the general turbulent field in the cloud. It is during the final stages of evolution of the disk that we expect planet formation to occur.

4.4.1. Outflows. Addressing first the nature of the outflowing material seen as broad wings on the CO lines, we note that several arguments suggest that the outflowing matter is strongly clumped. What is the nature of these clumps and what is their fate? Resolution of the clumps spatially will be possible with the Millimeter Array. It seems clear that the matter seen in the broad CO wings is generally matter swept up from the ambient cloud. The stellar (or disk) wind is presumably faster, less massive, and at least partially ionized, as reflected by the detection of infrared recombination lines or radio continuum emission from many of the stars driving outflows. Study of these winds will be an important task for the MMA (see Sec. 3.2). Further, since many winds appear to be optically thick at centimeter wavelengths, observations of the free–free emission at millimeter wavelengths may be more effective in elucidating their structure. Also, the recombination lines (e.g., H39α or H40α near 2.6 mm) may be strong enough to detect, providing velocity information on the ionized component. If

part of the wind is neutral and molecular, then very broad wings ($\Delta v \approx 300$–1000 km s^{-1}) may appear on the CO profiles. Such wide lines, while consistent with many theoretical models, would rule out some disk-driven wind models.

At an early stage in their evolution, most (and perhaps all) stars eject material back into the interstellar medium, often in two oppositely directed jets. The mechanism by which this ejection takes place is still uncertain. It is likely, however, that the outflows play an essential role in star formation, in that they carry angular momentum away from the young star, allowing further accretion from a surrounding disk.[19]

Recent observations[20] of weak, extremely high-velocity wings on CO and HCO^+ lines show that molecular gas is accelerated to velocities of more than 100 km s^{-1} in some outflows. This molecular gas probably is entrained in a high-velocity, largely atomic wind which is ejected from the central star. Heroic efforts at Arecibo and the VLA have barely detected the H I emission associated with this gas. By contrast, observations of CO—particularly the $J = 2$–1 and $J = 3$–2 transitions—with the MMA will produce 0.″1 images of this high-velocity gas.

Images of two CO transitions and perhaps of other species such as HCO^+ or H_2O will establish the temperature, density, and ionization fraction within these high-velocity winds, and will allow direct comparison with infrared-array images of shocked H_2, S II, and other tracers of shocks. The MMA images will establish whether the molecular gas is confined to high-density clumps or bullets within the outflows. Measurements of the proper motions of such clumps, when combined with radial velocity measurements, will allow three-dimensional modeling of the outflows.

Surveys[21,22] of nearby IRAS cores have detected outflows toward nearly half of them. The increased sensitivity and resolution of the MMA should allow us to image on a linear scale of ten AU essentially all the outflows in Ophiuchus and Taurus. It will then be possible to determine at what point in the evolution of a protostar an outflow begins. By observing species such as SO, SO_2, and SiO, whose chemical abundances are apparently strongly enhanced in outflow regions, it will be possible to identify outflows even in complex regions such as clouds near the Galactic center. Studies of several transitions of less abundant isotopic species, such as ^{13}CO, will provide better determinations of the energy and momentum injected back into the parent molecular clouds. Thus, we can test whether outflows help to support such clouds against gravitational collapse.

4.4.2. Water. Although water masers have long been recognized as tracers of high-density shocks, their disconnectedness from thermal tracers in the cloud has obscured their interpretation. If thermal water could be observed, important data on the fundamental oxygen chemistry of clouds could also be obtained.

[19] *Cf.* Shu, F. H., Lizano, S., Ruden, S. P., and Najita, J. (1988), *Astrophys. J. (Letters)*, **328**, L19.

[20] Lizano, S., Heiles, C., Rodríguez, L. F., Koo, B., Shu, F. H., Hasegawa, T., Hayashi, S., and Mirabel, I. F. (1988), *Astrophys. J.*, **328**, 763.

[21] Tereby, S., Vogel, S. N., and Myers, P. C. (1989), *Astrophys. J.*, **340**, 472.

[22] Myers, P. C., Heyer, M., Snell, R., and Goldsmith, P. (1988), *Astrophys. J.*, **324**, 907.

Recent observations of emission at 183 GHz of the 3_{13}–2_{20} line of water (upper level 204 K) indicate that this emission may occur both thermally and in connection with *and at similar strengths to* the 22 GHz water masers (upper level 643 K).[23] Cernicharo and his colleagues demonstrated that this line may be observed during good weather from a high mountain site, such as is planned for the MMA, in spite of the telluric extinction. Also, from the same high site (> 9000 ft), the 5_{15}–4_{22} line of water (upper level 470 K) should have a similar opacity, though it has not yet been detected. A ground-based multi-transition study of water should be possible with the MMA. Although atmospheric transmission may not exceed 50% in these lines,[24] the presence of very strong masers should permit phase-closure techniques to operate so effectively that images of the thermal water emission may be obtained. Last, note that within the next few years the Submillimeter Wave Astronomy Satellite (SWAS) will conduct a survey of water emission in the ground-state line at 557 GHz. The study of water emission—tracing active regions of star formation and the oxygen chemistry of star-forming clouds—will play a major role in MMA science.

4.4.3. Pre-stellar Disks. The densities and temperatures in disks surrounding protostars are quite different from those found in other interstellar molecular material—Beckwith *et al.*[25] find mean temperatures of 150 K for a sample of disks around pre-main-sequence stars. Measurements of the chemical structures of these disks can be made if several species are observed. A recent interferometric study has shown remarkable differences in the distributions of different molecular species across the prominent disk in Orion.[26] The 80-AU region centered on IRc2 has recently been found to have the ring-like structure of an expanding rotating disk when viewed in the light of SiO maser emission.[27] In the moderately massive disk about the embedded young star IRAS 16293–2422, NH_3 has been found to vanish in dense central regions where SO emission concentrates.[28] The physical conditions, as well as the chemistry, must have pronounced radial variations in pre-stellar disks.

4.4.4. Magnetic Fields. Although it is well recognized that magnetic fields play important roles in various astrophysical processes, observations of the magnetic-field structure of circumstellar disks and molecular clouds are sorely lacking. Recent efforts by Novak *et al.*[29] and others have succeeded in measuring 100-μm linear polarizations in Orion of up to 6% in relatively large beams. But despite intense effort, only about twenty positions have ever been observed in

[23]Cernicharo, J., Thum, C., Hein, H., John, D., Garcia, P., and Mattioco, F. (1990), *Astron. & Astrophys.*, in press.

[24]Schwab, F. R. and Hogg, D. E. (1989), *Millimeter-Wave Atmospheric Opacity and Transparency Curves*, Millimeter Array Memo No. 58.

[25]Beckwith, S. V. W., Sargent, A. I., Chini, R. S., and Gusten, R. (1990), *Astrophys. J.*, in press.

[26]Vogel, S. N. (1985), private communication.

[27]Plambeck, R. L., Wright, M. C. H., and Carlstrom, J. E. (1990), *Astrophys. J. (Letters)*, **348**, L65.

[28]Mundy, L. G. *et al.* (1990), *op. cit.*.

[29]Novak, G., Gonatas, D. P., Hildebrand, R. H., Platt, S. R. and Dragovan, M. (1989), *Astrophys. J.*, **345**, 802.

Orion. Other work has yielded magnetic-field directions from observations at 400 μm and 1.3 mm, again only in a very few sources and with large ($> 20''$ beams). The MMA will address this deficiency in two ways:

(1) The measurement of linear polarization, which gives the direction of the transverse component of the field. If the dust grains in the disk are elongated and are aligned by the magnetic field, these observations should indicate the direction of the field in the disk and cloud core. It will then be possible to study whether field and polarization alignments become more ordered in denser regions and whether there is any overall relation between the directions of the high-velocity outflows and those of the magnetic fields.

(2) The detection of the Zeeman effect, which gives the magnitude of the field along the line of sight. By differencing the left- and right-circularly polarized line emission of such a molecule as SO and measuring the frequency shift, the magnitude of the magnetic field can be mapped, so that it should be possible to trace the variation of field strength and compare it to density morphology.

These measurements will be carried out with very high spatial resolution and should resolve questions of averaging of the magnetic field when observed with larger beams.

The angular resolution of the MMA for the first time opens up the possibility of tracing the magnetic field through the star-formation process to the point of evaluating its effect on disk evolution, angular-momentum dissipation, and formation of multiple-star systems. The MMA will be particularly valuable in studying how magnetic fields influence the formation of circumstellar disks. High-resolution MMA images of the velocity and density structure in the disks will be combined with polarization data to well constrain models of disk dynamics and evolution. In addition, MMA observations of molecular ions may shed light on the ionization fraction of the gas in circumstellar disks and cloud cores, a very poorly known, but important, parameter.

4.4.5. Protoplanetary Disks. Recent observations have shown that disks exist around many young stars.[30,31] These disks control the accretion onto the stars; eventually the remnant material is the stuff from which planets form. The sizes of disks are determined by the angular momentum of the accreting gas and typically lie between 50 and 1000 AU, giving angular sizes of $0\farcs3$ to $6''$ in the nearest molecular clouds. Clearing of the inner disk, the region of planetary formation, has apparently been indirectly observed in some systems. For comparison, the solar system has a diameter of 80 AU out to the orbit of Pluto, and comets lie at distances of 1000 AU (the Kuiper belt) to 50,000 AU (the Oort cloud).

Imaging such disks with high angular resolution is critical to our understanding of star formation and the development of planetary systems. The

[30] Strom, K. M., Strom, S. E., Edwards, S., Cabrit, S., and Skrutskie, M. F. (1989), *Astron. J.*, **97**, 1451.

[31] Beckwith, S. V. W., Sargent, A. I., Chini, R. S., and Gusten, R. (1990), *Astrophys. J.*, in press.

MMA has the unique combination of sensitivity and angular resolution to make such observations possible for the first time.

Studies with the OVRO interferometer have shown the existence of Keplerian rotation in several gaseous disks on a scale of 1000 AU[32] and have also revealed that in many objects there is a denser inner disk detected in dust continuum emission. In one case, L1551–IRS5,[33] the dense disk was shown to have a radius of 45 ± 20 AU, and a column density of several hundred grams per square centimeter.

MMA images of dust continuum from a disk will reveal the column-density structure on scales down to 10 AU, the diameter of Jupiter's orbit. Spectral-line measurements on similar scales will indicate the temperature, density, and chemical composition of the disk. Velocity information will give us the rotation curve for the disk as well as the mass of the central star, a vital piece of information in understanding stellar evolution.

As a specific example, for HL Tau the peak flux density in the $J = 1$–0 CO line is $\sim$ 4 Jy, and the 2.7-mm continuum flux density is 100 mJy. The sensitivity and resolution of the MMA should permit the routine detection of such objects. For example, with a 1-km baseline, the resolution of $J = 2$–1 CO, at 1.3 mm, is $0\farcs25$. At 160 pc, the distance of the Taurus cloud, this corresponds to 40 AU. Thus it will be possible to compare the speckle infrared-dust measurements ($\sim 0\farcs3$ resolution) directly with the gas properties. It will also be possible to sample the velocity structure much closer to the star and determine whether we are observing gas in Keplerian orbits. Adopting a typical stellar mass of $1M_\odot$, the velocity at $0\farcs13$ radius and 160 pc distance should be 6 km s^{-1}. With 20-kHz filters, the velocity resolution will be 0.05 km s^{-1}, and velocity gradients will be readily discernible. Extrapolating from the HL Tau case, it should be possible to observe a reasonable statistical sample of such objects and ascertain their properties. In particular, it may be possible from the velocity information to ascertain which are likely to be protoplanetary disks and which are incipient binary systems. Numerous T Tauri stars and related objects are obvious candidates for such studies. In addition, it may be possible to make continuum measurements at 1.3 mm of the nearby disk-like structures around β Pictoris, Vega, and Fomalhout and further investigate their nature, and to study the $\sim$ 150 similar sources discovered in the IRAS database and more which are expected from SIRTF.

4.4.6. Global Properties of the Galaxy. The great sensitivity of the Millimeter Array will make possible the detection of many of the objects discussed above, throughout our Galaxy. Thus one can imagine undertaking an entire class of projects involving Galactic structure. And thus one could use the MMA to *identify protostellar clumps, bipolar outflows, or disks throughout our Galaxy.* These would serve first as test points for Galactic kinematics studies. Since the MMA would resolve most of the extended emission along the line of sight,

[32]Sargent, A. I. and Beckwith, S. V. W. (1989), in IAU Colloquium No. 120, *Structure and Dynamics of the Interstellar Medium*, G. Tenorio-Tagle, J. Melnick and M. Moles, Eds., (Springer-Verlag: Berlin), p. 215.

[33]Keene, J. and Masson, C. R. (1990), *Astrophys. J.*, in press.

revealing only the compact structures, some of the confusion in current CO surveys would be alleviated. In a similar way, SiO masers could be used to trace the velocity field in the Galactic bulge, addressing questions of bulge rotation and velocity dispersion.

The star-formation rate could be studied as a function of galactocentric radius R_{GC}, using the MMA to count protostellar cores and bipolar outflows. With sufficient statistics, one could also use the observed relations between the CO mass-loss rate and the stellar mass in order to investigate the initial mass function and its variation with R_{GC}.[34]

Of particular interest in studies of galactic structure will be the investigation of the inner regions of our Galaxy, particularly the inner several hundred parsecs, where gamma-ray observations suggest that the usual relation between CO luminosity and mass breaks down.[35] Knowledge of the nature of the gas and of star formation in this region is vital to efforts to understand nuclear starbursts in other galaxies. Again, protostellar fragments or bipolar outflows will be used to probe the velocity field, thus constraining the gravitational potential and mass distribution, as well as the star-formation activity, in this region.

The MMA will also prove a valuable probe of the inner three parsecs of the Galaxy, where many fascinating phenomena have been discovered. The MMA will be able to probe the major components in this region (ionized gas, neutral molecular gas, and dust). For this work, one would like $0''.1$ resolution.

Finally, a more speculative possibility is to study the accretion disk around the black hole. Images with very high angular resolution of the H- and He-recombination lines could push the studies of rotational velocities much closer to the center than possible with infrared techniques, constraining the mass of the central object. In addition, hyperfine lines (analogous to the 21-cm line, but at millimeter wavelengths) may be produced from H-like ions of heavy elements (via the Sunyaev–Curasov effect) in the very hot gas of the accretion disk.

5. Astrochemistry

The field of astrochemistry began in the late 1930s and early 1940s when optical astronomers first detected the simple diatomic species CH, CH^+, and CN in tenuous clouds of gas and dust in our Galaxy, but modern astrochemistry started only two decades ago with the radio detections of the polyatomic molecules NH_3, H_2O, and H_2CO in dense Galactic clouds. These polyatomic detections were a quantum leap forward for the study of astrochemistry because: (1) they motivated the radio and infrared detections of many other new molecular species; and (2) they shattered the theoretical dogma of the day that only diatomic molecules could exist in the harsh environment of the interstellar medium, which led to new and successful models of interstellar chemistry. Today there are over ninety interstellar and circumstellar molecular species known, including molecules with ten, eleven, and thirteen atoms, three small three-membered rings, ionic species

[34] Levreault, R. M. (1985), Ph. D. Thesis, The University of Texas at Austin.

[35] Blitz, L., Bloemen, J. B. G. M., Hermsen, W., and Bania, T. M. (1985), *Astron. Astrophys.*, **143**, 267.

such as protonated N_2 and protonated HCN, molecules which contain potential tracers of stellar processing such as sodium and chlorine, and three molecules containing the amide group. Indeed, so many species contribute to the spectroscopic richness of star-forming molecular clouds that, when one is observing with the limited spatial resolution of a single antenna, the density of molecular lines greatly impedes the proper interpretation of the physics and chemistry of these regions. The combination of high spatial resolution, state-of-the-art sensitivity, and an unprecedented frequency range which will be available in the MMA will set the stage for the next great leap forward in astrochemistry: the ability to spectroscopically probe the very cores of circumstellar, interstellar, and extragalactic cloud regions and observe the driving chemical reactions that steer processes ranging from the formation of galaxies, through the formation of stars, to even possibly the formation of life itself.

The Millimeter Array will:

- Provide, at 10 AU resolution, images of the chemical gradients in protostellar nebulae for nearby star-forming regions (e.g., in Taurus and Ophiuchus).
- Obtain arcsecond-resolution mosaiced images of molecular abundances throughout GMC complexes; study the changes induced by fragmentation/condensation.
- Yield the spatial variation of dissociative shock chemistry in outflow sources.
- Enable study of photochemical processes via spectral-line images of circumstellar shells.
- Probe dust formation and destruction via emission-line images of the refractory molecules that mark dust catalysis and spallation.
- Provide high-resolution data revealing processes connected with low-temperature chemistry (deuteration and grain-mantle processing) in cores of dense molecular clouds.

5.1. Ion–Molecule Chemistry. The earliest successful model of interstellar chemistry was based on ion–molecule reactions, a theory which over the years has provided a satisfactory picture of many of the smaller interstellar molecules that are known, and some of the larger ones. This chemistry is based on the cosmic-ray ionization of H_2 and of He, producing H_2^+, H_3^+, and He^+ ions which interact with many neutral atoms and molecules to produce a wide variety of molecular ions. Neutral species are formed from these ions by recombination with electrons. The major successes of this chemistry have been the semi-quantitative as well as qualitative predictions for HCO^+, HCS^+, and $HCNH^+$, as well as for neutral species such as HCN, HNC, OH, CCH, CN, and others. The tremendous chemical fractionations with respect to deuterium of several species (DCO^+, N_2D^+, DCN, etc.) are perhaps best explained by ion–molecule chemistry, though it is now clear that a number of mechanisms participate in this process.

However, ion–molecule chemistry is beset by major uncertainties (e.g., unknown reaction rates, input atomic abundances, time-dependent effects, and the role of dust). These presumably explain the difficulties in predicting the abundances of species such as C I and the cyanopolyacetylene chains, N_2H^+,

NH_3, and the abundances of most sulfur-containing compounds. In addition, there is ample evidence that clouds are far more heterogeneous, chemically, than would be expected on the basis of standard ion–molecule chemistry. The limits of applicability of this most important interstellar chemistry have not been delineated despite their importance to many of the conclusions reached in molecular astrophysics, because of limitations both in basic laboratory data and in observations—especially those involving spatial resolution and sensitivity. Recent advances in laboratory instrumentation and in theoretical methods made possible by the continuing explosion in computational hardware promise to radically alter the current lack of data, and they have already demonstrated the importance to interstellar chemistry of:

- Steep temperature dependences of many ion–molecule reaction rates at temperatures below 80 K, in both an increasing and decreasing sense;
- Radical and bare-atom ion–molecule and neutral–neutral reactions in the gas phase; and
- Photodissociation self-shielding in molecules other than H_2 (e.g., in CO).

The steep abundance gradients which are observed, and the inferred time dependences, stress more and more the close interdependence between source dynamics, energy balance, and chemical composition. No longer can the ultimate stages of star formation be studied while ignoring the chemical changes, and vice versa.

An ideal astrophysical laboratory for the study of ion–molecule chemistry is provided by the *translucent* or high-latitude cirrus clouds that contain many core regions with densities typical of dark cloud cores. These clouds appear, like dark clouds, to have fairly simple morphologies at low spatial resolution, but they possess considerable chemical complexity. Studies of H_2CO, C_3H_2, HC_3N, C_2S, and NH_3 in their cores indicate fractional abundances from ten to a hundred times lower than found in *normal* molecular clouds. These abundances actually agree very well with predictions of ion–molecule chemistry models under steady-state conditions. One seemingly attractive feature of these cirrus cores was their structural 'simplicity'; but recent detailed studies of the 2–1 and 1–0 transitions of CO indicate that they, too, contain much fine structure, unresolved in single-antenna beams, that may prevent reliable abundance determinations. These determinations are particularly important for CO because of its unique role as an *endpoint* of astrochemistry. If CO is underabundant in cirrus cores, a metal deficiency for these high-latitude clouds would be indicated, whereas a *normal* CO abundance would indicate the importance of chemical evolution, shocks, or elemental abundances to the chemistry of the vast majority of all studied molecular clouds, viz., those in the Galactic plane.

Ion–molecule chemistry in star-forming molecular clouds—in particular, those containing O–B associations—is considerably more complicated. For example, the tendency for images of a given region made in different molecular lines to differ from one another was already evident in single-antenna studies. It is even more pronounced in interferometric images. The most well-studied

region to date involves the Orion/KL cloud core, of which over a dozen images have been generated in various millimeter-wavelength emission lines, as well as in the dust continuum. Maps of species from the BIMA, Nobeyama, and OVRO arrays, along with centimeter-wavelength studies at the VLA in species as diverse as CO, CS, SiO, HCN, HDO, SO_2, NH_3, CH_3OH, etc., have now been produced on spatial scales approaching 2–3″ with superb spectral resolution. The differences in the images are clearly not caused by excitation effects, but seem rather to reflect true chemical variations within the nebula. Orion is one of the closest and most energetic sources in the sky, with total angular extent less than 60″; to study adequately other regions of this sort clearly requires the observation of many different molecular species, at high angular resolution.

Indeed, although current observations are beginning to reveal representative sizes of clumps in star-forming regions, much refinement is needed before abundances of molecules can be reliably estimated. Relative abundances of molecular species are, of course, the basis of any comparison with theories of astrochemistry. Crude abundances are often estimated under the strong assumptions of local thermodynamic equilibrium (LTE) conditions and of optically thin lines. We now know that in many complex species (C_2H_5OH, C_2H_5CN, CH_3CHO, CH_3OHCO, C_2H_3CN, $(CH_3)_2O$) those transitions with small dipole components are often observed by single antennas to be far stronger than expected on the basis of LTE. Either the assumption of low optical depth is wrong, in which case the abundances of these species are up to a hundred times larger than currently believed, or these transitions are selectively pumped, and can provide important information on the interactions of molecules with neutral particles and with radiation fields. Either possibility undoubtedly implies highly compact clumps, requiring observations with the highest spatial resolution.

Little is currently known about distant star-forming regions, because of limited spatial resolution. In general, OH, H_2O, H_2CO, and CH_3OH masers show that massive star-forming regions contain dense clumps of gas. Other indicators of such clumping now exist, but serve only to illustrate how much more needs to be learned to understand the chemistry. VLA studies of NH_3, and BIMA observations of HCN, HCO^+, and other species, seem to suggest *typical* star-forming clumps of size 0.05 to 0.1 pc, densities of 10^6 cm^{-3}, and temperatures of 100–200 K. Dozens of such clumps may be found in a single small region, as in NGC 7538.[36] Our estimates of these typical properties are affected by current instrumental limitations; the star-forming clumps may actually be denser and smaller.

5.2. Low-Temperature Chemistry. Relative to the giant molecular clouds, the dark clouds are nearby, less massive, and colder, and they lack the formation sites for massive stars. Cold dark clouds can therefore be excellent laboratories for astrochemistry, since high spatial resolution allows accurate determination of local physical conditions. Although there are some basic similarities between the abundances observed in dark clouds and in the more quiescent portions of large, warm molecular clouds, there are also fascinating differences, both among

[36] Pratap, P., Batrla, W., and Snyder, L. E. (1989), *Astrophys. J.*, **341**, 832.

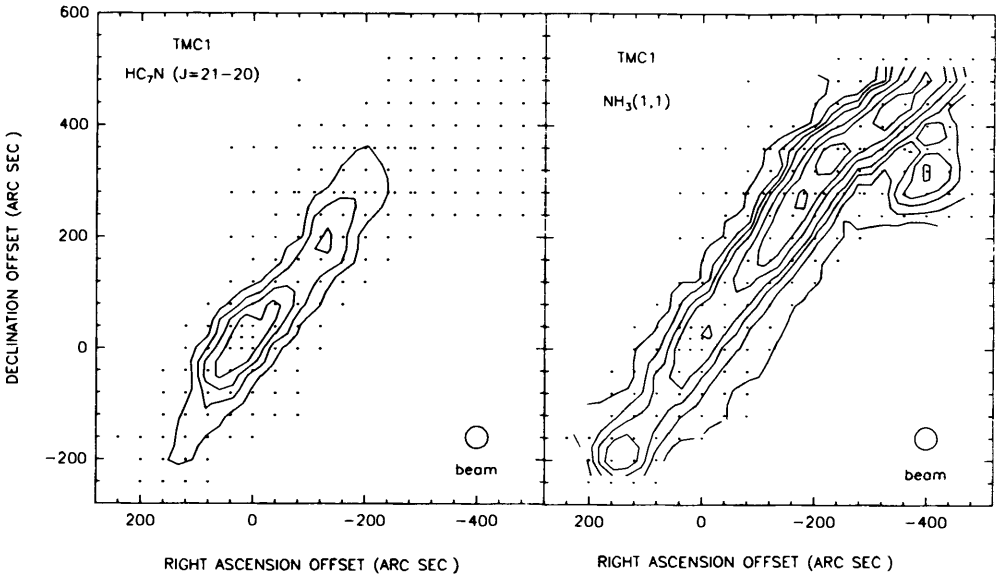


Figure III-12. Comparison of the distribution of two molecules in TMC 1 [C. A. Olano, C. M. Walmsley, and T. L. Wilson (1988), *Astron. & Astrophys.*, **196**, 194]. *Left panel:* Distribution of peak brightness temperature for $J = 21$–20 HC_7N, with contour levels at 0.5, 1.0, 1.5, and 2.0 K. *Right panel:* Distribution of peak brightness temperature in the main component of $NH_3(1,1)$. Contour levels are 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, and 5.0 K. The points represent the observed positions.

clouds and, apparently, within a given cloud. The reasons for such variations are not well understood, but are a subject of active study. They could signify that chemistry in the dense and cool clouds may not have reached a steady state. Determination of the hydrogen and carbon isotopic ratios in these clouds can also provide a critical test of the chemical models.

A well-studied example of a cold cloud with abundance variations is TMC 1, a chain of tiny condensations, or *clumps*, in which almost two score of molecules, including large linear carbon-chain molecules, have been detected. TMC 1 has been mapped in the lines of many molecules, and it is probably the dense cloud whose physical conditions are best known. Observations at high spatial resolution (see Fig. III-12) reveal that the half-resolved core clumps, whose structures are best visible in the high-excitation lines of HC_3N and C_3H_2, have sizes of 0.01 pc and masses of $1M_\odot$. These cores, which could be low-mass protostars, seem perfectly aligned—the apparent signature of magnetic confinement. They have the same low temperature and very similar densities, yet quite different abundances in species such as C_3H_2, HC_3N, HC_5N, NH_3, and HCO^+.

Finally, cold dark clouds have been fertile sites for recent searches for new molecules, including both heavy organic species, such as CH_3C_4H, and *non-terrestrial* species, such as C_3H and C_3S. Many of these have not been thoroughly mapped, so their spatial scales are not known. It seems likely, however, from comparisons with better known molecules, that the prime requisite in searches for additional species will be sensitivity—the antenna temperatures for the detections of C_3O and CH_3C_4H were on the order of only 30 mK.

5.3. Shock Chemistry. A number of molecules—CO, HC_3N, HCO^+, and HCN—have been observed in very-high-velocity molecular outflows associated with several late-type stars, in particular the *protoplanetary* nebula CRL 618.[37,38] Since these molecules could not survive the direct acceleration, which certainly involves shocks, to such high velocities—they must have formed in the post-shocked gas. As such they are therefore the clearest examples of shock chemistry. Their abundances can be compared to those predicted by current shock-chemistry models,[39] which find a sequential formation of H_2, CO, HCN, HCO^+ in the cooling post-shocked region. Better angular resolution than is currently available is needed to understand the origin and morphology of such high-velocity outflows and the related chemical processes (in particular, the locations of different molecules).

Clear evidence for shock chemistry is also found in high-mass star-forming regions. For example, HCN, SO, SiO, and HCO^+ all exhibit *plateau* or high-velocity emission in Orion; the SiO is extremely compact and centered on IRC2, while the HCN and SO are somewhat extended, and the HCO^+ is very widely distributed. Detailed investigations of the individual velocity channels have enabled dynamical comparisons of the various sources to be made—which in this case reveal that IRC2, in addition to driving the bipolar outflow traced in HCO^+ and vibrationally excited H_2, also heats the surrounding dense medium known as the *hot core*, best seen in molecules such as HDO or NH_3. Interactions of the stellar outflow with the surrounding medium can be seen at many distance scales in various types of maser activity, ranging from SiO masers which lie within 35 AU of IRC2, to the OH and H_2O masers which are scattered around the inside of the dense surrounding disk, to methanol masers visible far from IRC2 where the outflow strikes the ambient cloud material.

Understanding the local effects of shock chemistry in those cold dark clouds that harbor formation sites of low-mass stars such as the Sun is also crucial. Indeed, one of the most exciting areas of contemporary astronomy is the characterization of the birth process of low-mass stars, which may ultimately provide new insights into the origin of the solar system. Many unique chemical processes—such as accretion shocks, grain melting, etc.—will become important in the high-density, high-temperature environment of the protostellar nebula, some of which may be observable *only* with an instrument as discriminating as the MMA. For example, the predicted sizes of the disk structures around protostellar cores are at most a few hundred AU, so that the study of a flattened stellar nebula even in one of the nearest objects would require resolutions of $0\farcs2$ or better. For chemical species with strong radial abundance gradients (i.e., very refractory molecules or species produced in the accretion shock or within protoplanetary envelopes), the source sizes are smaller still. The beam dilution encountered with single-antenna observations, even at sub-millimeter

[37]Cernicharo, J., Guélin, M., Martín-Pintado, J., Peñalver, J., and Mauersberger, R. (1989), *Astron. & Astrophys.*, **222**, L1.
[38]Knapp, G. R., Sutin, B. M., Phillips, T. G., Ellison, B. N., Keene, J. B., Leighton, R. B., Masson, C., Steiger, W., Veidt, B., and Young, K. (1989), *Astrophys. J.*, **336**, 822.
[39]Neufeld, D. A. and Dalgarno, A. (1989), *Astrophys. J.*, **340**, 869.

wavelengths, is sufficiently large that the method by which the chemical evolution of protostellar nebulae occurs cannot be understood. The great sensitivity of the MMA is needed to enable us to address these questions.

5.4. Photochemistry. Study of the chemistry of circumstellar shells is particularly fascinating because it is driven by non-equilibrium processes—specifically, the diminishing influence of the stellar radiation field with radial distance from the central star, and the increasing exposure to the interstellar radiation field with time. Observations at high spatial resolution will provide an essential check on chemical models that predict molecular abundance variations as functions of radius within the circumstellar shells of evolved stars. The *freeze-out* model, in which chemical species are formed under conditions of thermodynamic equilibrium in the high-density, high-temperature photospheric region and are then propelled outward by the stellar wind with frozen abundances, no longer accounts for the observational data. In particular, the model fails to predict the large abundances of carbon-chain radicals in a number of carbon-star envelopes, the detection of HNC, a clearly non-equilibrium species, and the observations of molecules in very-high-velocity outflows (those with 50 km $s^{-1} \leq v \leq 200$ km s^{-1}).

The most likely explanation of the formation of the carbon-chain species and HNC involves ion–molecule reaction networks in the outer portion of the circumstellar shell initiated by interstellar UV-photons.[40] In this model, species such as CCH, HC_3N, and HNC should be confined to the outer envelope, while the *equilibrium* species such as CS and the refractory species such as SiO should peak near the star. This view has been largely confirmed in the case of IRC+10216, a prominent, nearby carbon-rich circumstellar envelope, by the beautiful aperture-synthesis images made with the BIMA interferometer,[41] and by the spectral-line survey observations of the IRAM 30-m telescope.[42] In this source, the reactive carbon-containing species CCH and C_3N indeed exhibit a hollow-shell distribution with a sharp cutoff at a radius of 10^{17} cm, whereas SiO, SiS, and CS each show a strongly peaked central concentration. Figure III-13 illustrates the layered chemistry in such a circumstellar envelope. Curiously, the recently discovered SiC free-radical,[43] a very refractory species, has a markedly shell-like distribution. This shows that it is most likely a product of photodissociation of the abundant SiC_2 molecule. Another means of probing the photodissociation and photoionization processes in circumstellar shells will involve high-angular-resolution images of the self-shielded CO in several isotopic forms. The MMA is the ideal instrument for such studies.

Even in giant molecular clouds there is extensive evidence that different types of chemistry occur side by side. In M17SW, which lies at the interface of the H II region and the molecular cloud, the ultracompact H II region indicates

[40]Glassgold, A. E., Mamon, G. A., and Huggins, P. J. (1989), *Astrophys. J. (Letters)*, **336**, L29.

[41]Bieging, J. H. and Rieu, N.-Q. (1988), *Astrophys. J. (Letters)*, **329**, L107.

[42]Bieging, J. H. (1990), preprint.

[43]Cernicharo, J., Gottlieb, C. A., Guélin, M., Thaddeus, P., and Vrtilek, J. (1989), *Astrophys. J. (Letters)*, **341**, L25.

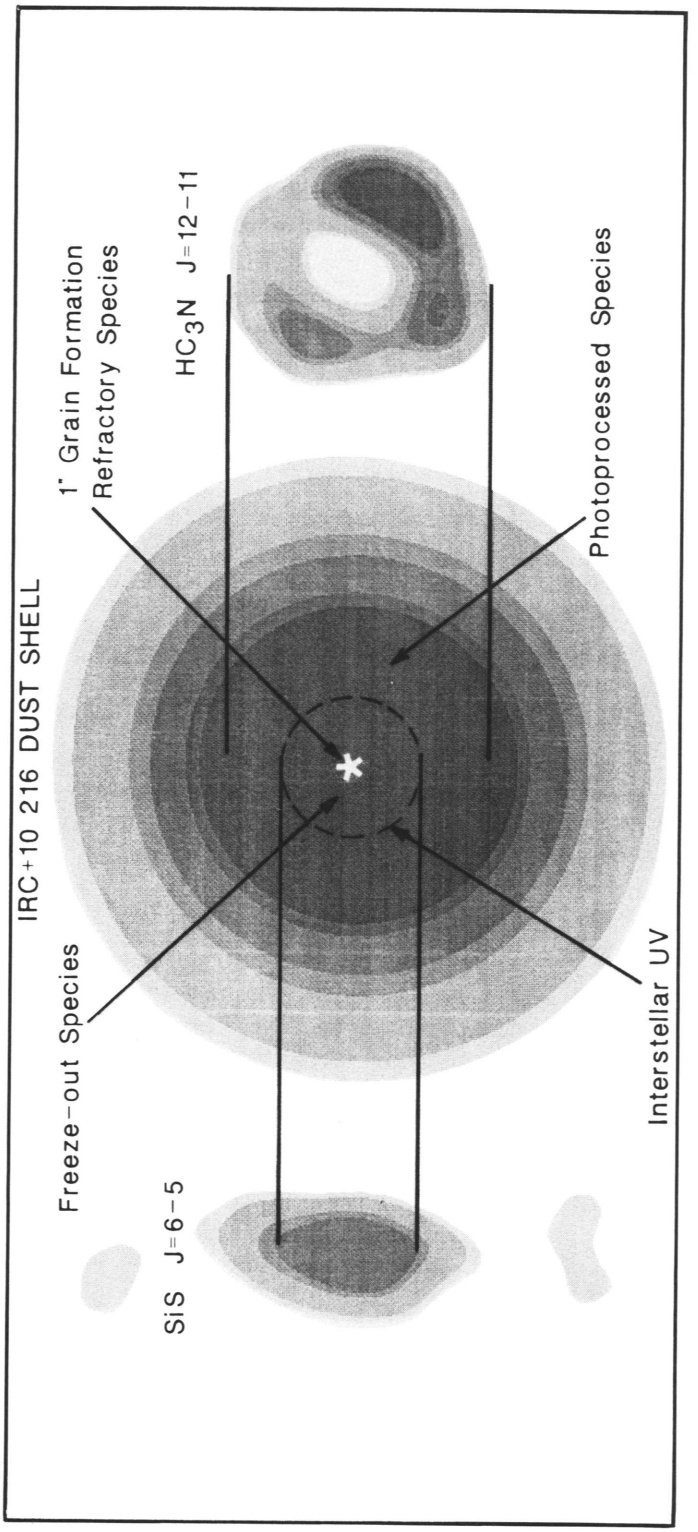


Figure III-13. This schematic illustration shows the radial density gradient in a circumstellar shell and the zones where the dominant chemistry changes. The SiS and HC_3N images are from the OVRO interferometer [Likkel, Morris, Masson, and Wootten (1987), unpublished].

the presence of a strong shock. Species such as C_3H_2 and $C^{34}S$ peak 1′ away, while several NH_3 clumps of 0.05 pc size are scattered throughout the region. It is likely that such clumps have envelopes heated by the UV from the H II region and cooler cores.[44] It is also likely that the apparent diffusion of atomic carbon throughout the molecular cloud, in conflict with chemical models that predict a definite transition zone (CO/C I/C II), is a manifestation of the penetration of the UV between and among many small clumps. Thus astrochemistry requires high spatial resolution for two major purposes: (1) to test qualitative predictions of the 'simple' structural aspects of the chemistry, such as transition zones; and (2) to determine brightness temperatures of molecular species by determining clump sizes, necessary to derive quantitative abundances. Only the MMA, with its large collecting area and sub-arcsecond imaging capability, will enable the large number of GMC star-forming cores known in the Galaxy to be studied with the kind of spatial detail now available for only the nearby Orion/KL region.

5.5. Extragalactic Chemistry. Nearly a dozen molecules, ranging in complexity from CO to CH_3OH, have now been detected in other galaxies, a dramatic increase over those observable just a few years ago. The study of at least several extragalactic molecules is important because they provide essential tools for determining whether, and perhaps why, the chemistry of other galaxies differs from that of the Milky Way. Even so fundamental a species as CO is already known to be relatively lacking in several nearby and well-known galaxies such as M31. It is not known whether molecular chemistry, as exemplified by CO, always is enhanced in spiral arms within a galaxy; the study of such arms in even the nearest galaxies requires resolution beyond that available with single-antenna millimeter-wave telescopes.

Isotopic ratios are fundamental to nuclear processing in stars, in that they depend on the stellar mass distribution, the rate of star formation, and the rate of return of stellar material to the ISM. Models of the Milky Way predict gradients in certain key isotopic ratios ($^{12}C/^{13}C$, $^{14}N/^{15}N$, $^{16}O/^{18}O$) with galactocentric radius, and it is difficult, because of distance ambiguities, to establish whether such gradients exist in our Galaxy. Observations of other galaxies will provide this information more clearly, and they will also indicate what differences there are between galaxies. Such differences, in turn, may reveal fundamental differences in the nature of star formation in galaxies. Optical studies do not provide such information, because of the severe blending and extinction affecting the various isotopic lines. Further, events such as photodissociation and accretion onto the central engines of AGNs are well matched to molecular probes in the millimeter- and submillimeter-wavelength regions. The high spatial resolution of the Millimeter Array and its tremendous sensitivity will be needed to obtain the necessary information in even the nearest galaxies.

5.6. Astrochemistry in the Solar System. The chemical compositions of planetary and satellite atmospheres are data of fundamental importance in plan-

[44]Stutzki, J., Stacey, G. J., Genzel, R., Harris, A. I., Jaffe, D. T., and Lugten, J. B. (1988), *Astrophys. J.*, **332**, 379.

this material have the potential to constrain even the details of stellar evolutionary models.

As the hydrogen in the core of an intermediate-mass star is exhausted and the star reaches the red-giant branch for the first time, ignition of hydrogen-shell burning causes the convective envelope to penetrate inward to material which has experienced hydrogen burning. Mixing this material to the stellar surface leads to an increase (roughly a doubling) of the ^{14}N abundance, an increase in ^{13}C relative to ^{12}C such that the ratio $^{12}C/^{13}C$ = 20–30, and no change at all in the ^{16}O abundance. The second phase of 'dredge-up' for stars more massive than about $3M_\odot$ occurs at the onset of He-shell burning; the convective envelope reaches inward to the helium core and mixes to the surface an enhanced abundance of ^{14}N. The star quickly evolves along the AGB as shown in Figure III-14, increasing its luminosity and driving a stellar wind at a mass-loss rate[45]

$$\dot{M} = -1.26 \times 10^{-20} \eta L_s R/M\,, \tag{8}$$

which rapidly increases with the increasing luminosity, decreasing mass, and increasing radius that accompanies this evolutionary phase. (Here L, M, and R are in expressed in solar units.) The coefficient η is a number on the order of unity. Note that both the mass and the fraction of the initial stellar mass ejected as circumstellar material into the ISM by the wind is large—viz., $3.3M_\odot$, or 80% of the initial stellar mass for the $4M_\odot$ star illustrated in Figure III-14. Frozen into the wind are the chemical abundances and isotopic ratios from the two 'dredge-up' episodes mentioned above.

Still later, expansion of the stellar atmosphere, driven by the enhanced luminosity of the core helium burning, lowers the pressure in the hydrogen-burning shell to an extent that hydrogen-burning ceases. This continues until the helium is exhausted in the core. The atmosphere then falls back inward, heating, until hydrogen ignites in the outer H-rich shell. Interior to this, both the core and the He-rich shell continue to heat without burning, until the He-shell can itself re-ignite. He-ignition again increases the stellar luminosity, lifts the atmosphere—which again quenches the H-burning shell—and the cyclic process repeats. The emergent stellar luminosity exhibits a gradual increase, punctuated by 'thermal pulses' which abruptly interrupt the increase. Since the dominant energy source for the thermal pulses is the triple-α process, the principal product is ^{12}C, which is convected to the surface, increasing the $^{12}C/^{13}C$ ratio (the third 'dredge-up'). Between pulses, ^{12}C is processed further to ^{14}N by the C–N–O cycle; all this material is carried away by the wind.

Mass ejection at the peak of the AGB branch is extremely violent. A 'super-wind' of $\sim 10^{-3} M_\odot \mathrm{yr}^{-1}$ ejects the stellar envelope in its entirety, terminating both the thermal-pulse phase and the third 'dredge-up'. Again the history of this event is frozen into the composition of the wind.

Since the 'dredge-up' episodes occur ordered in time and each alters the chemical and isotopic abundances in a distinctive way, the effects of each of these

[45] Reimers, D. (1975), in *Problems in Stellar Atmospheres and Envelopes*, B. Baschek, W. Kegel, and G. Traving, Eds., (Springer: Berlin/Heidelberg), p. 229.

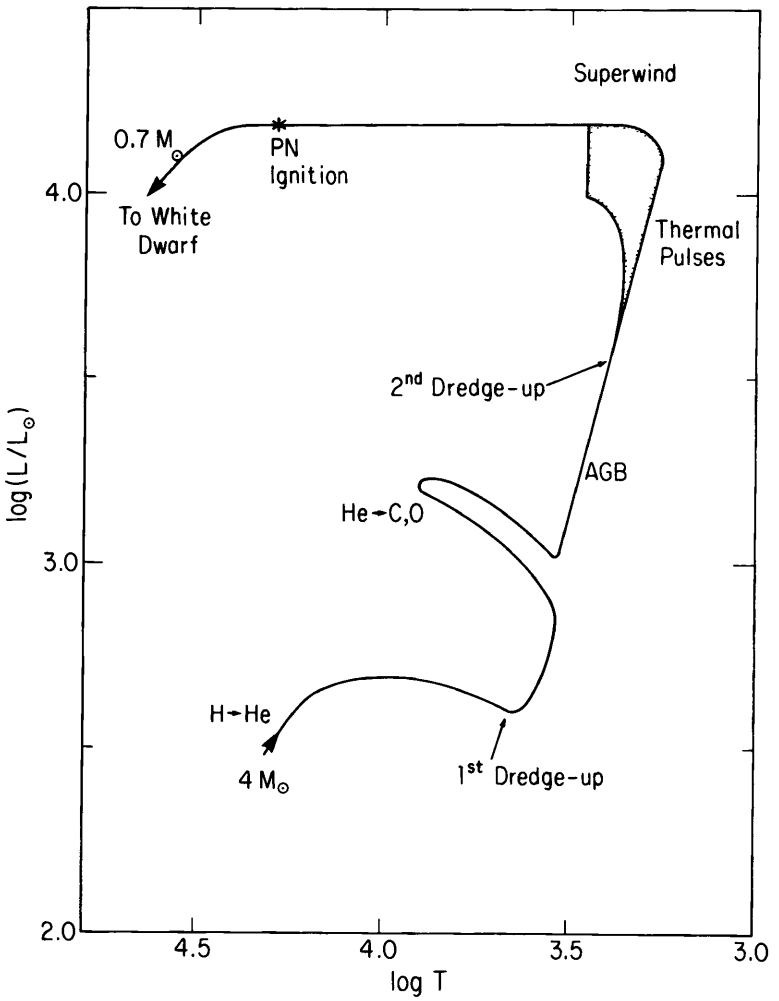


Figure III-14. Post-main-sequence evolution of a $4M_\odot$ star. [After I. Iben, Jr. (1985), *Quart. J. Roy. Astron. Soc.*, **26**, 1.]

processes can be evaluated by observations of circumstellar shells. Figure III-15 is a schematic representation of three observable ratios whose radial gradients are important diagnostics of nucleosynthesis and convective mixing in a $4M_\odot$ AGB star. MMA observations of such gradients, together with observations of the density distribution as a function of distance from the star, will allow us to investigate the mass-loss rate, understand its variation with time, and determine the relative importance of the quiescent wind to the superwind phase on the $\sim 3.3M_\odot$ ejected into the ISM by this evolving $4M_\odot$ star.

In this regard, one of the crucial observations for stellar evolution in recent years was made by Mufson, Lyon, and Marionni,[46] who detected CO emission from the planetary nebula NGC 7027. Since then, circumstellar clouds have been found associated with some twenty planetary nebulae.[47] These observations

[46]Mufson, S. L., Lyon, J., and Marionni, P. A. (1975), *Astrophys. J. (Letters)*, **201**, L85.
[47]Huggins, P. J. and Healy, A. P. (1989), *Astrophys. J.*, **346**, 201.

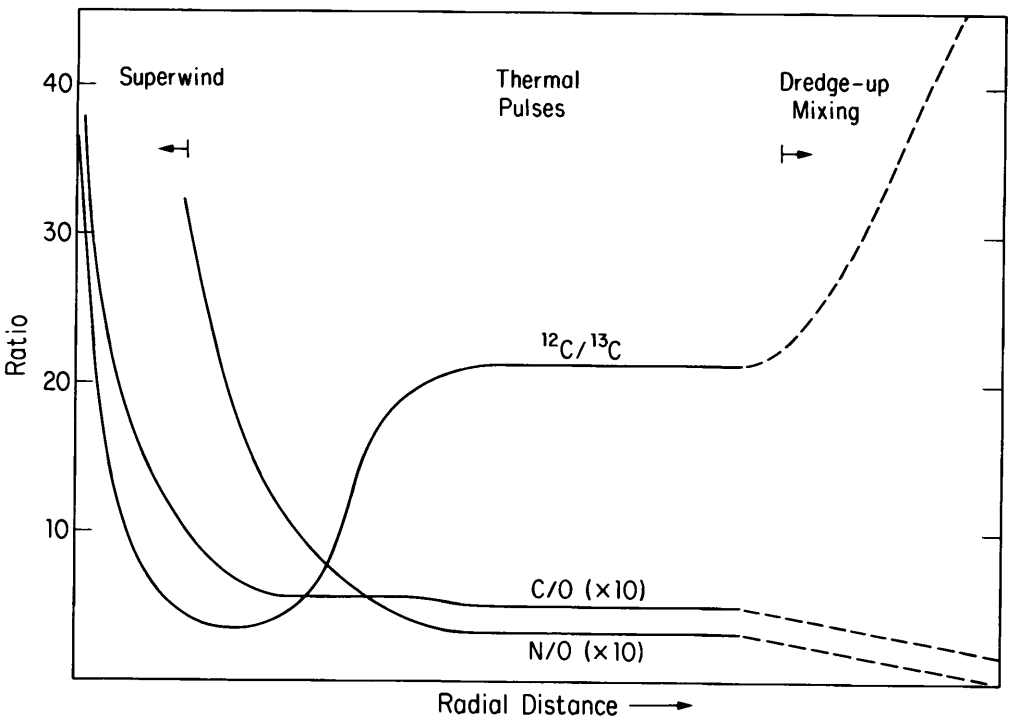


Figure III-15. Schematic representation of the expected chemical and isotopic gradients in the circumstellar shell of an evolved $4M_{\odot}$ AGB star. [After A. Renzini and M. Voli (1981), *Astron. & Astrophys.*, **94**, 175.]

show that it is indeed the mass lost during the AGB phase which is destined to become a planetary nebula as the central star collapses, heats up, and begins to ionize the circumstellar shell. Images of the molecular-line emission at high resolution clearly show the shell-structure of the surrounding molecular cloud (see Fig. III-16).

The high resolution available with the Millimeter Array will allow the detailed study of many such shells. Photochemistry in these environments will be studied; the shell kinematics will be imaged; and the evolutionary histories of the stars during their transitions to the planetary nebula stage will be determined. Finally, the measurement of the shell masses of a large number of planetary nebulae, coupled with an examination of their luminosities and galactic kinematics, should yield a good upper limit for the progenitor mass of white dwarf stars, or, conversely, a lower limit for the mass of supernova progenitors.

The simple picture described above, of a planetary nebula being the post-mass-loss ionization of a circumstellar shell, already presents one glaring inconsistency deriving from the limitations of the instruments used to make the observations. That is, at the spatial resolutions available with single-antennas, most circumstellar envelopes appear to be circularly (hence spherically) symmetric, but the majority of planetary nebulae are not so. This characterization will need to be clarified and re-investigated using the higher resolution available with the MMA; the statistics of non-circular symmetry should prove very

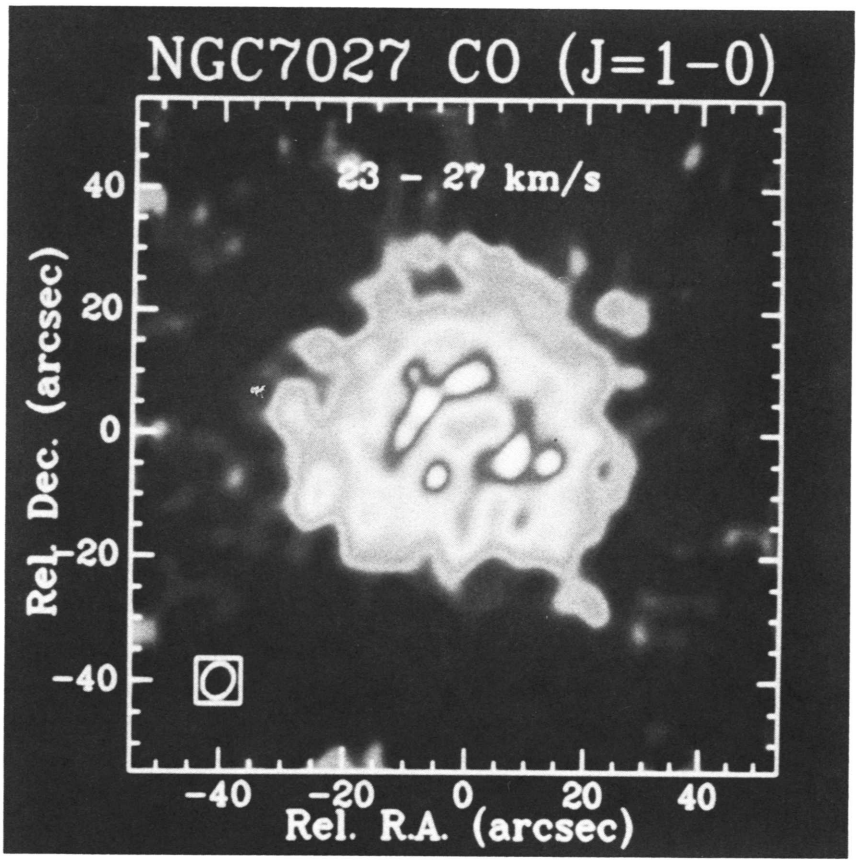


Figure III-16. This intensity map of ^{12}CO (J = 1–0) at 115 GHz toward the planetary nebula NGC 7027 was produced by the Berkeley–Illinois–Maryland Array (BIMA) and the NRAO 12-Meter Telescope. Emission was averaged over a velocity range $v_{\rm LSR}$ = 23–27 km s^{-1} centered on the systemic velocity, at which the gas flow is perpendicular to the line of sight. The molecular gas emission reveals an expanding shell structure of diameter $\lesssim 20''$ exterior to the ionized nebula. The neutral envelope shows considerable structure on scales $\lesssim 5''$ ($\approx$ 0.03 pc). The inset illustrates the 50% contour of the CLEAN beam.

interesting, since they can be compared with expectations from the shaping of the circumstellar envelopes in binary-star systems. The subsequent shaping of the planetary nebula during and after ionization has received much attention of late, and most scenarios invoke a hot fast wind from the central degenerate star, channeled by the surrounding anisotropic gas envelope (see, e.g., the recent review by Kwok[48]). Recent high-sensitivity molecular-line observations of transition objects such as CRL 618 and CRL 2688 have shown the presence of very-high-velocity *molecular* winds, which occur in objects as cool as CRL 2688 in which ionization has not yet begun. The CO (3–2) and CO (2–1) line-profiles

[48]Kwok, S. (1987), in *Late Stages of Stellar Evolution*, S. Kwok and S. Pottasch, Eds., (Reidel: Dordrecht), p. 149.

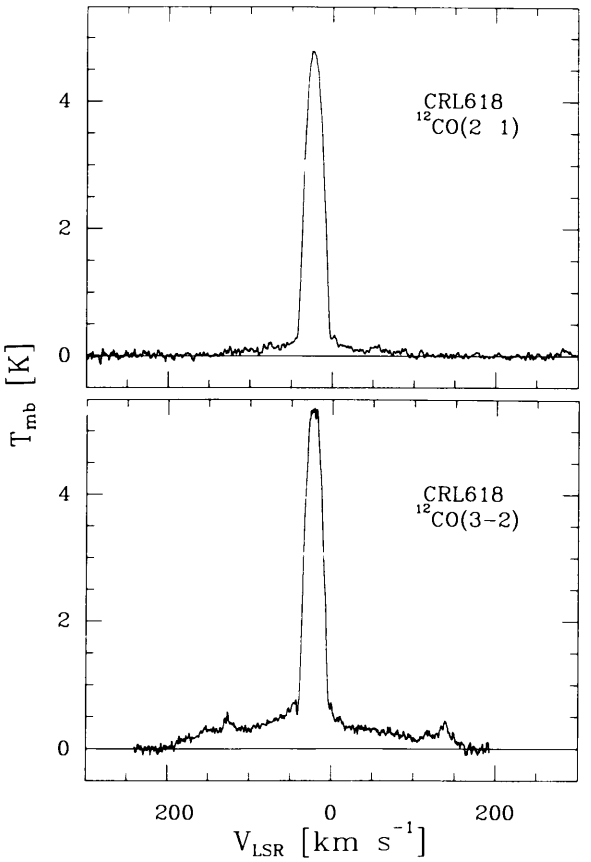


Figure III-17. The very-high-velocity CO wind from the late AGB star CRL 618. The prominence of the wings at CO (J = 3–2) is indicative of both high excitation and a very compact emission region. [From C. F. Gammie *et al.* (1989), *op. cit.*]

of CRL 618[49] are shown in Figure III-17.[50] These fast winds are very difficult to explain within any plausible scenario unless they constitute the final ejections of the stellar envelopes. There is much to be learned from high-spatial-resolution observations of this phenomenon, such as whether the fast winds are spherically symmetric or bipolar, their relations to the structure seen in optical, infrared and H_2-line images, and their chemistry. The studies described in this paragraph offer the possibility of observing in detail the final evolution of a star across the H–R diagram; and this may be a phenomenon which happens fast enough that it can be observed in 'real' (i.e., human) time.

6.2. Grain Formation. Winds from cool evolved stars are probably the dominant source of refractory dust grains in the interstellar medium. The grains manifest themselves through thermal emission extending from the far-infrared through millimeter wavelengths. At around 1-mm wavelength the emission is certainly optically thin, so that high-resolution images of the thermal contin-

[49] Gammie, C. F., Knapp, G. R., Young, K., Phillips, T. G., and Falgarone, E. (1989), *Astrophys. J. (Letters)*, **345**, L87.

[50] *Cf.* also Cernicharo, J., Guélin, M., Martín-Pintado, J., Peñalver, J., and Mauersberger, R. (1989), *Astron. & Astrophys.*, **222**, L1.

uum from such winds will be an excellent tracer of the dust distributions. The high continuum sensitivity of the MMA will make possible direct imaging of the dust-condensation zones for those AGB stars within a few hundred parsecs, at resolutions $< 0\farcs1$. High-frequency performance of the MMA is especially critical here, since the dust emission increases at least like ν^3 and since the angular resolution scales as ν^{-1}. Grain growth is expected to be most rapid at distances of a few times 10^{14} cm, so such observations will require the best-possible resolution.

Accurate dust mass-loss rates are extremely important in assessing the total contribution of cool star winds to the interstellar dust. Although far-infrared observations have been used to derive these rates,[51] continuum observations at longer wavelengths ($\lambda \approx 1$ mm) will place much tighter constraints on the dust emissivity as a function of wavelength. The high sensitivity of the MMA in the continuum should permit studies of a large sample of stars for which far-infrared (i.e., IRAS) flux densities are known, and so should provide an accurate census of the rates of dust return by cool evolved stars.

A complementary approach to the question of grain formation is the study of the gas-phase distribution of refractory molecules (e.g., SiS and SiO), which are incorporated into the solid particles. The distribution of these molecular species is an indicator of grain growth. For example, Bieging and Rieu[52] have used the BIMA millimeter-wave interferometer to determine the SiS abundance distribution in the envelope of the carbon star IRC+10216 and conclude that the SiS abundance (relative to H_2) declines rapidly as the wind reaches $\sim 3\times10^{15}$ cm radius, which implies that $> 98\%$ of the silicon is incorporated into grains within sixty years of ejection from the photosphere. Models of the silicon chemistry, which include adsorption onto grains, indicate that the SiS abundance is sensitive to the rate at which molecules stick to grain surfaces. High-resolution molecular-line studies of cool-star envelopes in abundant refractory species like SiS and SiO could therefore be used to constrain models of grain growth and to determine microscopic parameters such as 'sticking coefficients' for particular species. By observing a wide variety of stars, the dependence of the grain-formation process on stellar composition, wind density (i.e., mass-loss rate), and temperature can be investigated in detail.

6.3. Molecular Processes. The rich molecular content of circumstellar envelopes is an important subject of investigation for the MMA. The flexibility of the instrument in terms of spatial resolution is a major asset to studies of gas-phase chemistry, and we expect that much can be learned about how molecules are formed and destroyed throughout the envelope.

In the inner envelope (typically for radii $< 10^{16}$ cm), where temperatures and densities are high, neutral reactions are expected to be important. This region has previously been largely inaccessible because, in even the nearest stars of interest, the angular sizes are too small to be resolved by existing telescopes.

[51] Sopka, R. J., Hildebrand, R., Jaffe, D. T., Gatley, I., Roellig, T., Werner, M., Jura, M., and Zuckerman, B. (1985), *Astrophys. J.*, **294**, 242.
[52] Bieging, J. H. and Rieu, N.-Q. (1989), *Astrophys. J. (Letters)*, **343**, L25.

The MMA, however, will be able to image the distributions of important species in this region, for typical envelopes, to distances of > 10 kpc. Of special interest are the refractory and inorganic molecules indicative of high-temperature chemistry, such as the metal halides,[53] which have so far been found only in the nearest and brightest objects. Spectral-line imaging of the inner envelopes at high angular resolution may also give essential information on the processes by which the winds are accelerated to terminal velocity.

In contrast to inner-envelope studies, observations of the outer parts (typically $> 10^{16}$ cm) will be well suited to the MMA in its more compact configurations, which maximize the surface-brightness sensitivity. From the standpoint of circumstellar chemistry, the outer envelope is the region where the ultraviolet optical depth to the exterior is low enough that the interstellar radiation field can penetrate the envelope and initiate a complex sequence of ion–molecule reactions. The importance of photochemistry has been demonstrated by detections of a host of organic species, whose abundances cannot be explained by equilibrium chemistry in the stellar photosphere, but must be created by ion–molecule reactions occurring at large distances from the star.[54] Imaging of the envelope of IRC+10216 has shown that relatively complex species like HC_3N are distributed in shells which extend to radii of several times 10^{16} cm and are clearly products of photochemical reactions.[55] Figure III-18 shows images made with the BIMA millimeter-wavelength synthesis array of three different chemical species—SiS, HC_3N, and C_3N—and illustrates the radial stratification which results from various chemical processes in the envelope. The MMA will permit the extension of such studies to a much larger number of chemical species, and to many more stars. Such work would test critically the photochemical theories by showing the effects of varying stellar chemical composition (e.g., C/O ratio), wind density, dust opacity, etc.

An equally important scientific result from MMA studies of circumstellar photochemistry is likely to be improved values for quantities which are very difficult to measure in the laboratory, such as the photodissociation rates of certain molecules and the rates for molecular reactions involved in the formation or destruction of observable species. In this respect, circumstellar envelopes may prove to be an excellent chemical laboratory for studying gas-phase ion–molecule chemistry. Since many of the same reactions are important in interstellar clouds, the chemical analysis of these envelopes should have a wider applicability to the interstellar medium.

6.4. Spectral-Line Formation and Radiation Transfer. The MMA will provide detailed information on critical questions related to processes affecting the observed molecular lines. The major theoretical areas include molecular excitation mechanisms and radiative transfer in stellar envelopes. It has long been recognized that excitation through the vibrational (and possibly electronic) transitions probably dominates the formation of the millimeter-wavelength ro-

[53] Cernicharo, J. and Guélin, M. (1987), *Astron. & Astrophys.*, **183**, L10.
[54] See, e.g., Olofsson, H. (1987), in *Late Stages of Stellar Evolution, op. cit.*, p. 149.
[55] Bieging, J. H. and Rieu, N.-Q. (1988), *Astrophys. J. (Letters)*, **329**, L107.

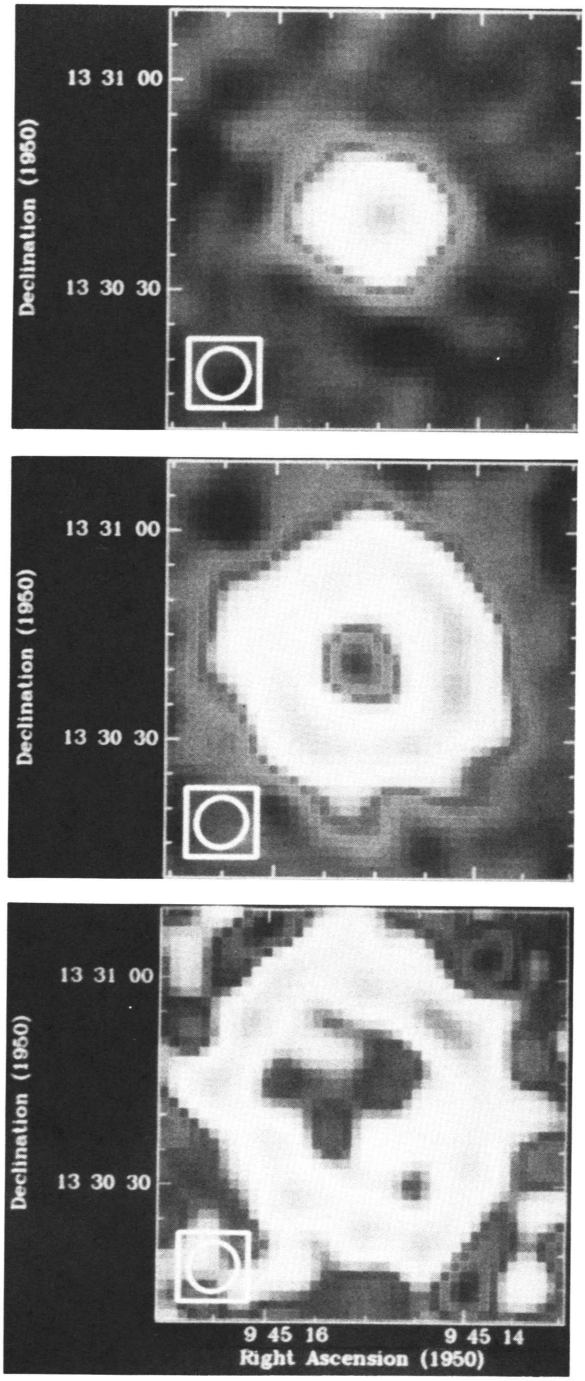


Figure III-18. The distribution of SiS *(top)*, HC_3N *(middle)*, and C_3N *(bottom)* in the circumstellar gas surrounding the AGB carbon star IRC+10216. These three images, each of 7″ resolution, were produced simultaneously by the BIMA telescope. SiS is a freeze-out species which condenses early in the outflow near the star; HC_3N is a product of ion–molecule chemistry in the cooler regions of the outflow; and C_3N is produced by photodissociation of HC_3N by interstellar UV radiation penetrating the shell from the exterior. [From J. H. Bieging (1990), in preparation.]

tational lines of many observed species.[56,57] Since essentially all of the astrophysical information discussed above depends on understanding the molecular excitation, this issue is critical to effective use of the MMA. Here again, the sensitivity and frequency agility of the MMA will provide powerful diagnostics for understanding molecular excitation. These diagnostics include studies of rotational lines of molecules in excited vibrational states, which are already detectable in some cases.[58,59,60,61] The MMA will provide resolved images of such vibrationally excited lines, which will critically test the theory of infrared excitation.

An additional important test may be the polarization of rotational lines in the ground vibrational state. Theoretical models[62,63] predict that for well-resolved images, linear polarizations of several percent will be present in the rotational lines of diatomic molecules as a consequence of the anisotropic infrared radiation field from the central star. It is therefore important that the MMA have a capability for polarization mapping in molecular lines with velocity resolution of ~ 1 km s^{-1}.

Some abundant species that are readily thermalized (especially CO) will be coupled to the gas kinetic temperature rather than to the infrared radiation field. Previous efforts to model the temperatures of circumstellar envelopes have not accurately predicted the temperatures in the outer envelope. Additional heat sources, such as photoelectric emission from dust grains in the outer envelope, where the interstellar UV photons can penetrate, may be an important energy input to the gas. High-resolution images from the MMA of collisionally excited molecules will be of importance in understanding the thermal balance of circumstellar envelopes.

Finally, early work on modeling circumstellar envelopes was (and continues to be) based on assumptions of spherical symmetry and r^{-2} density gradients. These assumptions greatly simplify the treatment of radiative transfer, yet we already know from recent millimeter-wavelength aperture-synthesis and single-antenna mapping that circumstellar envelopes are often highly non-spherical and have multiple velocity components, which implies deviations from r^{-2} density laws. A 'gallery' of accurate high-resolution MMA images of molecular lines in many envelopes will be a major stimulus to improving the theoretical treatment of radiative transfer.

6.5. The Astrophysics of Stellar Masers. Astrophysical masers are a valuable, but sometimes difficult to interpret, probe of the structure and kinematics

[56]Morris, M. (1975), *Astrophys. J.*, **197**, 603.
[57]Deguchi, S. and Uyemura, M. (1984), *Astrophys. J.*, **285**, 153.
[58]Ziurys, L. M. and Turner, B. E. (1986), *Astrophys. J. (Letters)*, **300**, L19.
[59]Turner, B. E. (1987), *Astron. & Astrophys.*, *183*, L23.
[60]Guélin, M., Cernicharo, J., Navarro, S., Woodward, D. R., Gottlieb, C. A., and Thaddeus, P. (1987), *Astron. & Astrophys.*, **182**, L37.
[61]Yamamoto, S., Saito, S., Guélin, M., Cernicharo, J., Suzuki, H., and Ohishi, M. (1987), *Astron. & Astrophys.*, **323**, L149
[62]Deguchi, S. and Watson, W. D. (1984), *Astrophys. J.*, **285**, 126.
[63]Goldreich, P. and Kylafis, N. D. (1982), *Astrophys. J.*, **253**, 606.

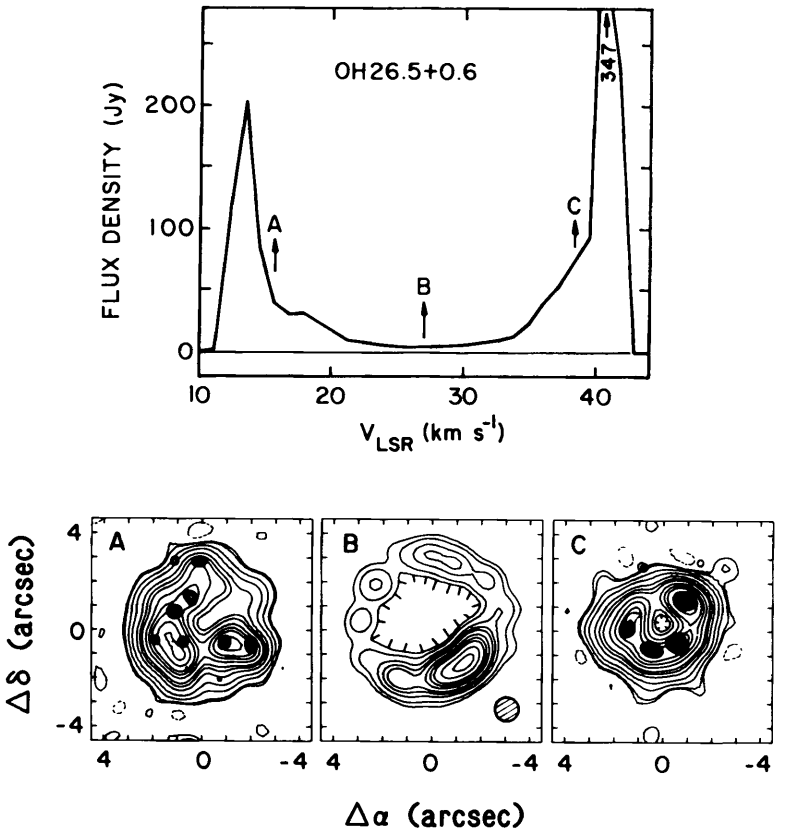


Figure III-19. VLA images of 1612-MHz maser emission from the oxygen-rich AGB star associated with the OH 127.8−0.0 circumstellar shell. [From P. F. Bowers and K. J. Johnston (1990), *Astrophys. J.*, in press.]

of the high-density regions in circumstellar envelopes. A complex spectrum of very-high-surface-brightness velocity components is often observed in multiple transitions of OH, H_2O, and SiO. Each velocity component usually originates from a single spot which is separated from others in a larger masing region. The SiO maser is observed only in the cool, extended envelopes of late-type stars and Orion. These late-type stars, Mira variables, and even more extreme oxygen-rich types, are a significant Galactic population and lose mass at very high rates. They are a primary source of processed material returned to the ISM. Indeed, astrophysical masers may be a statistically significant indicator of stellar mass loss. The sizes of maser regions correlate with several independent indicators of mass loss.[64] Currently this correlation is best for extreme OH/IR stars, which, although important, are only a small part of the larger picture. Figure III-19 shows a map of the OH-maser emission at 1612 MHz from the OH/IR star OH26.5+0.6. The brightnesses of SiO masers in the 'normal' population of Mira variables make it possible to extend this correlation to a sample of objects which is vastly larger and more significant on a galactic scale.

[64]Bowers, P. F. (1989), in *Proceedings of the Sixth Cambridge Workshop on Cool Stars, Stellar Systems and the Sun*, G. Wallerstein, Ed., in press.

Maser excitation is not well understood, but the existence of OH, H_2O, and SiO masers in the same stars hints at a common mechanism. Infrared pumping via similar vibrational/rotational transitions in the intense part of the stellar continuum may be the common factor. SiO, with its large number of observable transitions, may be an ideal probe. The expectation is that by observing both the distribution and the variation of maser structure in several transitions, the base of the mass-loss region and the driving mechanism for the outflow may be studied. A large number of MMA images, of a meaningful sample of stars and transitions, will be needed in order to provide definitive answers.

7. Planetary Science

The Millimeter Array is an instrument which is sorely needed for competitive research in planetary science. The most valuable observations the instrument will offer us, which cannot be done at any other existing or proposed telescope, are accurate (better than 1% accuracy) center-to-limb observations of planets, and observations of large objects ($> 30''$) as well as small ($\sim 1''$) objects, such as satellites and asteroids. The unique capability of fast imaging is another feature of particular importance to planetary work, since planets vary on relatively short time scales, because of either rotation, revolution around the Sun, or wind effects.

Some of the most exciting projects made possible by the MMA will be to:

- Image the varying atmospheric winds and thermal profiles on Mars and Venus, using CO and its isotopes.
- Resolve phosphine emission in the Great Red Spot, HCN on Titan, and volcanic emission on Io.
- Measure properties of subsurface layers of hundreds of asteroids.
- Image the thermal emission from Pluto and Charon.
- Expand cometary research by providing unobscured images of comet nuclei: study subsurface layers and atmospheric structure; and facilitate precision astrometry.
- Image impulsive events and sporadic molecular emission from comets.
- Detect 'proto-Jupiters' in nearby stellar systems and determine their masses and chemical compositions.

7.1. Planetary Atmospheres. Observations at millimeter wavelengths allow one to derive abundances of gases as functions of altitude, together with the temperature vs. pressure profiles of the atmospheres. The observations are needed for lines of different species, and preferably also for different transitions. At different transitions one typically probes different altitudes in the atmosphere. Hence, one can derive an altitude distribution for the gas's abundance. The line intensity is a measure of the convolved abundance and temperature profiles in the atmosphere. The shape of the line depends upon the pressure and abundance at the relevant altitude levels. A few examples illustrate these points.

7.1.1. CO on Venus and Mars. Mapping studies of CO in the atmospheres of Venus and Mars offer the possibility to study the diurnal, latitudinal, and

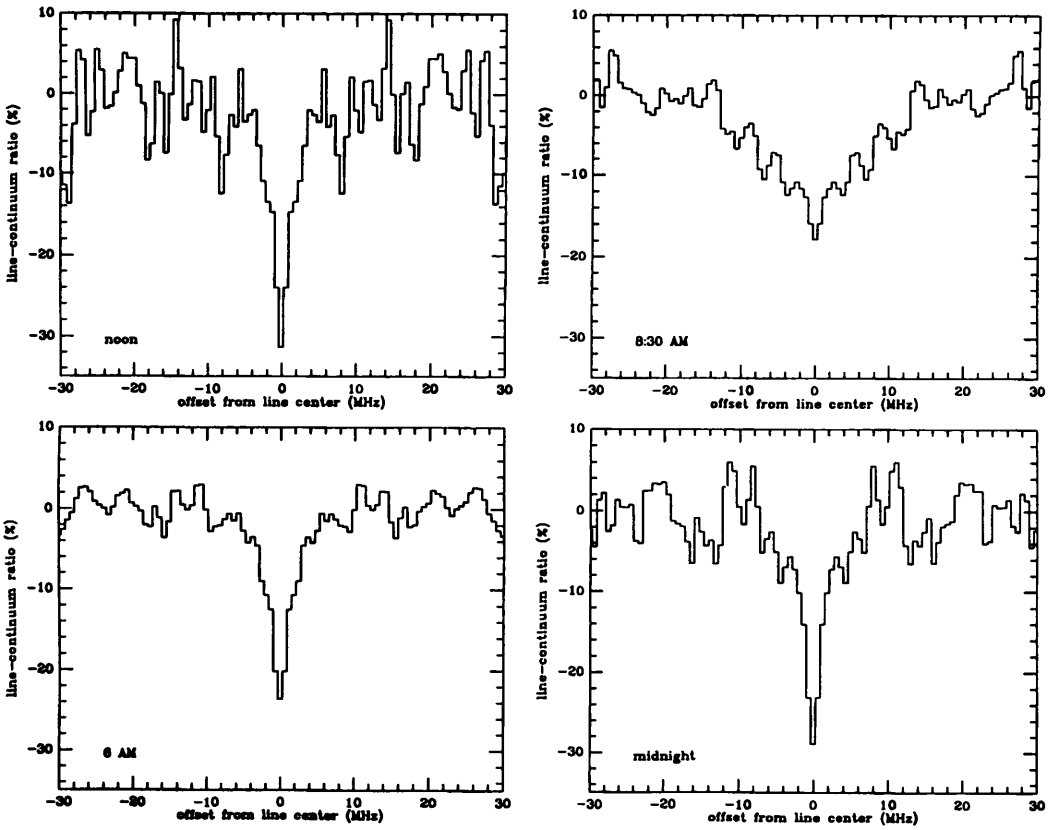


Figure III-20. CO (J = 1–0) profiles at locations marked by the crosses in Fig. III-21. [From I. de Pater (1991), *Ann. Rev. Astron. Astrophys.*, in preparation.]

seasonal variations of the atmospheric temperature and CO abundance in a direct way. In the upper atmosphere of Venus, strong diurnal CO-abundance gradients exist. Maps of the J = 1–0, 2–1 and 3–2 lines may be inverted to yield the longitudinal and latitudinal CO abundance and temperature variations in the atmosphere, at altitudes of 80–120 km. Such studies provide the fundamental input to models of the general wind circulation and photo-chemistry of this region in the atmosphere. Direct measurement of winds would substantially improve wind-circulation theories. Such measurements are possible by observing the Doppler shifts of lines near the limb of the planet. These shifts can be measured at the limb if wind velocities are $\gtrsim 100$ m s^{-1}.

The CO abundance in the atmosphere of Mars is not strongly altitude dependent, and, therefore, one can use the CO lines at millimeter wavelengths to sound the atmosphere for its temperature profile. Maps of the planet Mars can then be used to produce global maps of the Martian temperature profile, in order to study latitudinal and seasonal variations as well as the effects of large-scale meteorological phenomena (e.g., global dust storms) on these profiles.

Both BIMA and OVRO have mapped the J = 1–0 CO transition on Venus. The lines vary across the disk in a manner consistent with single-antenna observations (see Fig. III-20). However, it takes $\sim$ 2 weeks to do the job using

a three-element interferometer. The line varies from grid cell to grid cell—and undoubtedly varies somewhat in time. The MMA is necessary for the fast, high-resolution imaging that is warranted for detailed studies of Venus's atmosphere. In addition, one would like to image at different CO transitions simultaneously, to unravel both temperature structure and CO abundance with altitude, at all positions on the disk.

Winds can be measured from the Doppler shifts of the lines. BIMA and OVRO made the first attempts at this as well. One expects day-to-night winds in the mesosphere and above ($z \gtrsim 100$ km), and retrograde winds in the cloud layers ($z \approx 60$ km). CO lines can be used to determine wind velocities at different altitudes and to trace the transition between these two regimes. The existing interferometric data imply ~ 100 m s^{-1} winds, flowing from the day- to the night-side.

The continuum image at 3 mm (see Fig. III-21) shows a strong day/night asymmetry. Is this a static feature—does it vary in time? Fast imaging, possible with the MMA, will provide the answer.

To be able to image fast and follow cloud features across the disk is important and feasible for Venus and the giant planets. Note that the features are of low contrast against a high-surface-brightness continuum.

7.1.2. Trace Constituents. Besides merely establishing the presence of trace elements, it is important to define their altitude distributions in order to define their places of origin—e.g., in the deep atmosphere, brought up by convection; or influx from outside the atmosphere (e.g., oxygen from rings or satellites); or high up in the atmosphere produced, for example, by photolysis by lightning.

In particular, center–limb observations of such gases will contain clues as to the levels in the atmosphere at which they exist. For the case of a typical temperature vs. pressure profile as shown in Figure III-22, the gas that is distributed near and above the troposphere (so in the stratosphere) may appear in absorption at the center of the planet, while appearing in emission at the limb. So center–limb observations may seem limb brightened in the line center, while they seem limb darkened outside the line. Examples of the important molecules are (a) PH_3, H_2S, CO, and HCN, on the giant planets; (b) sulfur molecules (SO_2 and SO) and ozone, in Earth-like atmospheres (Venus and Mars); (c) nitriles and hydrocarbons like HCN, HC_3N, and organic molecules, on Titan; and (d) SO_2 and SO on Io.

7.2. Solid Surfaces. At different wavelengths one probes to different depths in the crust—roughly anywhere from a depth of several wavelengths, to several hundreds of wavelengths (for pure water-ice). The measured brightness temperatures depend upon the physical temperatures of the subsurface layers and on the radiative transport of the thermal emission outward. The physical temperatures depend upon the solar insolation and the thermal conductivity and inertia of the material. Radiative transfer outward is limited by the emissivity and the absorption/scattering characteristics of the material, both of which are highly dependent upon wavelength. Thus, probing the surface layers at different wavelengths allows mapping of the temperature distribution with surface

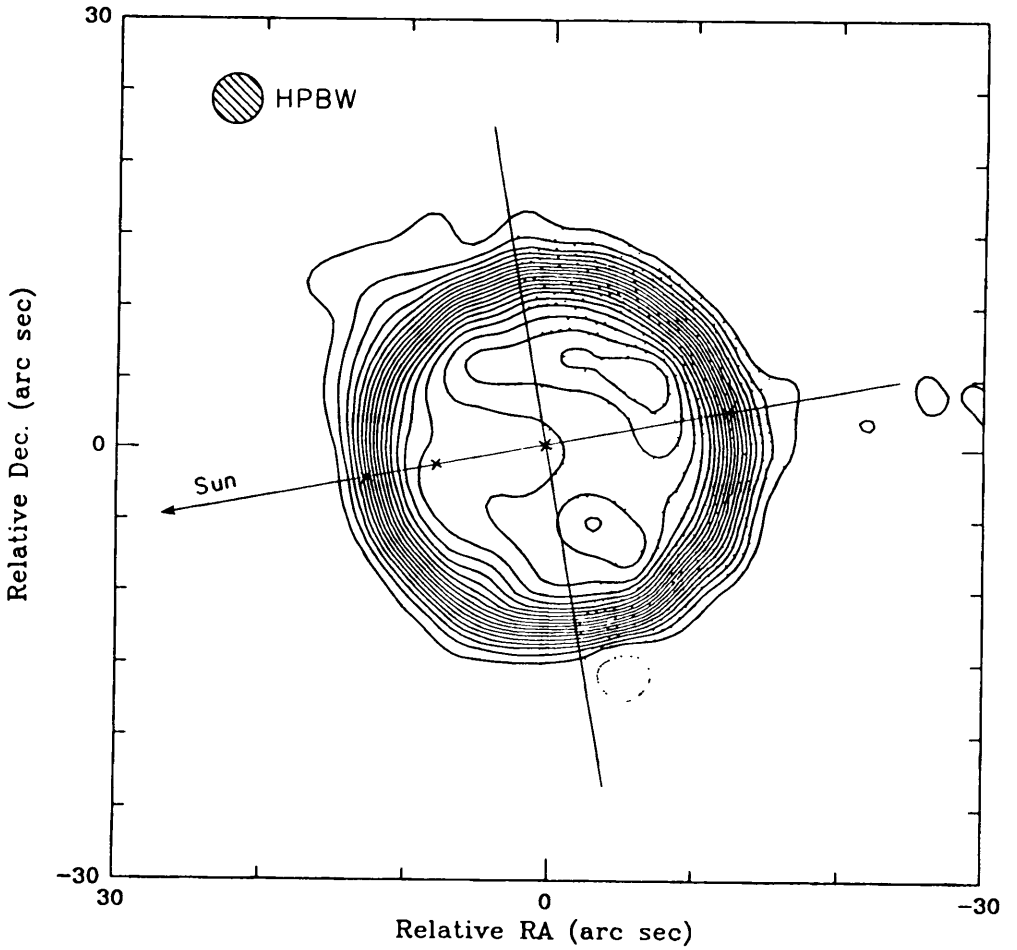


Figure III-21. Venus at 3 mm. Small crosses mark the locations of the CO profiles shown in Fig. III-20. [From I. de Pater (1991), *op. cit.*]

depth. This, together with high-resolution images in intensity, as well as in polarization, will ultimately allow one to deduce the substance and composition of the material (e.g., solid rock, loose dust, gravel; ice in clumps, wet ice; etc.). The importance of the millimeter-wavelength observations is that one typically probes in the region of the diurnal solar heat wave—this gives the information necessary to derive the thermal characteristics (conductivity, thermal inertia, and dielectric constant) of the material, properties unique for each substance.

The following are a few specific examples.

7.2.1. Satellites and Asteroids. The Galilean satellites each have their own unique properties. We have only to recall pizza-like Io with its volcanoes and hot spots, and the Voyager photographs of each satellite. Spectra of the satellites at infrared to centimeter wavelengths give the information on the subsurface layers of the bodies—information which cannot be derived in any other way except by drilling a hole in their 'ground'. In particular, the spectra of Europa and Ganymede are exceptional when compared to that of the moon. Brightness

IV. IMAGING

1. Introduction to Imaging

One of the prime roles of the Millimeter Array will be imaging of objects on size scales ranging from many arcminutes down to one tenth of an arcsecond. The highest resolution will come from the largest configuration, with baselines of some few kilometers, using now-standard imaging techniques that have been developed over the past ten to twenty years.[1] Objects will be observed for some period of time, perhaps up to four hours, and images constructed via Fourier synthesis, followed if necessary by deconvolution and self-calibration to correct for defects caused by limited sampling of the Fourier plane and fluctuations in the seeing. The lowest resolution will come from a 'mosaicing' mode in which the MMA operates somewhat like a conventional single-antenna facility: the MMA in its most compact configuration, with baselines up to about 70 meters, is pointed successively to a grid of points spanning the region to be imaged. This mosaicing mode is dictated by the small size of the primary beam at millimeter wavelengths. The algorithms for constructing an image from the data thus collected were developed just five years ago but have been extensively tested and used since then. For weak objects, an image is formed by linear combination of a number of Fourier-synthesized images, while for brighter objects, on which limited Fourier-plane sampling may be important, a multi-pointing deconvolution algorithm must be used, incorporating perhaps self-calibration. For both modes of imaging, the excellent Fourier-plane coverage generated by the forty elements of the MMA makes deconvolution substantially less important than it is for existing arrays such as the VLA and VLBA. For example, in the compact configuration the coverage will be essentially complete for many observations, and deconvolution will be reserved for cases where the image-plane signal-to-noise ratio exceeds about 100 : 1. This complete coverage is also vital for cases where deconvolution is not appropriate, such as the study of structure in molecular clouds, and for time-critical applications, such as imaging solar flares or imaging during periods of exceptionally low atmospheric opacity. Finally, good Fourier-plane coverage allows high-fidelity imaging of objects such as planets, which usually are troublesome targets for interferometric imaging.[2]

In this chapter we discuss these two forms of imaging: high-resolution, 'conventional' imaging, which is relatively well understood; and the mosaicing mode, which is treated in more detail. The latter gives rise to a number of questions: How are short spacings to be measured? How robust is the mosaicing mode? How important are various instrumental effects such as pointing errors? We have conducted an extensive program of observational tests and simulations to answer these and other questions about mosaicing. We will discuss the general results and show specific examples of mosaicing.

---

[1] *Cf.* Thompson, A. R., Moran, J. M., and Swenson, G. W., Jr., *Interferometry and Synthesis in Radio Astronomy*, Wiley, New York, 1986.

[2] de Pater, I. (1991), *Ann. Rev. Astron. Astrophys.*, in preparation.

2. The Array Configurations

The MMA will have four principal configurations: three large configurations having east–west diameters of about 3 km, 900 m, and 250 m, and a maximally compact configuration with maximum east–west diameter of $\sim$ 70 m. These will be stretched north–south to obtain good Fourier-plane coverage over a range of declinations. We shall call these the A-, B-, C-, and D-configurations. Since the larger configurations must be tailored to the site geography, we have deferred detailed investigations of the configurations, though we have demonstrated the feasibility of good large configurations even on irregular surfaces such as the Magdalena Mountains site (see Fig. IV-1). The form of the compact, filled D-configuration will be independent of site selection. Since the key goal is to maximize surface-brightness sensitivity while avoiding excessive shadowing, the elements are packed with a filling factor of about 50% (see Fig. IV-2).

3. High-Resolution Imaging

The prime distinction between high- and low-resolution imaging is that the former will principally be of a single primary beam of the array antennas, while the latter will be of many such primary beam areas. High-resolution imaging is therefore identical in principle and very similar in practice to imaging with centimeter-wavelength arrays such as the VLA and VLBA.[3] The array measures the visibility function corresponding to the region of sky in the primary beam pattern of the elements. This can be converted into an image of the region by inverse Fourier transformation. The 'dirty' image thus constructed is affected by two sources of error.

First, the incomplete sampling of the Fourier plane leads to scattering of power in the 'dirty' image into sidelobes. For very weak objects, and typically for single spectral channels, this is insignificant compared to the noise level, but for bright objects it may limit the dynamic range and fidelity of the image. Deconvolution algorithms such as 'CLEAN'[4] and the Maximum Entropy Method (MEM)[5] allow correction for the power scattered into sidelobes, at the cost of some assumptions about the source structure. These assumptions, being compactness for CLEAN and smoothness and positivity for MEM, are quite acceptable for the radio sky as seen at centimeter wavelengths, but may be somewhat less so for the complex emission expected at millimeter wavelengths in, for example, imaging of molecular clouds. However, the large number of array elements, 40 compared to 27 for the VLA and 10 for the VLBA, makes the Fourier-plane coverage of the MMA far superior to that of any major centimeter-wavelength array. This, in turn, means that for the MMA, deconvolution is less important than for centimeter-wavelength arrays.

[3]For a complete account of single-pointing imaging see, for example, the lectures in Perley, R. A., Schwab, F. R., and Bridle, A. H, Eds., *Synthesis Imaging in Radio Astronomy*, A.S.P. Conference Series Volume 6, Astronomical Society of the Pacific, San Francisco, 1989.

[4]Högbom, J. A. (1974), "Aperture synthesis with a non-regular distribution of interferometer baselines", *Astron. & Astrophys. Suppl. Ser.*, **15**, 417.

[5]E.g., Narayan, R. and Nityananda, R. (1986), "Maximum entropy image restoration in astronomy", *Ann. Rev. Astron. Astrophys.*, **24**, 127.

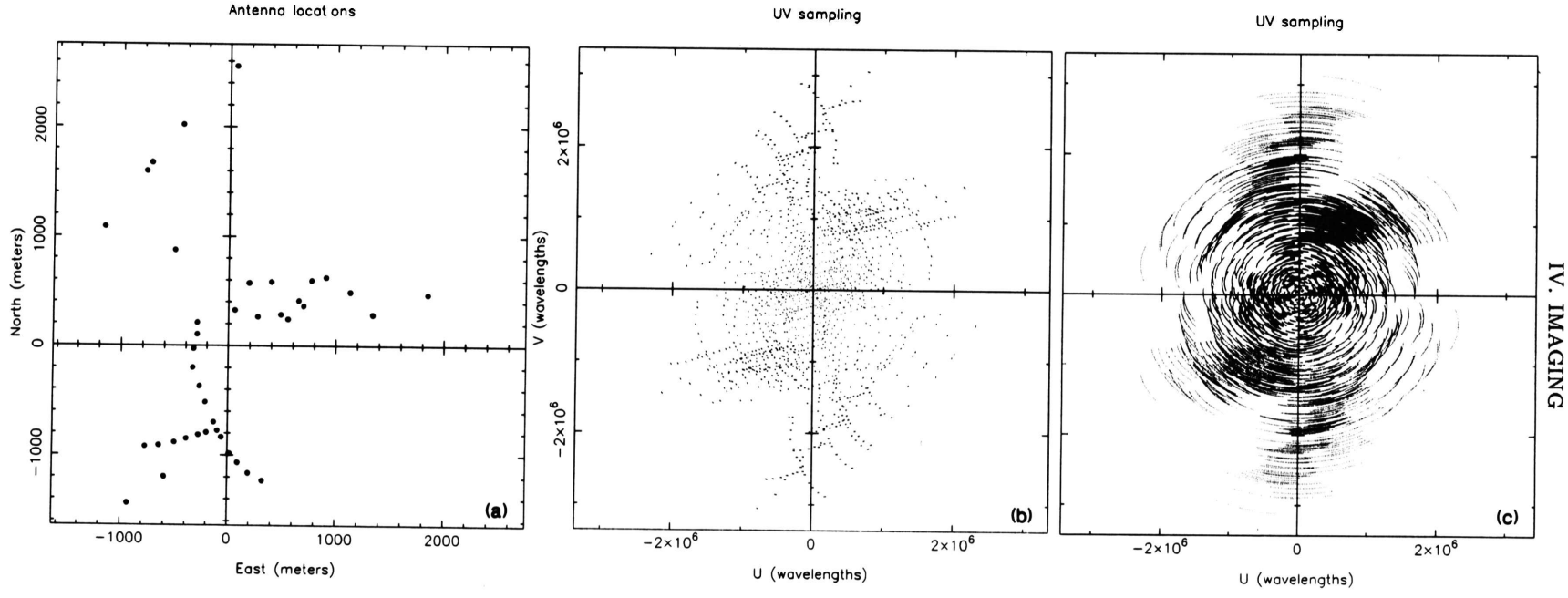


Figure IV-1. The A-configuration adapted to the uneven topography of the Magdalena Mountains site: *(a)* element locations; *(b)* Fourier-plane coverage for a snapshot observation; and *(c)* Fourier-plane coverage for a two-hour integration at declination 30°.

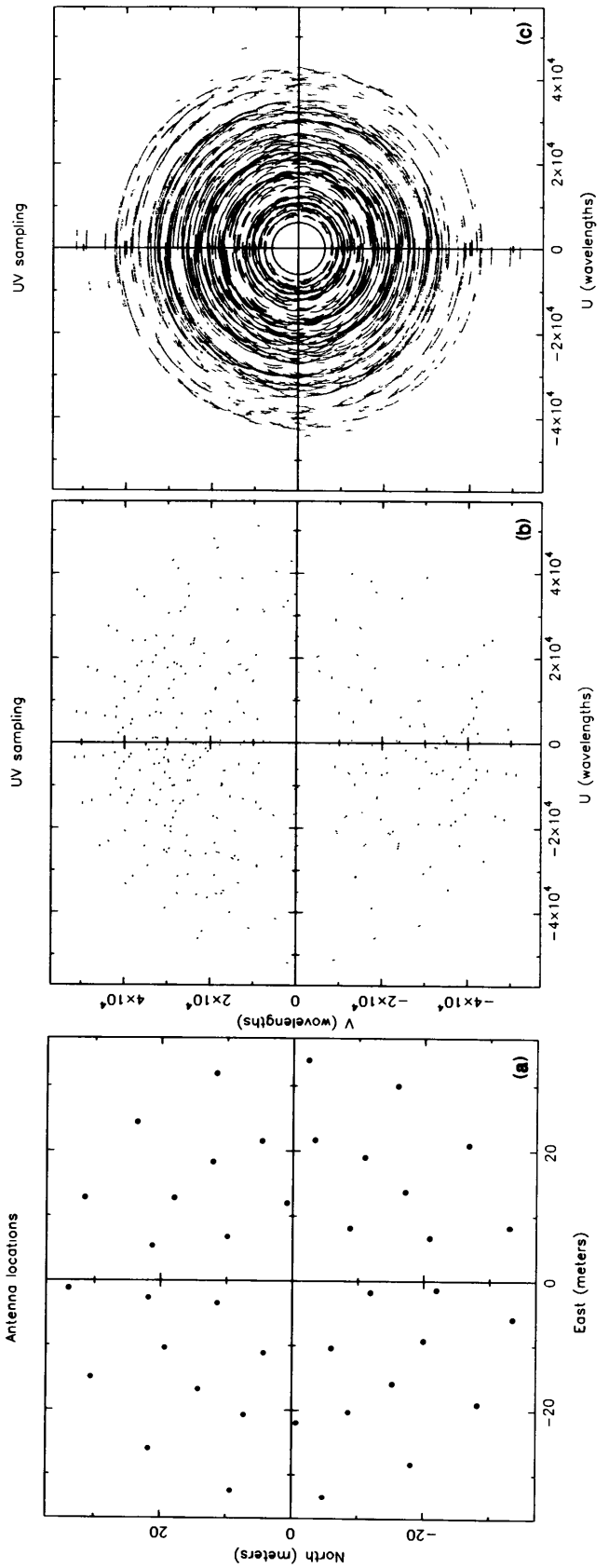


Figure IV-2. The D-configuration: *(a)* element locations; *(b)* snapshot coverage; and *(c)* two-hour integration at declination 30°.

The second source of error relates to signal propagation through the troposphere. The atmospheric opacity and refractivity both fluctuate on time scales ranging from seconds to days. Hence the measured visibility function may contain errors in both amplitude and phase due to atmospheric propagation. Since the atmosphere is essentially non-dispersive at millimeter wavelengths, the phase error increases linearly with frequency whereas the opacity shows a more complicated behavior with frequency (see Chapter II), the general trend of which is to be less favorable at high frequencies. A high site is essential to minimizing these fluctuations, as is rapid calibration of the array, but, even so, some calibration errors will remain. These will limit the dynamic ranges of deconvolved images. However, since the errors are tied to array elements, self-calibration[6] can be used as a countermeasure. The basic principle of self-calibration is that the array calibration constitutes a set of free parameters in the imaging.[7] This extra freedom allowed can be balanced in one of two ways: either via assumptions about the object structure similar to those used in deconvolution, or by forcing redundancy of measurement of the visibility function. The latter is simpler computationally and is model independent, whereas the former does not require any special geometry for the array, an important consideration in the larger, possibly mountain-based configurations of the MMA where self-calibration is most important. Either form of self-calibration requires that the object be sufficiently bright that the signal-to-noise ratio on most baselines is greater than unity. Since this may be too stringent an assumption in some cases, it will be possible to use the ability of the correlator to operate simultaneously at a number of widely separated frequencies in order to (self-)calibrate on a strong part of the spectrum and transfer the derived calibration parameters to a weaker part. A further optimization for spectral-line observations would use a frequency-dependent model rather than simply a model of the strongest line as is current practice for objects with weak continuum. Finally, at least in the smaller configurations, solving for the phase of a physically clustered set of antennas would result in a considerable improvement in signal-to-noise ratio. In the smallest configuration, solving for a simple wavefront tilt will often suffice.

There are a few factors which make MMA high-resolution imaging much simpler and easier than VLA imaging. First, the Fourier-plane coverage will be more extensive and more uniform, improving the quality of the 'dirty' images. Second, the computational image sizes are smaller. Third, wide-field problems such as non-isoplanatism and non-coplanar baselines, which plague the VLA at low frequencies, will be negligible.

4. Low-Resolution Imaging

4.1. Method. For low-resolution imaging, the limited field of view of the array elements becomes restrictive. 'Mosaicing' large fields of view is now common practice at the VLA and will be widely used for the MMA. In mosaicing, the fundamental observable is not simply the visibility function of the sky as seen

[6]See, e.g., Pearson, T. J. and Readhead, A. C. S. (1984), "Image-formation by self-calibration in radio astronomy", *Ann. Rev. Astron. Astrophys.*, **22**, 97.
[7]See, e.g., Lecture 9 in Perley, Schwab, and Bridle (1989), *op. cit.*

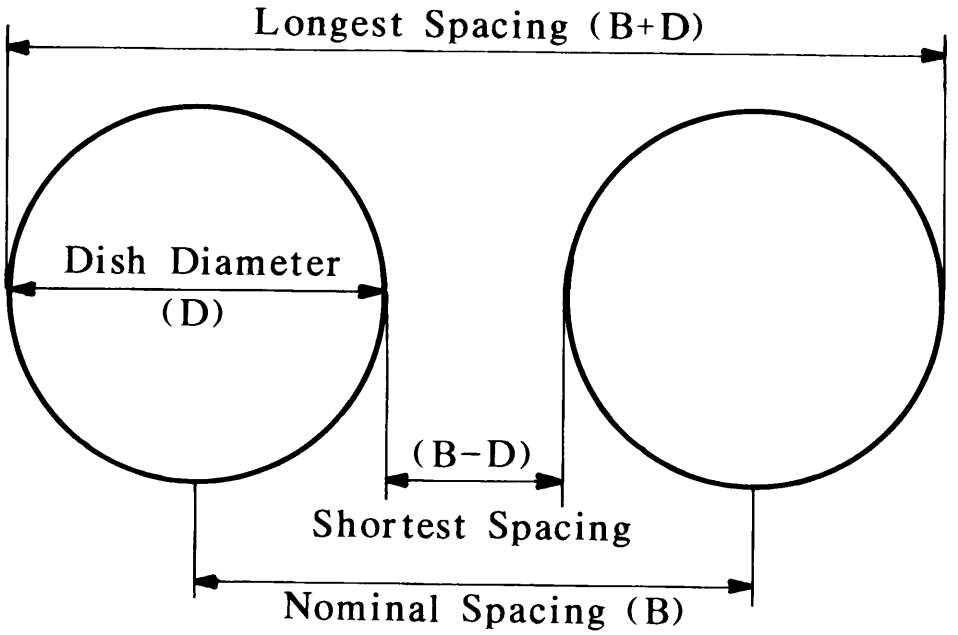


Figure IV-3. A pair of array elements, each of diameter D meters, with a nominal baseline of B meters, actually measure spacings ranging from $B - D$ to $B + D$. Either the Ekers–Rots scheme [Ekers, R. D. and Rots, A. H. (1979), "Short spacing synthesis from a primary beam scanned interferometer", in *Image Formation from Coherence Functions in Astronomy*, C. van Schooneveld, Ed. (Reidel: Dordrecht), p. 61] or mosaicing will recover these spacings.

by an antenna, but rather the visibility for a given pointing position of the array elements. By varying the pointing position to span a given region, and by sampling well the Fourier plane for each pointing, it is possible to obtain essentially all the information required to image the region.[8] We illustrate this point using Figure IV-3, which shows schematically the Fourier components measurable using a two-element interferometer. Although the nominal spacing is B meters, the finite diameter D of the array elements means that all spacings ranging from the minimum $B - D$ up to the maximum $B + D$ are present in a single measurement of the visibility. In fact, the measured visibility function is just a linear combination of all these Fourier components. To untangle these various Fourier components, we simply point the interferometer to a number of places on the sky, separated by $\lambda/(2D)$, and Fourier-transform. Furthermore, Fourier transformation of the total power measured by the interferometer elements will yield the Fourier components on spacings up to D meters. An illuminating special case is that of square antennas with touching edges: clearly all Fourier components from zero up to $B+D$ meters are present, as they must be since this is just another description of a 'single dish'. For real, circular antennas, arranged in a

[8]Cornwell, T. J. (1988), "Radio-interferometric imaging of very large objects", *Astron. & Astrophys.*, **202**, 316; also Lecture 15 in Perley, Schwab, and Bridle (1989), *op. cit.*

packed two-dimensional configuration, all spacings will be present, though with a somewhat lumpy distribution of sensitivity. By taking care with the design of the configuration it is possible to obtain all spacings up to the maximum baseline measured by the array. We can therefore see that, in this description of imaging wide fields of view, the distinctions between 'single dishes' and interferometric arrays become blurred. The packed configuration is essentially a big, multi-fielding 'single dish' with an efficiency lowered by the packing factor but with concomitantly slightly higher resolution. Hence we can measure all the required Fourier components from interferometric and total-power measurements using just the array elements. No large single antenna is required to sample spacings up to B meters. We have therefore adopted this 'homogeneous array' concept for the MMA. The array elements themselves will be outfitted with total-power measuring systems, which will allow beam switching or chopping for continuum observations.

Let us make this more concrete: the visibility function sampled at pointing positions $\mathbf{x}_p$ is the Fourier transform of the true sky brightness $I(\mathbf{x})$ weighted by the primary beam A,[9]

$$V(\mathbf{u}, \mathbf{x}_p) = \int A(\mathbf{x} - \mathbf{x}_p)\, I(\mathbf{x})\, e^{2\pi i \mathbf{u}\cdot\mathbf{x}}\, d\mathbf{x}\,. \tag{1}$$

For a 'single dish', $\mathbf{u}$ is zero and then,

$$V(0, \mathbf{x}_p) = \int A(\mathbf{x} - \mathbf{x}_p)\, I(\mathbf{x})\, d\mathbf{x}\,. \tag{2}$$

In a typical mosaicing observation, one will use the whole MMA, pointing to each position in a grid spanning the object and separated by $\lambda/(2D)$, dwelling on each for some minutes of time. An analog of the usual 'dirty' image can be formed from a suitable linear combination of the 'dirty' images for each pointing $I_p^{\rm dirty}$:

$$I_{\rm linear}(\mathbf{x}) = \frac{\sum_p \frac{A(\mathbf{x}-\mathbf{x}_p)}{\sigma_{I,p}^2} I_p^{\rm dirty}(\mathbf{x})}{\sum_p \frac{(A(\mathbf{x}-\mathbf{x}_p))^2}{\sigma_{I,p}^2}}\,. \tag{3}$$

For objects observed with poor signal-to-noise ratio, this image will be entirely adequate. However, for brighter objects some deconvolution will be required. There are two principal ways of doing this deconvolution. First, for moderate dynamic ranges, one can use two excellent approximations: first, that all the synthesized beams are roughly the same, and second, that the array diameter is much greater than the antenna diameter. Furthermore, in the most compact, D-configuration of the MMA, the coverage of the Fourier plane is completely filled if one accounts for the finite diameter of each array element. It can be shown that to this level of approximation, a simple *linear* deconvolution scheme is adequate. This deconvolution is accomplished by division in the Fourier plane

[9] *Cf.* Thompson, Moran, and Swenson (1986), *op. cit.*

of the transform of the image $I_{\rm dirty}$ by the Fourier transform of an effective 'dirty' beam given by:

$$B_{\rm linear}(\mathbf{x}) = \frac{\sum_p \frac{A(\mathbf{x}-\mathbf{x}_p)}{\sigma^2_{I,p}} A(\mathbf{x}_p) B_p(\mathbf{x})}{\sum_p \frac{(A(\mathbf{x}-\mathbf{x}_p))^2}{\sigma^2_{I,p}}}. \quad (4)$$

For even brighter objects, the two approximations mentioned above will break down and the deconvolution scheme must then correct for different 'dirty' beams for each pointing. In currently used mosaicing algorithms, this nonlinear deconvolution is performed implicitly rather than explicitly: samples of the extended visibility function $V(\mathbf{u}, \mathbf{x}_p)$ are used as constraints in an optimization approach[10] to deconvolution such as the Maximum Entropy Method. We anticipate that linear mosaicing and deconvolution will be adequate to achieve dynamic ranges up to about 100 : 1 while the full nonlinear mosaicing procedure will be required for higher dynamic ranges, these numbers being dependent upon the (excellent) quality of Fourier-plane coverage of the compact configuration.

Beam switching, either by a nutating subreflector or a chopper, will be required to remove the variable emission from the atmosphere. The effect of beam switching is to completely remove the true zero-spacing point and to suppress other spatial frequencies near zero. In conventional 'single-dish' observing, an image can be reconstructed from such data using well-understood and thoroughly tested techniques developed for the Bonn 100-m telescope.[11] For 'single dish' observations, we will use these techniques, while for mosaicing, we will incorporate correction for beam switching into the two mosaicing techniques described here.

An important remaining point concerns the amount of time required to make the total-power measurements. Because of the necessity of removing the effect of atmospheric emission and the need for a tapered illumination pattern for total-power observations, we envisage that when sensitivity is important interferometric measurements and total-power measurements will be made asynchronously. There are two different ways of estimating the required observing time in these two modes, one based upon matching sensitivities and one based upon matching signal-to-noise ratios. If we ignore the necessity for beam switching, then matching the sensitivities of the interferometric and the total-power measurements requires that the latter be observed for a substantially shorter time (the ratio in time, $\sim 0.2:1$, being approximately the square of the filling factor of the array close to the center of the Fourier plane). The inefficiencies of beam switching multiply this number by a factor of about eight, so that the array spends 40% of the observing time making interferometric measurements and 60% making total-power measurements. However this argument is really appropriate only for cases where there is almost no signal at all, which is not

[10] Cornwell, T. J. (1988), *op. cit.*

[11] Emerson, D. T., Klein, U., and Haslam, C. G. T. (1979), *Astron. & Astrophys.*, "A multiple beam technique for overcoming atmospheric limitations to single-dish observations of extended radio sources", **76**, 92.

very interesting for mosaicing! This leads us to the second argument, based upon equalizing signal-to-noise ratios. In a typical mosaicing observation, the signal in the total power (in Janskys) will almost always be substantially greater that that in the interferometric measurements. To see this, let us consider three illustrative, but not exhaustive, cases:

(1) *A source which contains little fine-scale structure:* the signal on the interferometric spacings is very low compared to the total power.

(2) *A point source:* the signal will be about the same for total-power and interferometric observations.

(3) *A random brightness distribution of P resolution 'blobs' in a single pointing:* If F is the total power, then the rms visibility will be about $F/\sqrt{P}$, with peaks of several times higher. Suppose that each blob is the size of a resolution element and that the emission fills the primary beam; then for a packed array of N elements, P is approximately N^2, and so the signal-to-noise ratio of the total-power measurements will be about N times higher than that of the interferometric observations. Increasing the size of the blobs would decrease this ratio correspondingly.

Mosaicing is not important for cases one and two, the former requiring total-power measurements only, and the latter, interferometric measurements of a single pointing. Only case three is a good candidate for mosaicing. Here the time ratio is about $8/P$ (beam switching 'single dish' to interferometer) which is less than about 1% for unresolved blobs filling the field of view, or about 10%, say, for slightly more extended blobs. Our conclusion is that, even taking into account the inefficiencies of beam-switching methods, the time spent in total-power mode could be a least one or two orders of magnitude less than that spent in interferometric observations. Simply stated, mosaicing is most important for well resolved objects, and so, much more time must be spent on the interferometric observations on regions of the Fourier plane where the signal-to-noise ratio is necessarily lower.

4.2. Mosaicing Simulations and Observations.

The arguments made above indicate that the compact configuration of the MMA will be able to (a) image objects on all size scales, and (b) measure the total power sufficiently rapidly. We have augmented these theoretical considerations with an extensive series of computer simulations[12] and actual observations. The key conclusions are enumerated below.

(1) Mosaicing can be used to image complex objects spanning many primary beams. The excellent Fourier-plane coverage of the compact configuration of the MMA allows imaging of complicated, well filled fields. For mosaicing, the effective Fourier-plane coverage is best visualized by convolving the individual sample points with the sensitivity function of the array elements (i.e., the Fourier transform of $A(\mathbf{x})$). Figure IV-4 shows the coverage for a typical snapshot.

[12]M. Holdaway (1990), "Imaging characteristics of a homogeneous millimeter array", NRAO Millimeter Array Memo No. 61.

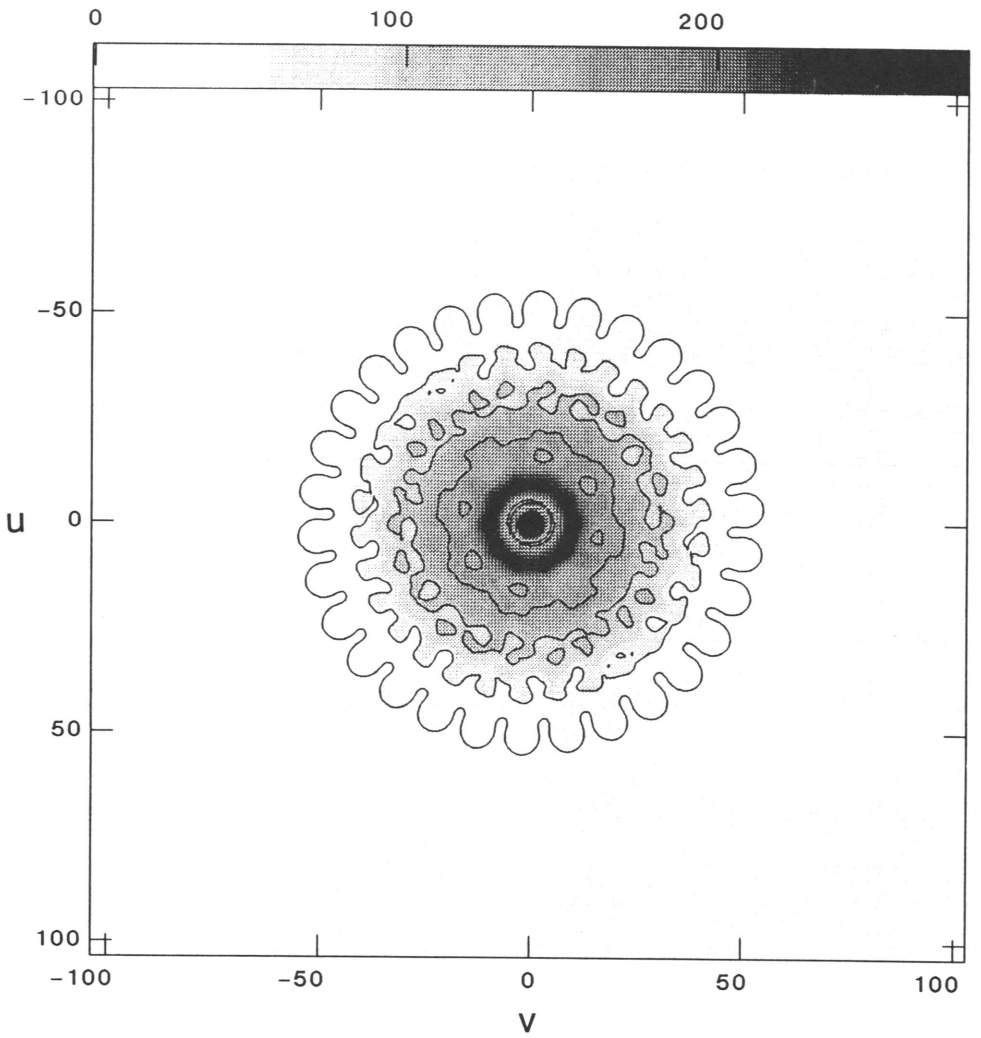


Figure IV-4. The naturally weighted, snapshot Fourier-plane coverage from the D-configuration, accounting for finite diameter of antennas. The units are kilo-wavelengths at 230 GHz, and the contour levels are 1, 10, 20, 30, 50, and 70% of the peak. Most important to note is that this function has no zeros inside the 1% contour, allowing linear deconvolution to remove the fluctuations in this transfer function.

The Fourier plane is effectively completely sampled out to about 70 meters, though some fluctuations in weight remain (these can be corrected in the Fourier transform). In many respects, the MMA is thus essentially equivalent to a single antenna of diameter 70 meters. Figure IV-5 shows a profile through the Fourier-plane coverage for this design of the compact configuration, and also the corresponding cut through a single antenna of diameter 72 meters. Note that the layout of the compact configuration could still be improved substantially: for example, the excess of samples at about 13 meters can be corrected by increasing the size of the outer ring. Similarly, more samples at about 9-meters baseline can be produced by increasing the number of elements in, and the size of, the middle ring. Hence, further fine-tuning is desirable. Ultimately, any remaining

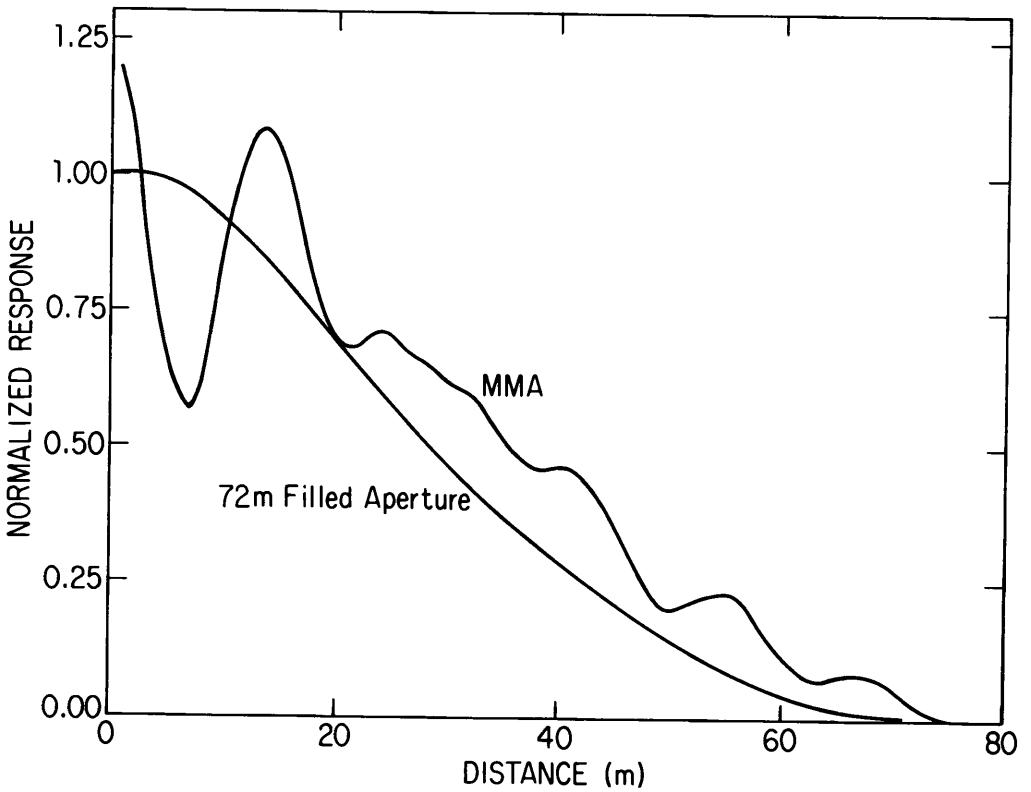


Figure IV-5. Representative cuts through the effective Fourier-plane coverage of the compact configuration and through the coverage corresponding to an equivalent single antenna.

fluctuations in the density of sampling will be corrected in the deconvolution step. The Fourier-plane coverage is excellent in that all spatial frequencies up to the maximum baseline are measured. Consequently, the fidelity of the mosaiced images can be extremely good. For example, although reconstruction of planets from interferometric data is notoriously difficult,[13] a simulation of the MMA performance on a planet (Fig. IV-6) achieves faithful reconstruction to within about 0.5% on the disk.

(2) At some level, pointing errors will limit the dynamic range attainable. This is an important factor in driving the antenna size smaller. We have extensively studied the effects of pointing errors, via both simulations and theoretical analysis. The importance of pointing errors clearly depends upon their characteristics. Because of the large number of array elements, substantial averaging of random pointing errors will occur, and so global and relatively constant pointing errors will be the most damaging, and also the most amenable to calibration and self-calibration. To illustrate this and other points, we have made reconstructions of a moderately complicated object, which is shown in Figure IV-7, observed at 230 GHz in the compact configuration. Forty-nine pointings are required to cover this object, taking about fifty minutes of observation. A representative reconstruction is shown in Figure IV-8. Figure IV-9 shows the dy-

[13] E.g., de Pater, I. (1991), *op. cit.*

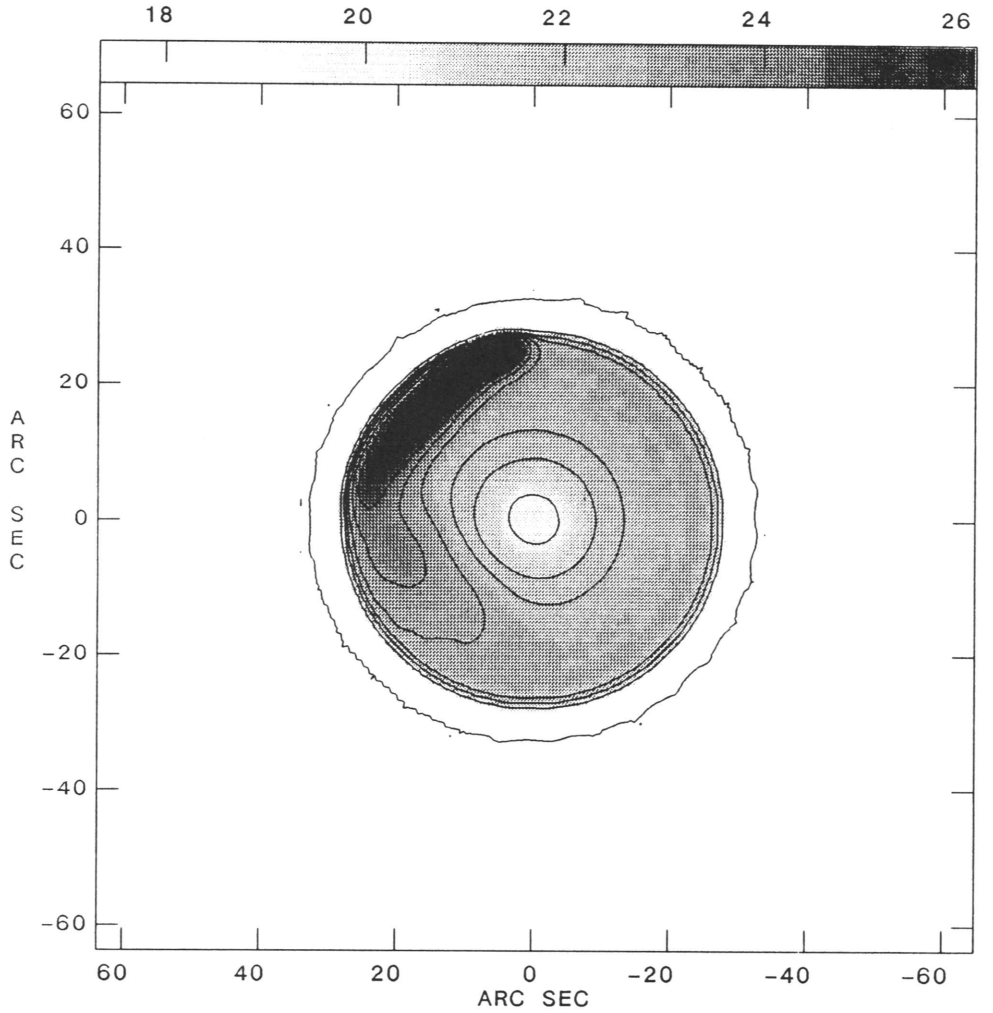


Figure IV-6. Reconstruction of a planet-like model imaged with the D-configuration at 230 GHz. The contour levels are at 0.1, 70, 75, 77.5, 80, 82.5, 85, 87.5, 90, 92.5, and 95% of the peak brightness.

namic range as a function of pointing error. A detailed pointing model was used, consisting of (a) a global constant-with-time pointing error for the whole array, (b) an initial random antenna-dependent, time-independent pointing error, (c) a drift in pointing common to all antennas, and (d) a purely random antenna- and time-dependent component. The components are chosen to be equal, except for (c) which is half the others. This model is based in part upon experience with the IRAM 30-m telescope.[14] Our pointing specification of FWHM/20 (i.e., $1''\lambda_{(mm)}$) is derived from these simulations. It is important to note the following two points. First, the pointing is important for reconstruction of all spatial frequencies, not just the lowest. This means that the pointing specifications cannot

[14] Altenhoff, W. J., Baars, J. W. M., Downes, D., and Wink, J. E. (1987), "Observations of anomalous refraction at radio wavelengths", *Astron. & Astrophys.*, **184**, 381.

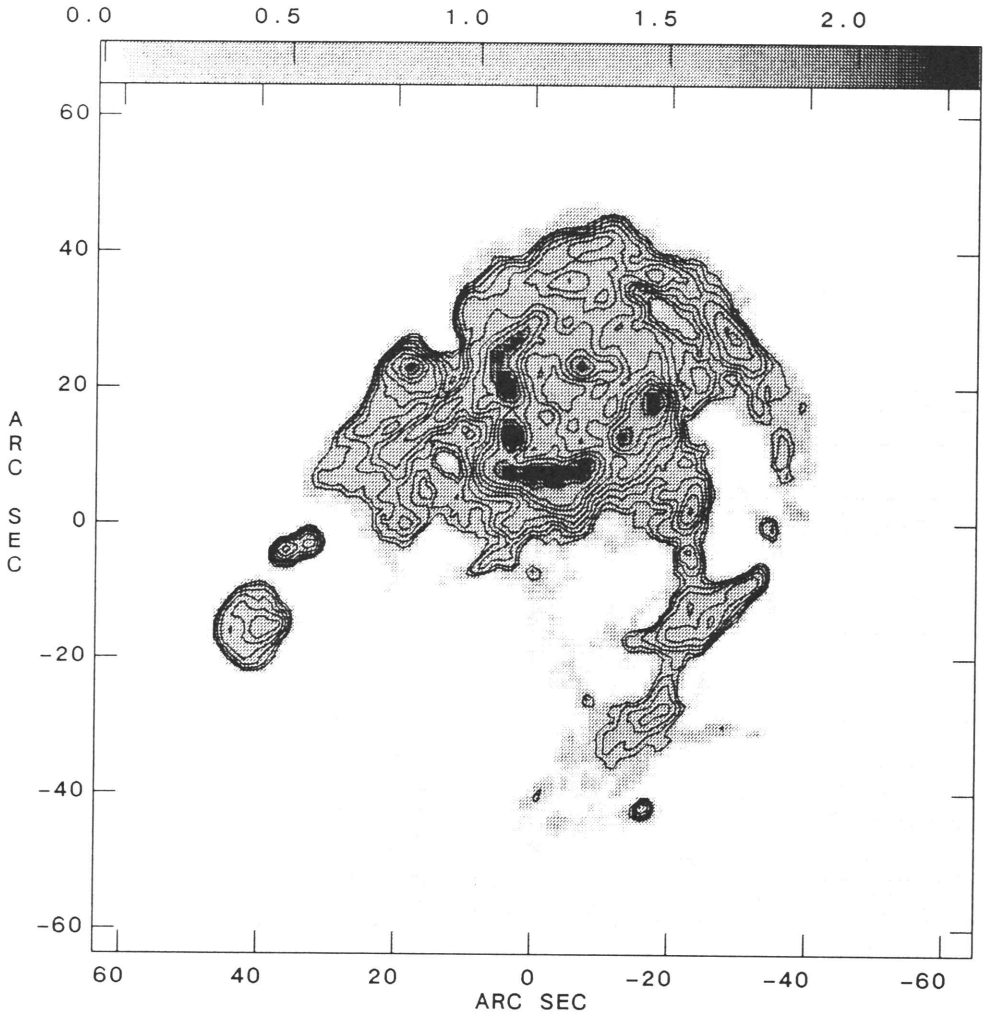


Figure IV-7. Model source. Contour levels are at 3, 4.2, 6, 8.4, 12, 16.8, 24, 33.6, 48, and 67.2% of the peak brightness.

be lowered by using a large single antenna for the total-power observations. Second, the effect of larger pointing errors is to decrease the dynamic range, roughly in inverse proportion to the error, rather than to destroy the image completely. While our assessment of the array performance does not assume calibration or self-calibration of pointing errors, we are confident that in practice it will be possible and that it will be worthwhile for the brightest objects. Also, strategies similar to those now employed in single-antenna observing may be important: for example, scanning a field a number of times will probably reduce the effect of pointing errors just as it does for single-antenna imaging.

(3) We have also investigated the effect of gain drifts and fluctuations on mosaicing in a more limited set of simulations than those performed for pointing errors. The net result is that mosaicing is quite robust with respect to gain errors. For example, random fluctuations of 5% per element per integration,

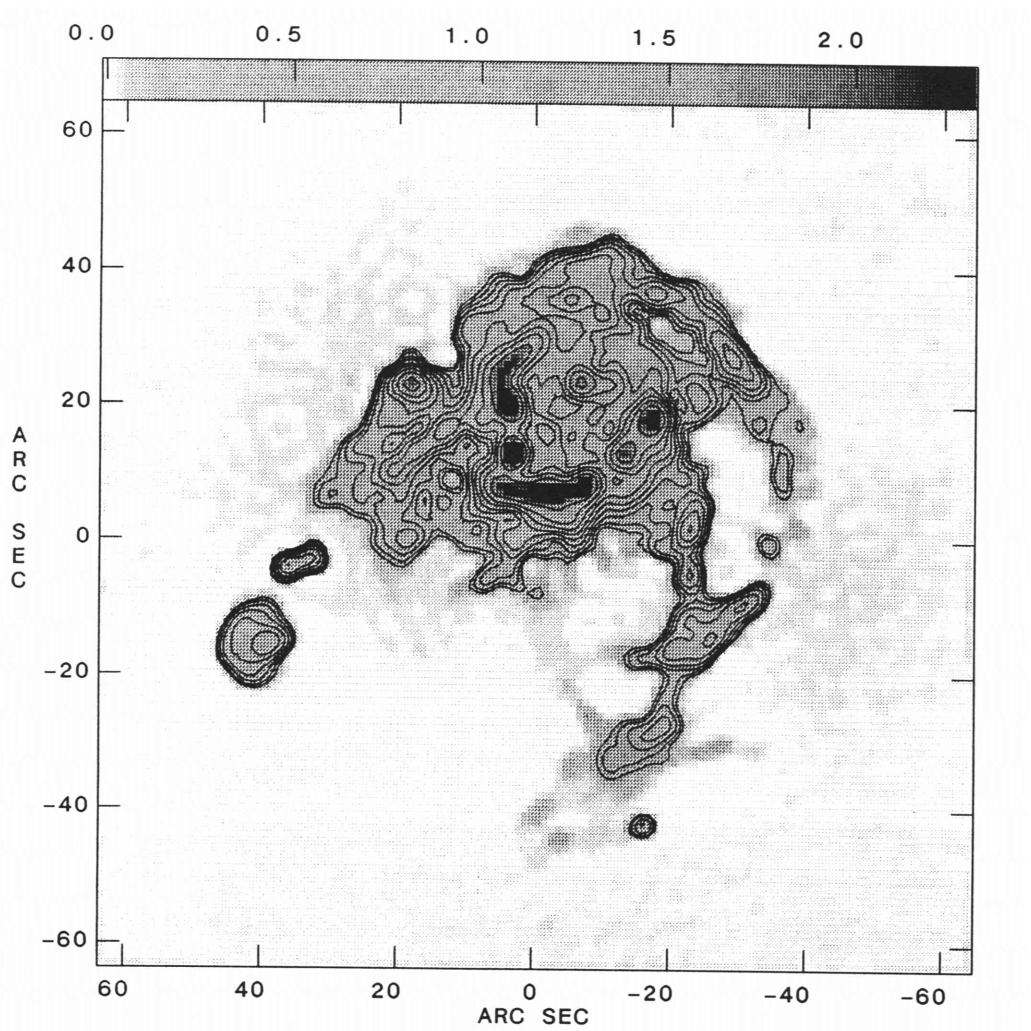


Figure IV-8. Reconstruction from D-configuration mosaic, including pointing errors. Contour levels are at 3, 4.2, 6, 8.4, 12, 16.8, 24, 33.6, 48, 67.2% of the peak brightness. Forty-nine pointings were used, spaced every 20″.

combined with a global drift of 5% over a complete observation, produce little discernible degradation in image quality, though at any given spatial frequency the fractional error is about 5%. Even with 20% gain errors, the apparent image integrity is maintained.

(4) The primary beam patterns will probably be known down to the few percent level in power. Simulations show that this will limit the dynamic range to less than about 800 : 1.[15] A similar argument shows that the surfaces of the array antennas must be quite accurately figured: to keep the error beam at a sufficiently low level, the surface accuracy must be about $\lambda/40$.

[15]Braun, R. (1988), "Mosaicing with high dynamic range", NRAO Millimeter Array Memo No. 46.

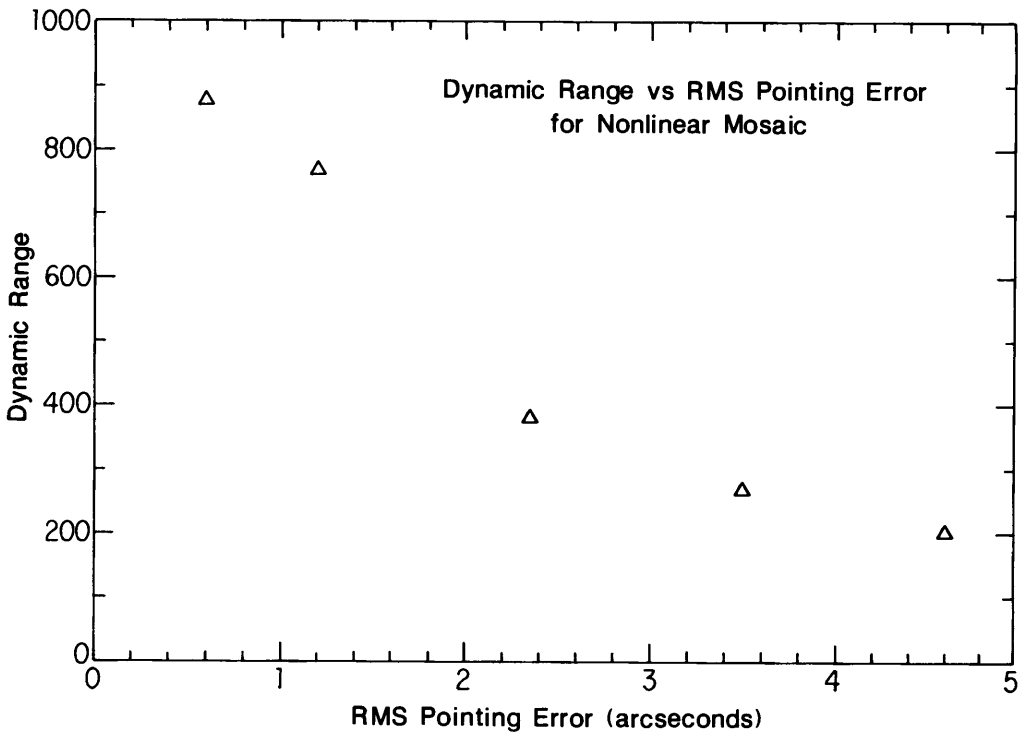


Figure IV-9. Dynamic range as a function of pointing error, for nonlinear mosaicing at 230 GHz.

(5) Incomplete sampling of the image plane, such as occurs when the emission spreads beyond the measured points, introduces some defects. An example of this is given in Figure IV-10, showing the reconstruction of our test object for complete sampling in the image plane (7 x 7), and samplings of 5 x 5, 3 x 3, and 1 x 1. Although some defects are evident, the overall quality of reconstruction deteriorates only slowly with increasingly incomplete sampling.

(6) At moderate signal-to-noise ratios of 100 : 1 or less, mosaiced images can be made using linear mosaicing followed by linear deconvolution, thus avoiding the extra computational costs and built-in structural assumptions of nonlinear deconvolution algorithms such as CLEAN and MEM. The sensitivity of such an image is then just what is calculated naively from the sensitivity of the individual antenna pairs and the primary beam patterns. This linear form of mosaicing will be most important for many spectral-line observations. The computational cost of this type of image is sufficiently small that a limited number of such images will be made available in real time as an aid to observing.

(7) The compact configuration has the ability to measure information on the same size scales as would be measured by a 15- or 20-meter diameter single antenna. Holdaway discusses the imaging quality on these size scales and shows that the residual errors are at the same level as, or lower than, those that would be obtained with a single antenna.

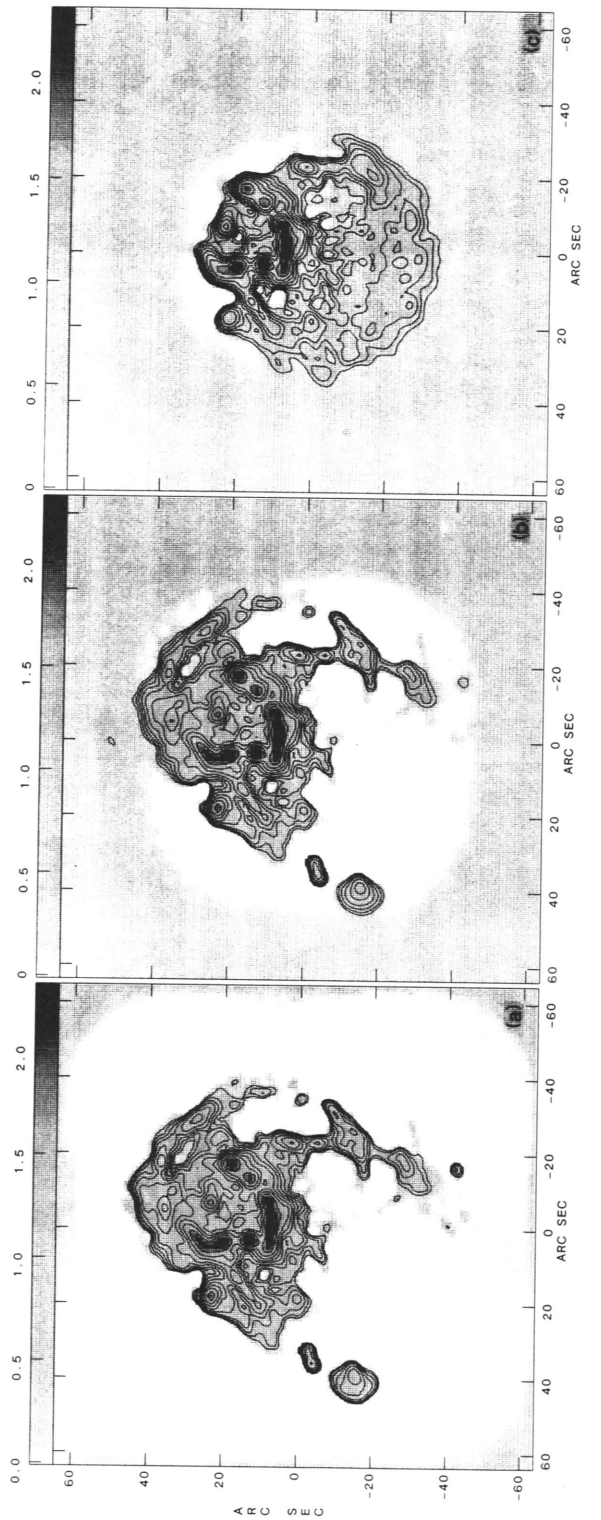


Figure IV-10. The effect of limited sampling on reconstruction of an image via mosaicing: *(a)* 25 pointings; *(b)* 9 pointings; *(c)* 1 pointing. Contour levels are the same as in Figure IV-8. The spacing of pointings was 20″, so, compared to Figure IV-8, only the inner regions were observed.

(8) In the few cases where very-high-dynamic-range mosaicing is required, a modified approach can be used.[16] Since the dynamic-range limitation will be most severe for the fine-scale structure, the high-resolution information from each field is processed separately rather than jointly, and an image is formed by merging the linear combination of the high-resolution separately deconvolved images with the low-resolution mosaiced image.

(9) The computing cost of nonlinear mosaicing is several times that required to image and deconvolve all the pointings separately, and it is several hundred times larger than the cost of linear mosaicing. The computing requirements for imaging are described further in Chapter V.

To summarize all these points, we show a simulation of a mosaiced image of a complex region with peak brightness of 20 K, both in continuum and spectral-line modes at 230 GHz (Figure IV-11). The pointing-error model is that described in item two above.

4.3. Imaging Augmented by a Large Single Antenna. For more than a decade, techniques designed to improve the sensitivity of interferometric observations to low-brightness, spatially extended emission have been in use and have undergone refinement. The ability of existing interferometers to measure visibilities at low spatial frequencies is limited, as described above, by the finite physical separation of the antennas, restrictions on the antenna placement, or the small total number of antennas—the Fourier plane is poorly covered at short spacings. For example, the filling factor of the VLA in its most compact configuration is less than two percent; for BIMA, OVRO, IRAM, or the Nobeyama array, it is less than or about ten percent. The brightness sensitivities of these sparse arrays improve dramatically if the interferometric data are complemented by total-power data from a large single antenna, whose diameter is substantially larger than the shortest separation between interferometer elements. And this is the technique that has been used for low-surface-brightness measurements with interferometers heretofore.

The distinction between a single antenna and an interferometer becomes much less readily discernible when one considers the MMA in its compact mosaicing configuration, which is 50% filled and in which the short spacings are well sampled. Nevertheless, it is worthwhile to ask whether, and to what extent, the MMA mosaic images would benefit by complementary information from a single antenna of diameter twice that of the array elements (viz., 16 m) or larger. This is a complicated question that can be answered neither by appeal to experience—for there never has been an array, at any wavelength, designed with total-power systems and a filled compact configuration—nor by one-dimensional geometric arguments. Instead, the answer must be sought by defining the observation to be made, or the classes of observations, and then simulating the array performance with and without the contributions of a large single antenna. Proper treatment of observational errors is an important part of the simulations. This is a major undertaking that will occur during the pre-construction phase of the MMA in

[16] Braun, R. (1988), *op. cit.*

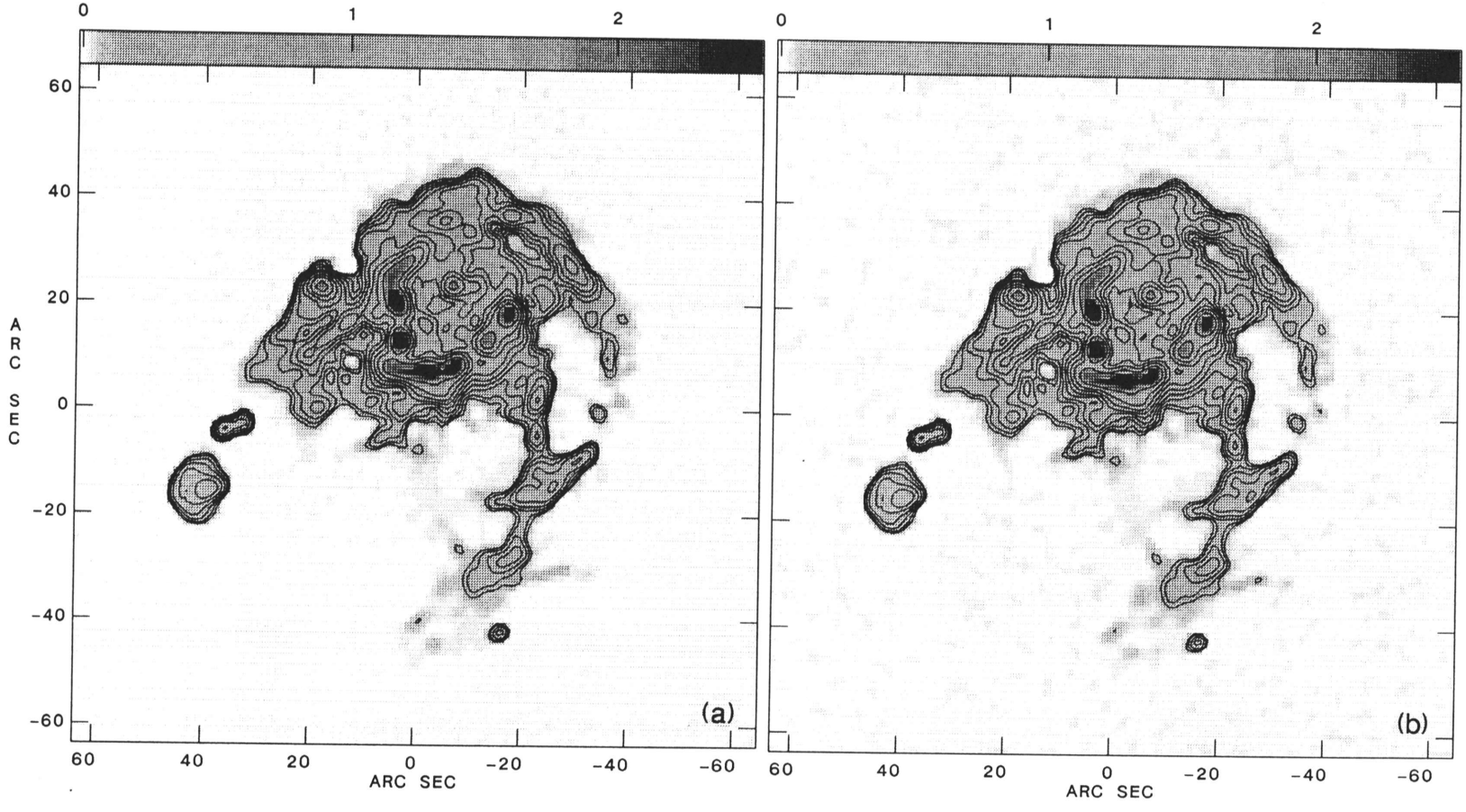


Figure IV-11. Full reconstruction from D-configuration observations at 230 GHz of standard model scaled to 20-K peak brightness: *(a)* in continuum; and *(b)* with 100-kHz spectral resolution.

collaboration with the community of MMA users and with careful attention to suggestions from that community.

The MMA design as presented here does not include a large single antenna for measurement of the visibilities at low spatial frequencies. Recent experimental tests, performed by augmenting VLA data with total-power data measured by a VLBA antenna,[17] provide added confidence in the MMA computer simulations. However, if as a result of the on-going imaging simulations described above it becomes apparent that an improvement to the MMA mosaic images can be realized by inclusion of data from a large single antenna, and if that improvement will justify the increased MMA operational cost and the burden associated with providing and maintaining one-of-a-kind instrumentation, then the array design will be modified accordingly.

[17]Langston, G., Uson, J., and Cornwell, T. (1990), in preparation.

V. COMPUTING

1. Overview

In the past, array-telescope computing has suffered from two difficulties: First, most arrays have lacked sufficient computing resources, at least in the analysis area, which has led to limitations in the science and to frustration amongst astronomers. Second, while large strides have been made in the sophistication of the imaging algorithms, less effort has been expended in easing the complexity of reducing data from array telescopes. For the VLA, both of these factors have been important in making spectral-line observations, in particular, much more burdensome for the observer than should have been the case for a state-of-the-art imaging instrument. The desire to avoid repetition of this unfortunate situation leads us to two basic goals for the Millimeter Array computing system: it must not constrain the scheduling of the MMA (i.e., sufficient power must be available for most experiments) and second, for the observer, the use of the MMA must be no more difficult than the use of a single-antenna instrument. In short, the computing systems must be fast enough and easy to use. These goals have impact in both the on-line, array-control and data-acquisition computers and in the off-line, imaging computers.

2. Array Control and Data Acquisition

The on-line computing requirements of the MMA will be slightly greater than those of the NRAO's VLA because of several factors: the increased number of telescopes, the increased capabilities of the correlator, the central role of mosaicing, and the importance of real-time scheduling to take advantage of periods of good 'seeing'.

The on-line system must perform the following tasks:

- Allow the array operator to control and monitor the whole system;
- Control of the array elements, the sub-arrays, the array as a whole, and the correlator;
- Monitoring of system parameters and data quality; and
- Archiving and transmission of data to the off-line systems.

There is also a need for real-time imaging, but this will be performed on the off-line computers, taking advantage of the direct optical-fiber connection of the MMA site with central NRAO facilities.

There are a number of feasible designs for the MMA on-line computer system, based, for example, either around a single monolithic computer or around a network of smaller computers and microprocessors. The former is slightly easier to design and program, while the latter has the advantage that the cost of spares is much lower. We will adopt a hybrid between these two: a number of separate but tightly coupled processors will be housed in one cabinet and will share

memory and other resources.[1] For control and programming, workstations will be provided for use by the operators and the programming group. This scheme has some of the virtues of the other two possibilities: minimizing the cost of peripheral devices and spare units, and easing programming through the use of shared memory. The level of processor required is modest. The overall cost including operator/programmer workstations, one high-level display workstation, and archiving hardware is about $200 k. The manpower required for programming the on-line system will be three individuals during the construction phase. Ongoing support will be provided from the operations budget.

The on-line system will distribute its data in two ways: first into an archive at the MMA site, and second, via an optical-fiber link to the computers at one of the other NRAO sites (see Section 4). We will have sufficient computer resources at both facilities for extensive monitoring of the data quality, via flexible displays of the data in many different forms, and via quick imaging, following initial real-time calibration and editing. This will be essential for maintenance, testing, and scheduling of the MMA.

3. Imaging Computer Requirements

The off-line computing requirements of the MMA will be somewhat less than those of the VLA as defined in the NRAO Array Telescope Computing Plan. In the large configurations, A-, B-, and C-, the computing required for MMA imaging is substantially less than that for the corresponding VLA configurations because, although the number of channels is increased by a factor of two, the number of pixels required is generally about an order of magnitude lower. Furthermore, the subtle and computationally burdensome errors which corrupt low-frequency, wide-field VLA images, such as non-isoplanatism and non-coplanar sampling, are not important at these high frequencies and small fields of view. In the smallest configuration, the D-configuration, the computing costs will be dominated by the need for mosaicing. The excellent Fourier-plane coverage of the compact configurations of the MMA will minimize the need for nonlinear mosaicing incorporating deconvolution, and instead, linear mosaicing augmented by linear deconvolution will be sufficient in many applications, resulting in large savings in computing.

From our existing experience with imaging, we can predict the computing time required for most observing projects. Here we give a number of illustrative examples. As a standard, we have chosen to use a machine of the level specified in the recent NRAO Array Telescope Computing Plan Addendum: a 2-Gflop, 2-Gbyte memory machine with a disk-to-memory transfer rate of about 1 Gbyte per second.

- *A four-hour full spectral-line observation in the highest-resolution (A-)configuration, with deconvolution of the continuum and self-calibration of the continuum or a bright channel:* This will take about one hour of run time. Adding deconvolution for each channel will increase this by a factor of fifty or so.

[1] This type of approach has been followed by the VLBA correlator project.

- *Mosaicing of 10,000 pointings using the smallest (D-)configuration and covering all channels:* This corresponds to a 30′ cloud imaged at 230 GHz. With a 20-second duration for a given pointing (10 seconds observing and 10 seconds setup, moving telescopes, etc.), the total observing time is about two days, while the required computer time is about five days for nonlinear mosaicing and about five hours for linear mosaicing.
- *Adding C-configuration to the previous example* will increase the run-time by about an order of magnitude, though correspondingly more observing time may be required to achieve adequate surface-brightness sensitivity.
- *Mosaicing of 100 pointings using the C- and D-configurations and covering all channels:* This corresponds to a few-arcminute cloud imaged at 230 GHz. With a 20-second duration for a given pointing, the total observing time is about ten hours, while the required computer time is about five days for non-linear mosaicing and about twelve minutes for linear mosaicing.

These examples are somewhat ambitious, and many projects will consume considerably less computer time. For example, a full B-configuration spectral-line observation will take about 10 to 15 minutes if deconvolution of each channel is not important, and about 8 to 9 hours if it is. Since the required computing varies over such a large range for the many different types of projects, to make a good prediction of the computing requirements we would need to be able to estimate the rates of occurrence of the various types of project. While this type of analysis is possible at centimeter wavelengths with the VLA, where the science is relatively more familiar and better understood, it is clearly inappropriate for a ground-breaking array such as the MMA. In addition, it is important to remember that the algorithmic costs will almost surely increase with increasing sophistication of the imaging. For example, self-calibration of the pointing errors or self-calibration with a frequency-dependent model would increase some of these examples' demands in run-time by several factors of two. Taking this into account, along with a margin of safety for the inevitable inefficiency of data reduction, we believe that a suitable machine or system would have the following components:

- A total CPU power of about 8 Gflop, costing about \$5 M in 1995. (See Section 5 for a discussion of the justification for this predicted price/performance ratio.)
- At least 300 Gbyte of off-line storage, costing about \$1 M in 1995.
- About ten high-performance graphics super-workstations for display and editing of data and for display of data cubes, costing a total of about \$500 k (1995). Efficient and capable data display of the large datasets and large cubes will require state-of-the-art display hardware and software.
- A large number (40–50) of high-performance personal workstations, costing about \$10–20 k (1995), for a total of \$0.5–1 M (1995).

For a number of simple projects, these workstations will be adequate.

- A system of mass-storage for processed datasets, similar in principle, for example, to the Common File System used at the National Center for Supercomputing Applications. This machine will provide backup and retrieval of large (~Gbyte) datasets to and from the fast-access disk storage. This is vital both for remote use of the MMA off-line computers and for easing the problems of controlling the use of the large amount of disk space envisaged. No widely accepted commercially available solution to this problem exists yet, but it may become available in the next five years. We expect that a satisfactory system would cost about $1 M.
- Local-area-networks are needed, with a hierarchy of capacities ranging from the Fiber-Distributed Data Interconnect (FDDI) or an enhanced Ethernet for the workstations, to a high-capacity system such as the High-Performance Parallel Interface (HPPI). Connections onto a national network such as the proposed National Research and Education Network (NREN) will be vital in the long run for providing remote observing and data-reduction facilities. We estimate the cost of networking as about $0.5 M (1995).
- For a number of small items such as image recording and hard-copy facilities, we would expect that about $200–300 k would be adequate.

The total cost of this system is between $8.7 M and $9.2 M (1995). Adding in the cost of the on-line system and a contingency of about 10%, we come to a total capital budget for computing equal to $10 M (1995). Some of the capabilities described here overlap with the recently proposed NRAO Array Telescope Computing Plan (1990). If that proposal is funded, then the MMA computer budget could be reduced to about $7 M (1995). The manpower required for programming the analysis software will be in addition to the staff currently engaged in such work for the VLA and the VLBA. We estimate that three new programming positions will be needed, starting during construction phase. Support staff are included in the operations budget.

In addition to the investment in hardware described above, it will also be vital to make progress in the analysis software. A major software goal will be simplicity of use: new approaches to software design will be vital in hiding the complexity of the MMA. In connection with this point, we note that NRAO has convened a Software Advisory Group to advise the Director on long-term goals and strategies for NRAO analysis software. This group consists of both internal and external members, having expertise and interests ranging from software design to astronomical data reduction.

4. On-Line Data Rates

Every integration period, the correlator will provide data from 780 baselines, each having a maximum of 1024 channel-polarization products. Allowing

eight bytes per (complex) baseline-channel-polarization product (probably IEEE single-precision format would be convenient), the total volume of data from an integration period is about 6.5 Mbyte. (Dumps of the autocorrelation data will produce about 140 kbyte.) The maximum dump *rate* is dependent on the minimum integration time, which is set by two constraints: first, Earth-rotation will smear the visibility function significantly if the integration time is equal to or greater than the time taken for the longest baseline to swing through an antenna diameter; second, mosaicing of a large field of view will require a low dwell time for each pointing. Ten seconds provides an acceptable lower bound for the integration time: giving at most 1.5% loss in intensity of an object near the first null of the antenna power pattern, and a noise level of about 4 mK for continuum in the D-configuration (0.5 K for galactic spectroscopy). The corresponding maximum dump rate is about 0.6 Mbyte/sec. For high-time-resolution observations, it will be possible to trade time resolution for frequency channels.

At the maximum data-dump rate, a typical spectral-line observation of a few hours' duration will require about 10 Gbyte of storage, and a day's archive will consume about 80 Gbyte. While suitable compact archive media exist now in, amongst other forms, optical disks and digital audio tapes, the dearth of standards and the rapid pace of developments mean that the exact choice of archive format should be deferred. We envisage that the data will be both archived at the MMA site and transmitted to more accessible NRAO computing facilities via optical fiber, and thence by NSF network to the user's home institution, if desired, for immediate analysis. The close connection of the imaging computers to the MMA means that these, too, can be considered adjuncts of the on-line system. The resulting ability to interact in real time with the data from the MMA will be vital, both for effective scheduling of observing time and for maintenance, test observations, and debugging. Initial data editing and calibration will be performed in real time so that a useful amount of immediate imaging will be possible.

5. Prediction of Cost of Off-Line Computers

Our strategy for deciding computing costs has been to predict the cost of a satisfactory system in 1995. The NRAO Computer Division has tracked high-performance computing over the past seven or eight years. Over that period, the typical price/performance ratio for mini-supercomputers (roughly speaking, a ~$1 M machine), has improved by a factor of two every two years. For example, the Convex C1- class machines, vintage 1986, will be superseded at the end of this year by machines about six times faster for the same cost. We can predict that by the end of this year, a 2-Gflop machine (actually consisting of a moderate number of processors, say 8) will be available for about $5 M, and thus using our rule of thumb, a Gflop will cost about $600 k in 1995. The current cost of high-performance disks is about $12 k per Gbyte, and this is also improving at about the same rate, so in 1995 the cost will be about $3 k per Gbyte.

VI. SITE SELECTION AND TESTING

1. Site Criteria and Evaluation

The scientific goals of the Millimeter Array require that it be located on a high, dry site offering the possibility of baselines at least 3 km in length. The atmospheric qualities are driven by the desire to minimize the loss of signal caused by atmospheric attenuation and the increase in system temperature contributed by emission from the sky. Both of these effects will be serious in the 1.3- and 0.85-mm bands and can be minimized only by locating the site at high altitude, above most of the water vapor. The specific need to observe at wavelengths of 1 mm and shorter requires that the total water vapor over the site be frequently 2 mm or less, and this in turn suggests that the site be at an elevation of at least 9000 feet. The physical size of the MMA is derived from the requirement of $0''.1$ resolution at 230 GHz.

Our approach has been to search exhaustively for such sites on U.S. territory with latitudes less than 36°, in order to maximize the sky coverage of the Array and to avoid the generally poorer weather patterns and increased amounts of snow characteristic of latitudes north of 36°. From a thorough study of topographical maps, we first assembled a list of the fifty sites south of 36° that have any land above 9000 feet. We believe this list to be complete. We then eliminated those potential sites that were clearly too small, were in national wilderness areas, or had some other overriding problem such as access. The few that survived this process were studied in more detail and were visited, if necessary, to assemble the needed information.

Sites north of 36° were also studied in the same way, except not quite as exhaustively. Here we concentrated on sites at least 3 km in extent. Several of these potential sites were visited, including the Aquarius Plateau in Utah and Grand Mesa and South Park in Colorado. None of these sites was found to be suitable, for various reasons.

We have also studied use of Mauna Kea and sites in northern Chile. In the case of Mauna Kea, the best potential site we can identify which is big enough to accommodate the MMA is well below the summit, at approximately 12,500 feet. Even this site is steeper than would be acceptable according to the criterion we used to eliminate sites in the Southwest, it has a limited horizon to the west, and it may not have a stable foundation on which to construct the antenna pads and the roads we would need. Also, it lies outside the area of the mountain reserved for astronomy. Nevertheless, we will continue to study it.

Another serious concern about sites outside the southwestern U.S. is the cost. First, construction either in Chile or in Hawaii comes at a premium, because of the higher costs of the necessary materials and the scarcity of workers with the necessary construction skills. Second, the MMA in New Mexico or Arizona can share facilities and personnel with other NRAO activities; the cost of replicating this infrastructure at a remote site is a premium paid once in

Figure VI-1. Geographic locations of potential MMA sites and other NRAO sites.

construction, and then again in operating costs throughout the life of the instrument. Thus, unless it can be shown that continental sites are inadequate for the MMA, we will not pursue these more adventuresome and expensive alternatives.

We have now restricted our studies to three Southwestern sites south of 36° that appear to meet the needs of the MMA. Two are in the Apache National Forest near Springerville, AZ and Alpine, AZ, respectively, and one is in the Magdalena Mountains near Socorro, NM. The relative locations of these and other NRAO sites are shown in Figure VI-1.

(1) The Springerville site is on a large (10 km × 10 km) mesa about 10 miles south of Springerville and Eagar, AZ at 9200 feet elevation. The site is mostly used for cattle grazing. It has a recreational area at its southern end (Big Lake and Crescent Lake) which is heavily used during the summer months. Because of this activity, the area of the plateau available to the MMA may be restricted.

(2) The Alpine site is a 3 km NS × 2 km EW site on the southern skirt of Escuadilla Mountain at 9800 feet elevation. It is a mostly open mountain meadow. North of the site is Escuadilla Mountain, which is a designated wilderness area. The site has some grazing, some logging, and some recreational activities on it. However, it is much less used than the Big Lake camping area on the Springerville site.

(3) The Magdalena Mountains site is mostly in the Langmuir Laboratory scientific reserve, which includes South Baldy Peak in the Cibola National Forest. The site is irregularly shaped but is about 5 km NS × 3 km EW and at 10,500 feet elevation. Langmuir Laboratory is a Congressionally mandated scientific preserve for atmospheric and astrophysical research. It is used mainly during the summer months for thunderstorm research and less actively year-round for optical astronomy.

All three of these sites are close enough to Socorro and the existing VLA/VLBA Array Operations Center (AOC) for the MMA operations support to be located there. The site in the Magdalena Mountains is closer to Socorro than the VLA or any of the VLBA sites. The other two sites are almost equidistant between Socorro and Tucson. Thus, for both of the Arizona sites, most of the higher-level MMA support could be based in either Socorro or Tucson, while an operations crew living in or near Springerville would maintain the MMA site facilities.

Each site has other advantages and disadvantages. For example, the site in the Magdalena Mountains has less snow than either of the two Arizona sites. A design goal of the project is to allow the possibility of expanding the MMA to longer baselines than 3 km sometime in the future. The Magdalena Mountains site would allow this by extending the Array onto the Plains of San Agustin at 7000 feet, which contains the VLA site and extends over 100 km from the MMA site, all of it line-of-sight. The Springerville site would allow 10-km baselines if other parts of the site are available, as mentioned above. The mountainous terrain around the Alpine site would not allow extensions to longer baselines than three kilometers.

The main advantage of the Arizona sites is that they allow a more flexible layout of antennas in the largest configuration of the MMA. The Springerville site would probably allow an optimum configuration. The Alpine site is very restricted but happens to have a shape allowing a fairly good configuration. In order to investigate this problem we have studied possible configurations for the largest MMA configurations on each of the three sites. Starting with the princi-

ples developed in our earlier theoretical work,[1,2] we have used the AIPS program UVSIM to generate the u-v coverage for a series of hypothetical configurations on each site. Each of these was analyzed for goodness of coverage until we found a solution consistent with the site topography which came closest to matching our goals. In Figures VI-2a, 2b, and 2c we show our best A-configuration layout for the MMA on each site. In Figure VI-3 we show the corresponding u-v coverage for a four-hour synthesis of a source at 30° declination. In Figure VI-4 we show the u-v coverage for a source at −30°. The beam parameters are summarized in Table VI-1 for −30°, 0°, 30°, and 60° declination. While the u-v coverage is best for the Springerville site, the u-v coverage is good for all the sites—in fact, better than the VLA's.

Table VI-1. Synthesized-beam major-axis and minor-axis full-widths at half-maximum, θ_1 and θ_2, at $\lambda = 1$ mm, for the Alpine (Terry Flat), Springerville, and Magdalena Mountains sites. (A-configuration and assuming natural weighting.)

Site	Declination	θ_1	θ_2	Position Angle
Terry Flat				
	−30°	0.″125	0.″089	15.°8
	0°	0.″095	0.″071	68.°9
	30°	0.″093	0.″059	77.°1
	60°	0.″094	0.″065	74.°5
Springerville				
	−30°	0.″116	0.″073	−11.°9
	0°	0.″078	0.″069	−60.°5
	30°	0.″075	0.″059	−74.°0
	60°	0.″075	0.″066	−67.°8
Magdalena Mts.				
	−30°	0.″134	0.″088	−8.°1
	0°	0.″090	0.″074	−70.°3
	30°	0.″089	0.″058	−79.°1
	60°	0.″089	0.″066	−73.°1

2. Atmospheric Tests

Probably the most critical parameter characterizing a site is the atmospheric transparency. A limited amount of low-resolution data (of both low time-resolution and low spatial resolution) is available from satellite work, and additional material is available for those sites which are close to radiosonde

[1] R. M. Hjellming (1984), "Evaluation of some initial possibilities for the large configurations of the proposed millimeter array", NRAO Millimeter Array Memo No. 21.

[2] T. J. Cornwell (1988), "A novel principle for optimization of the instantaneous Fourier plane coverage of correlation arrays", *IEEE Transactions on Antennas and Propagation*, **AP-36**, 1165.

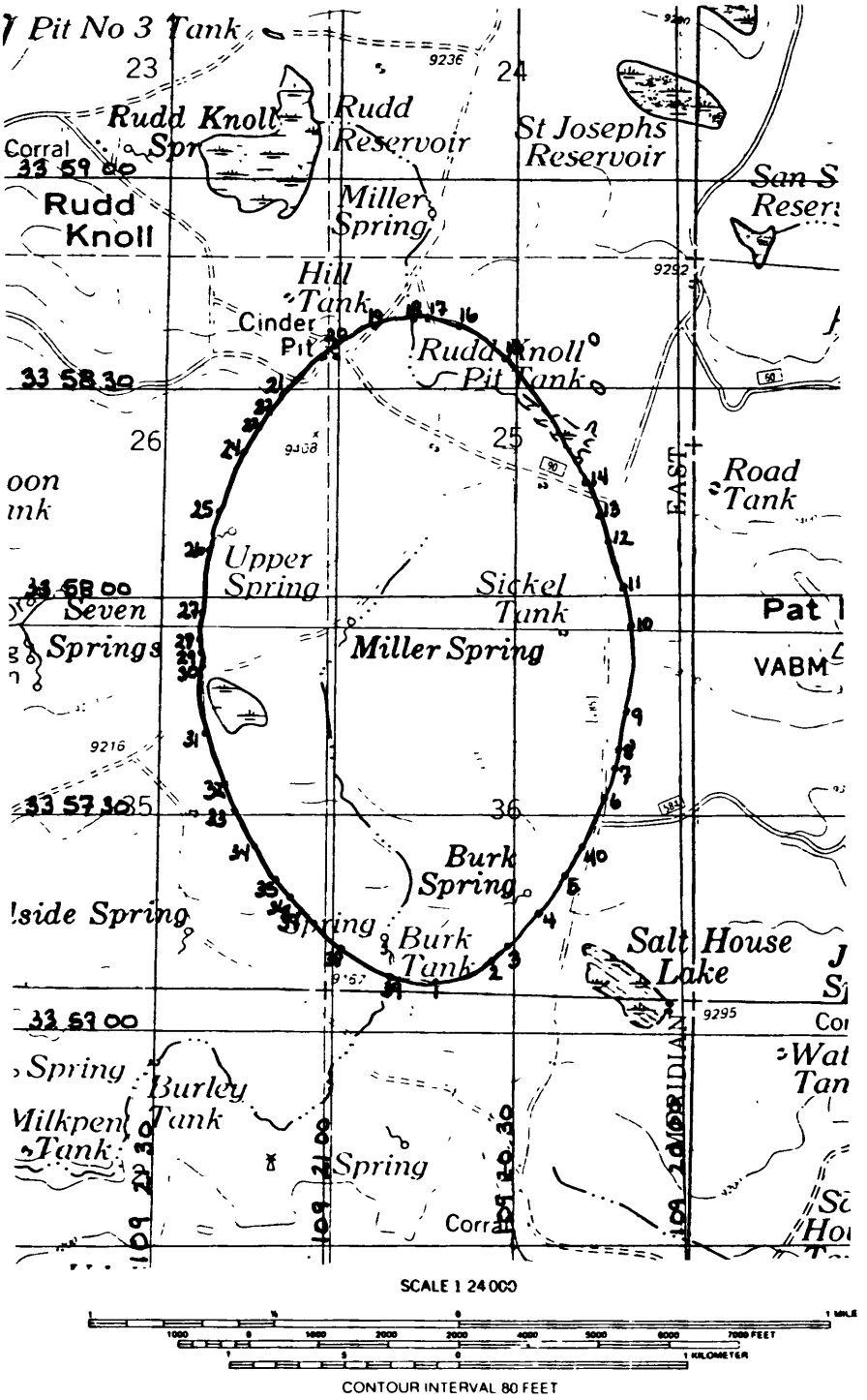


Figure VI-2a. Three-kilometer MMA configuration on the Springerville site.

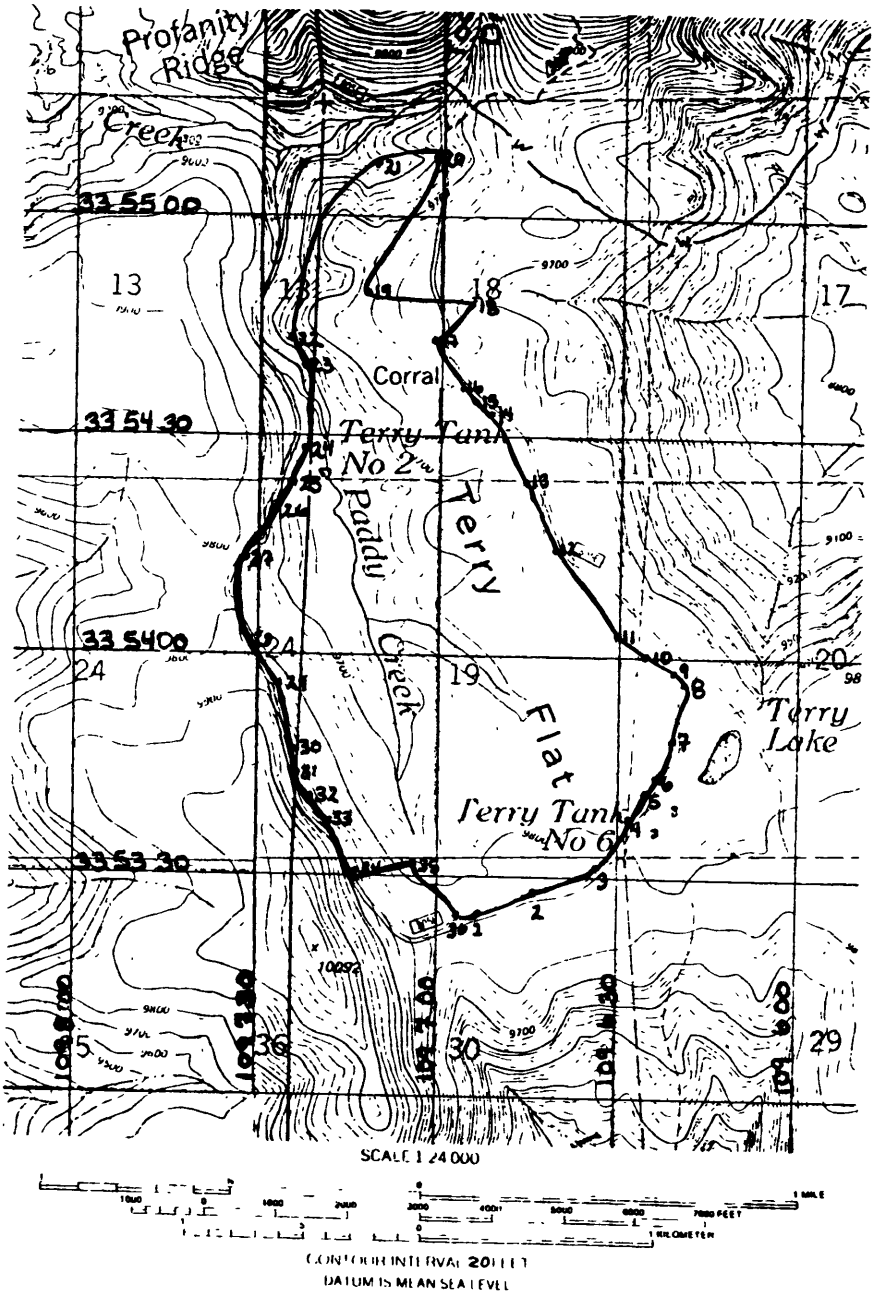


Figure VI-2b. Three-kilometer MMA configuration on the Alpine site.

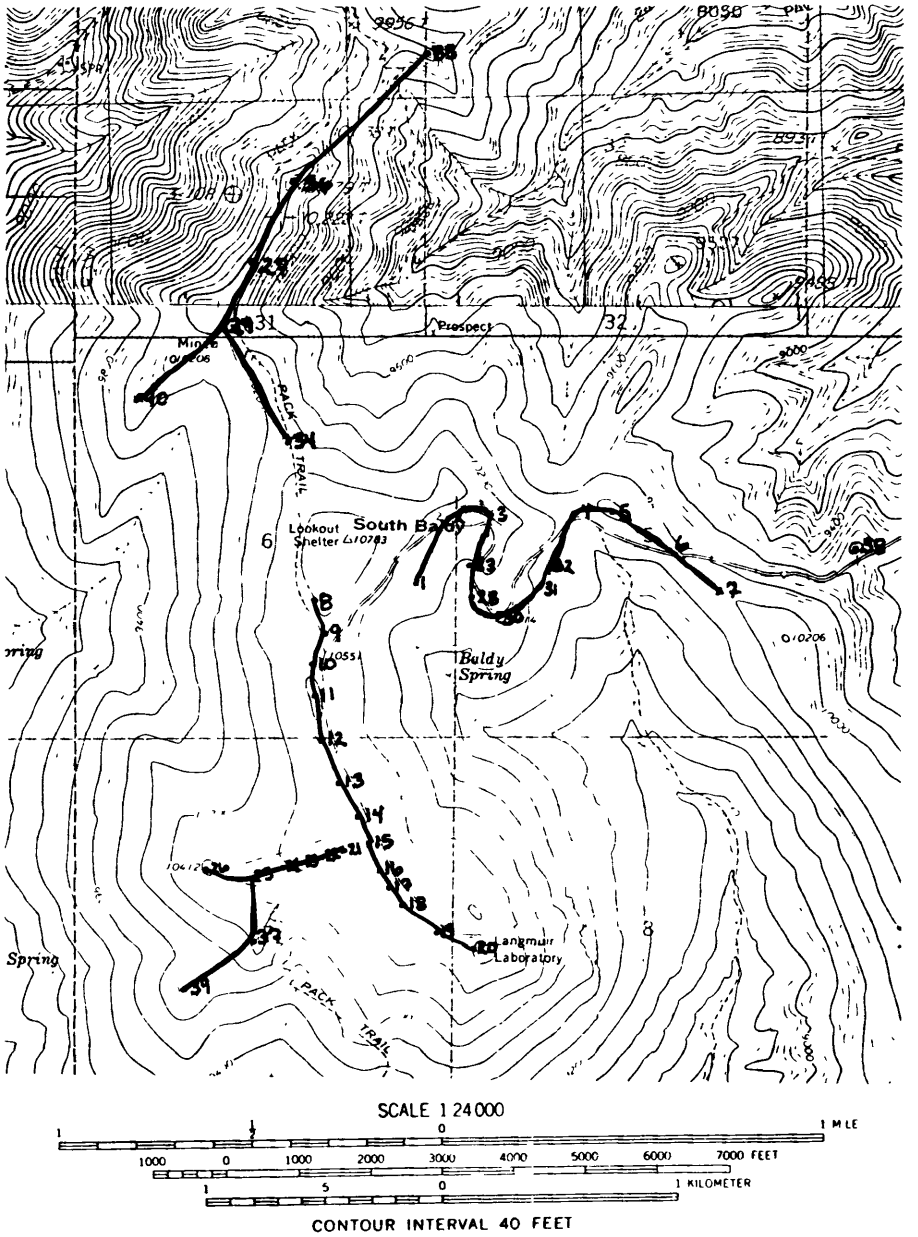


Figure VI-2c. Three-kilometer MMA configuration on the Magdalena Mountains site.

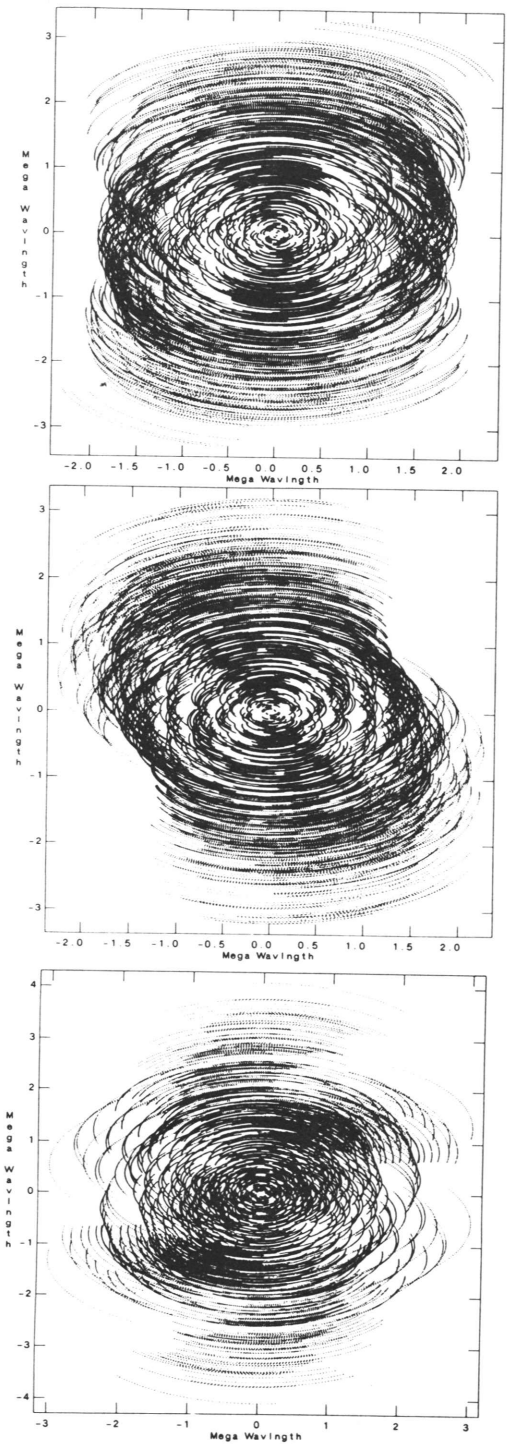


Figure VI-3. The u-v coverage for a four-hour synthesis observation of a source at 30° declination. *(Top)* The Springerville site; *(Middle)* the Alpine site; *(Bottom)* the Magdalena Mountains site.

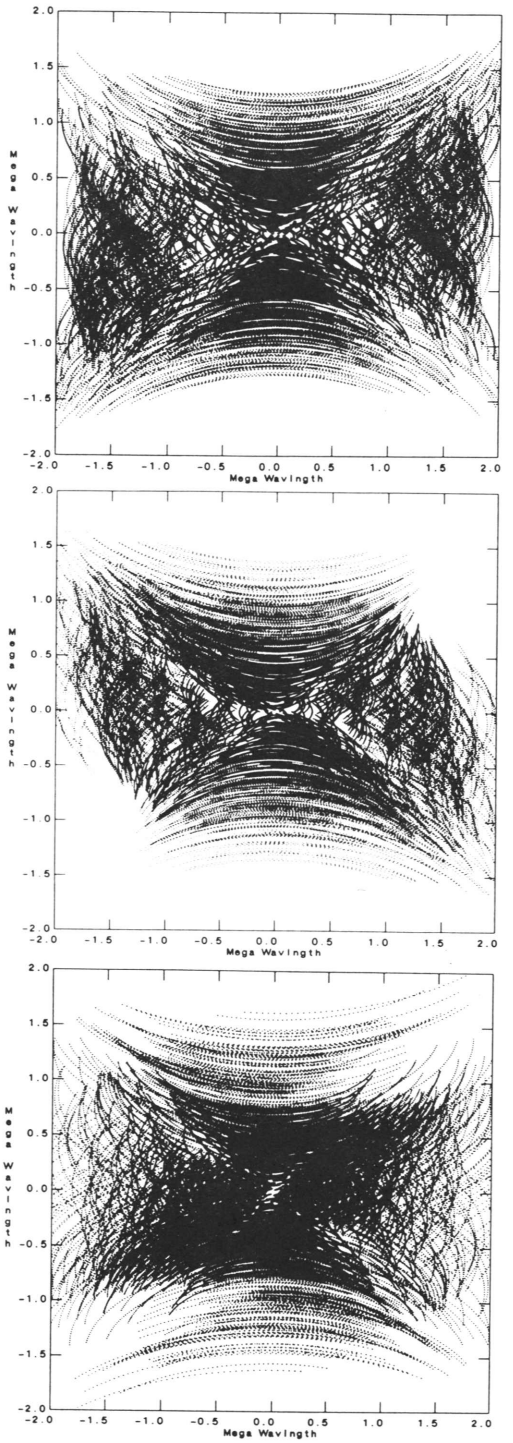


Figure VI-4. The u-v coverage for a four-hour synthesis observation of a source at $-30°$ declination. *(Top)* The Springerville site; *(Middle)* the Alpine site; *(Bottom)* the Magdalena Mountains site.

launch facilities. Some time ago, however, it became evident that the existing data were not sufficient for our needs. In order to decide between the sites we need to study the atmospheric opacity and phase stability at each site. To accomplish this task, we have constructed three 225-GHz tipping radiometers. A fourth has been built for use with the NRAO 12-Meter Telescope on Kitt Peak. One of these radiometers has been operating for the past four years at the Joint Observatory for Cometary Research on the Langmuir Laboratory site in the Magdalena Mountains. A second radiometer has been in operation at the Caltech Submillimeter Observatory on Mauna Kea since August 1989. The third will begin operation on the Springerville site in the summer of 1990.

The 225-GHz tipping radiometer is housed in a weatherproof aluminum enclosure and requires a 12-volt power source and a PC-compatible computer for control. In Figure VI-5 we show a photograph of the radiometer opened up to expose the electronic components. The radiometer package is a rectangular box of roughly one-cubic-foot volume. A machined parabolic reflector with approximately ten square inches of collecting area is mounted external to the box. This mirror may be positioned in 1?8 increments by computer control of a 200-step drive motor. Signal from the sky is reflected off the mirror and enters the radiometer enclosure through a one-inch diameter mylar window and illuminates another identical, fixed parabolic reflector. A rotating 'chopper wheel' is positioned at the common focus of the two reflectors, interrupting the signal path. This chopper wheel is actually a rotating, flat-bladed reflector which alternates the signal source between the sky and two temperature-regulated RF absorbers, one at 45° C and one at 65° C, that serve as calibration sources. From the signals generated by this configuration, the control computer then calculates the system gain from the ratio of the 65° C signal to the 45° C signal and from the actual measured physical temperatures of the two loads. Using these data it corrects the observed sky signal to an absolute temperature scale. From measurements at ten different elevations, equally spaced in $\sec z$ between 1.0 and 3.0 air masses, it then fits a curve to the data to derive the 225-GHz optical depth at the zenith.

In Figure VI-6 we show an example of the type of data now being obtained with these radiometers. We have plotted the fraction of the total number of hours in the period November 10, 1986 to April 15, 1987 for which the opacity at 225 GHz at the Langmuir Laboratory site has a value less than a given amount. Observations were made during about 79 percent of the time. The quartile values are 0.08, 0.13, and 0.23, as shown. For reference, these opacities correspond to 1.4, 2.2, and 4.1 mm of precipitable water vapor in the zenith path through the atmosphere.

Another way of looking at the same data is to ask how long the periods of very favorable atmospheric conditions persist. Typical 'full' synthesis observations would last for four hours with the MMA. In order to fit at least one such observation into the schedule, a somewhat longer period of 'good' weather is necessary. In Table VI-2 we have listed all the periods during the observations of Figure VI-6 when the opacity stayed at 0.1 or less for at least six hours. These

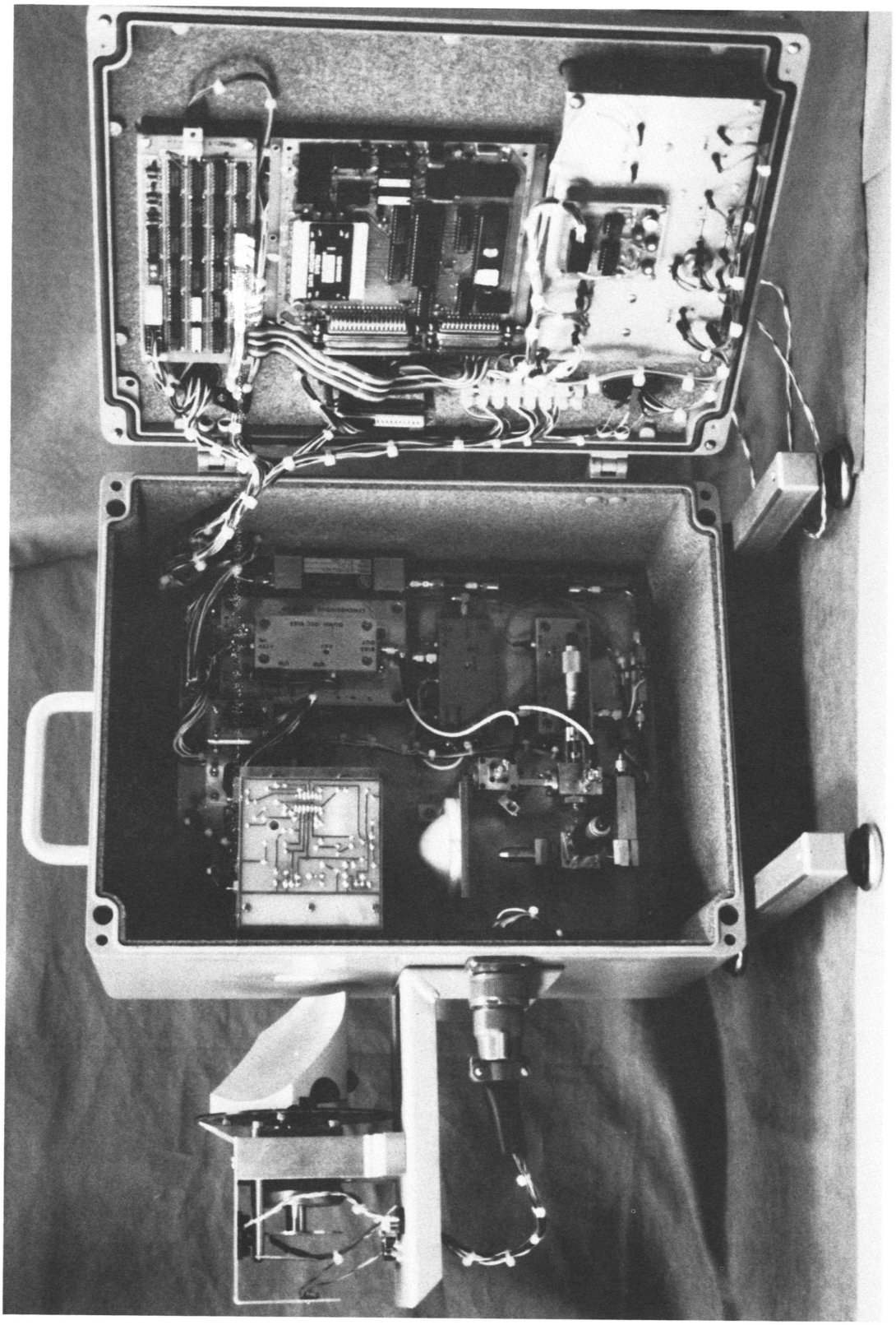


Figure VI-5. Photograph of one of the 225-GHz tipping radiometers used for atmospheric tests.

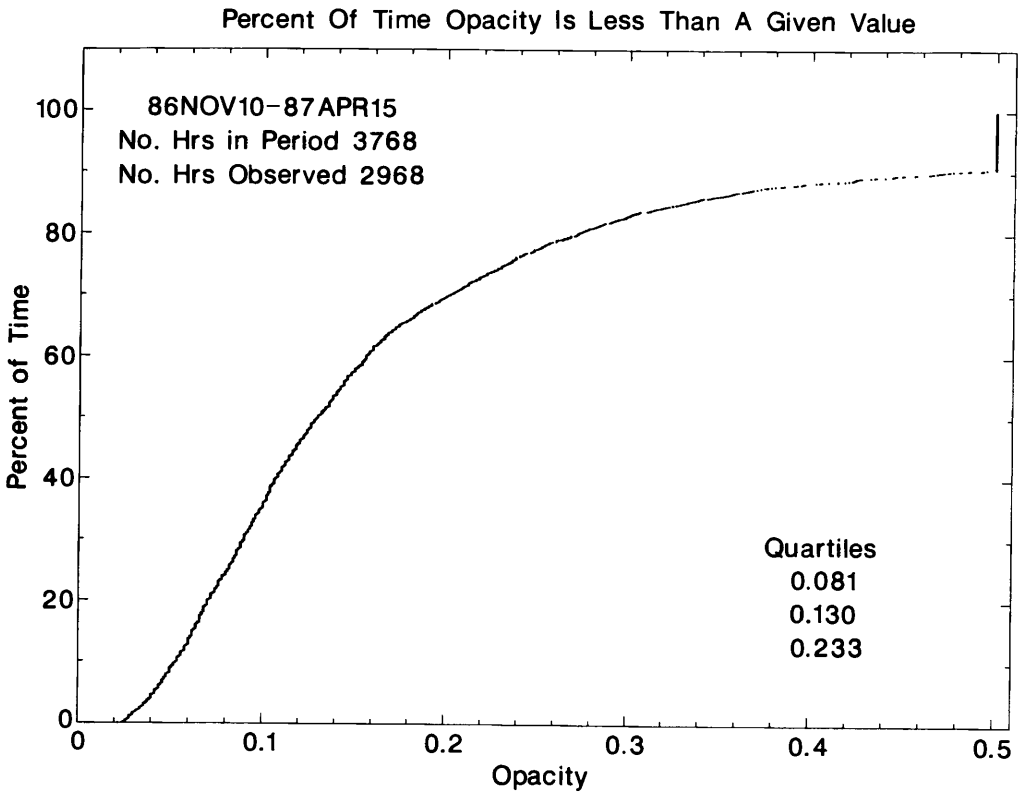


Figure VI-6. An example of data obtained at the Magdalena Mountains site by one of the three 225-GHz tipping radiometers which are in use for atmospheric tests of potential MMA sites.

are the periods during which the most-demanding observations could be carried out.

Besides tipping-scans to measure the opacity every ten minutes, the radiometer is also used to estimate the atmospheric stability. This is accomplished by observing the sky-temperature fluctuations at the zenith for an hour at a time. Water (both vapor and liquid) over a radio telescope produces both sky-brightness and path-length variations, which are simply related. In the absence of water droplets, one should be able to derive the path-length variations overhead by using the tippers to measure the spectrum of the observed temperature fluctuations. From this measurement, together with a wind velocity, one should be able to infer the phase fluctuations an interferometer would see.

In November and December 1988, we tested this concept by comparing 225-GHz sky-temperature fluctuations at the VLA site with the VLA phase stability at 2 cm. The atmosphere is essentially non-dispersive at any wavelength with an opacity low enough to make it astronomically useful. Thus the 2-cm VLA stability should be directly related to millimeter-wavelength phase stability. In Figure VI-7 we show results of these measurements. The plot compares the temperature fluctuations of the sky at 225 GHz, as seen by the tipper, with phase fluctuations measured at the VLA at a wavelength of 2 cm using an

Table VI-2. Periods of low opacity at the Magdalena Mountains site. These data were collected over the period November 10, 1986 through April 15, 1987. A 'good period' has opacity less than 0.10 for at least six hours; it can encompass time spans of up to three hours when the opacity is higher; however, the worst opacity must be less than 0.15 during a 'good period'. In all, there were 40 'good periods' during this time. The total number of hours involved in these periods is 980. An average 'good period' lasts 24 hours and has a median opacity of 0.069.

Run	Start	Stop	Duration	— Opacity —		
	y/ m/ d/hr	y/ m/ d/hr	(hours)	Median	Average	Minimum
1	86/11/10/11	86/11/12/ 4	42	0.068	0.071	0.038
2	86/11/23/22	86/11/24/ 4	7	0.072	0.075	0.062
3	86/11/27/ 4	86/11/28/21	42	0.076	0.074	0.037
4	86/11/30/ 1	86/11/30/13	13	0.067	0.068	0.036
5	86/11/30/20	86/12/ 2/19	48	0.060	0.062	0.037
6	86/12/ 3/ 9	86/12/ 3/20	12	0.090	0.089	0.079
7	86/12/11/22	86/12/15/17	92	0.055	0.059	0.025
8	86/12/19/ 1	86/12/19/ 6	6	0.083	0.086	0.079
9	86/12/23/ 9	86/12/23/14	6	0.046	0.046	0.039
10	86/12/26/ 2	86/12/27/ 3	26	0.064	0.068	0.044
11	86/12/29/11	86/12/31/21	59	0.055	0.054	0.028
12	87/ 1/ 2/12	87/ 1/ 3/ 7	20	0.070	0.068	0.031
13	87/ 1/ 4/ 8	87/ 1/ 4/16	9	0.090	0.091	0.086
14	87/ 1/ 8/21	87/ 1/ 9/ 7	11	0.053	0.054	0.035
15	87/ 1/20/18	87/ 1/23/ 5	60	0.080	0.075	0.036
16	87/ 1/24/11	87/ 1/24/20	10	0.079	0.079	0.066
17	87/ 1/25/15	87/ 1/27/ 8	42	0.067	0.069	0.050
18	87/ 2/ 1/ 5	87/ 2/ 1/14	10	0.088	0.087	0.078
19	87/ 2/ 2/ 2	87/ 2/ 3/ 1	24	0.080	0.078	0.058
20	87/ 2/ 3/ 6	87/ 2/ 3/11	6	0.098	0.094	0.080
21	87/ 2/ 5/16	87/ 2/ 6/ 7	16	0.052	0.053	0.044
22	87/ 2/21/15	87/ 2/23/ 5	39	0.047	0.053	0.034
23	87/ 2/26/21	87/ 2/28/ 3	31	0.060	0.061	0.035
24	87/ 2/28/ 7	87/ 3/ 1/21	39	0.082	0.076	0.034
25	87/ 3/ 3/ 9	87/ 3/ 6/10	74	0.042	0.049	0.024
26	87/ 3/ 6/13	87/ 3/ 6/21	9	0.054	0.060	0.044
27	87/ 3/11/ 2	87/ 3/11/12	11	0.076	0.078	0.059
28	87/ 3/12/ 2	87/ 3/12/16	15	0.076	0.080	0.060
29	87/ 3/13/ 3	87/ 3/13/ 8	6	0.090	0.090	0.082
30	87/ 3/13/20	87/ 3/14/12	17	0.064	0.070	0.050
31	87/ 3/15/ 5	87/ 3/15/10	6	0.082	0.084	0.072
32	87/ 3/17/22	87/ 3/18/11	14	0.085	0.085	0.065
33	87/ 3/18/21	87/ 3/19/12	16	0.075	0.078	0.064
34	87/ 3/23/ 9	87/ 3/23/16	8	0.049	0.051	0.036
35	87/ 3/26/19	87/ 3/27/ 7	13	0.068	0.065	0.022
36	87/ 3/28/ 6	87/ 3/28/11	6	0.061	0.068	0.054
37	87/ 3/30/ 9	87/ 3/31/17	33	0.069	0.067	0.040
38	87/ 4/ 2/16	87/ 4/ 3/12	21	0.060	0.059	0.026
39	87/ 4/12/20	87/ 4/14/15	44	0.054	0.056	0.027
40	87/ 4/14/21	87/ 4/15/13	17	0.068	0.071	0.053

effective antenna separation of 10 km. The solid line is the *theoretically predicted* relationship between temperature fluctuations and the VLA phase fluctuations for two antennas with uncorrelated atmosphere overhead. The turndown at $7^{\circ}.5$ in the plot is due to the expected contribution of instabilities in the VLA local-

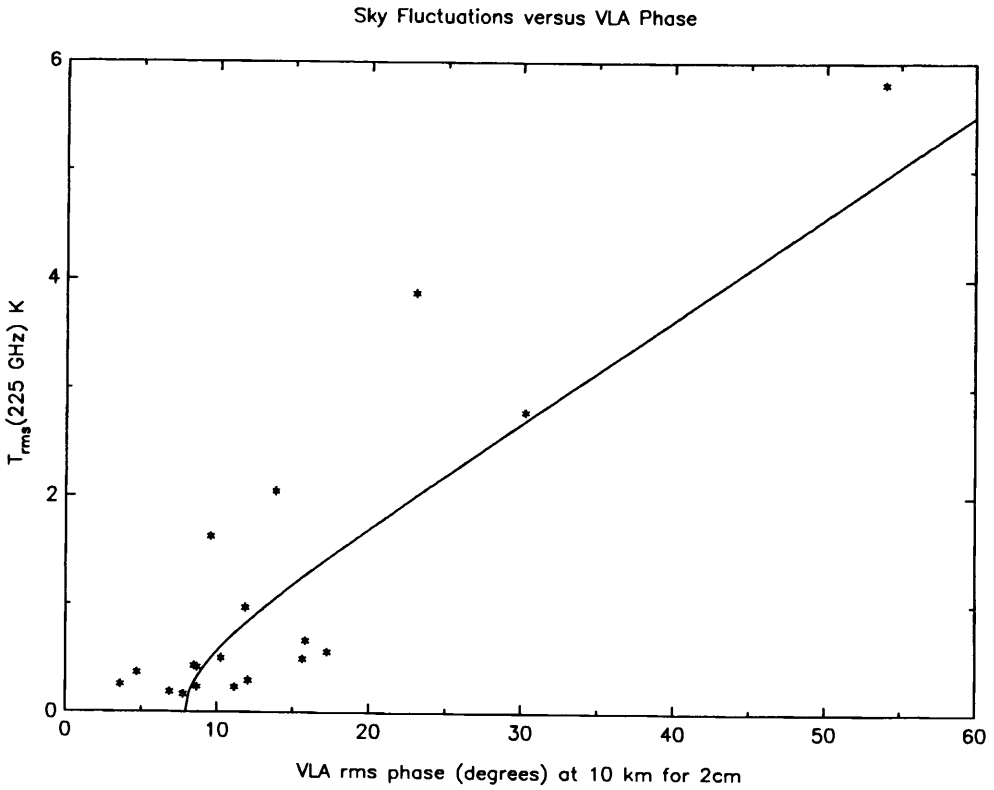


Figure VI-7. 225-GHz sky-temperature fluctuations, observed by a tipping radiometer, plotted against VLA phase fluctuations observed at 2-cm wavelength over a 10-km baseline. The solid line represents the theoretical prediction of the relationship that should exist.

oscillator system. The important point here is that the curve shown is not a fit but (except for the turndown at 7°5) is entirely specified by the theory. Thus it seems likely that we are measuring two closely related quantities.

Since the summer of 1989, once every five hours the tippers interrupt their opacity scans to make one-hour-long observations of the atmospheric stability. These observations suggest that both Mauna Kea and the Langmuir Laboratory site regularly have very stable conditions. The best observed conditions at Langmuir (20 mK rms) correspond to about four degrees of phase at 230 GHz on a 3-km baseline on time scales of four minutes.

3. Future Plans

In the future we will:

(1) Finish studying and testing the three candidate sites;
(2) Choose one of them as the prime site; and
(3) Obtain permission to begin construction.

Starting in the spring or summer of 1990, we will begin testing of the Springerville site. These tests will require us to run the 225-GHz tipping radiometer on a remote site without power, phone service, or existing buildings. For this purpose we have designed a solar and wind system to power the tipper and a

radio-telephone system for communications. Procurement of the hardware is now underway. We will begin setting up the system as soon as we are past the snow season this spring.

Further studies of the final antenna configurations and general site layout will also take place between now and the summer of 1991. We expect to be in a position to make a final choice of the primary site in the summer of 1991.

After the decision is made, we will begin the process of approval by the Forest Service for our choice. This process will involve an environmental impact statement and public hearings. It should take one-and-one-half to two-and-one-half years. Thus, under our plan, we should have full approval for construction by 1994.

VII. ANTENNAS

1. General Antenna Specifications

The general antenna specifications call for a transportable antenna that provides precision operation at frequencies at least as high as 350 GHz; its optical design must also support efficient operation in the 9-mm band. The antennas must support the demands of high-fidelity mosaiced imaging at all frequencies. They must also be reliable and easy to maintain.

2. Precision

Requirements for high-fidelity mosaiced imaging place tight restrictions on the pointing accuracies and beam patterns of the individual antennas. They also demand a high degree of antenna-to-antenna uniformity. In order to provide precision operation at the highest frequencies, the effective rms accuracy of the reflector surface—a combination of

(1) surface panel irregularities,
(2) panel setting errors,
(3) gravitational deformations,
(4) thermal deformations, and
(5) secondary and other reflector accuracies

—should be less than $\lambda 1.0$ mm/40 = 25 μm, corresponding to an aperture-efficiency reduction at the highest frequencies of no more than $\sim$ 10 percent. Further, the antenna pointing accuracy should be better than one-twentieth of the full 3-dB beam width, or about $1''\lambda_{(\mathrm{mm})}$. These specifications have been shown to be necessary in order to produce high-quality mosaiced images (see Chapter IV). They will also permit operation at higher frequencies with reduced, but still useful, efficiency.

The difficulties that may be expected in meeting these specifications are summarized in Figure VII-1. This plot of antenna diameter vs. rms surface accuracy illustrates the regions where gravitational deformations or thermal gradients across the structure will significantly impair the performance of the antenna. The point representing the proposed MMA antenna is just on the limit of the gravitational limit for a 'stiff' (non-homologous steel) structure and beyond both the typical daytime (ΔT = 5° C) and nighttime (ΔT = 1° C) temperature-differential limits for a structure made of steel. Considering gravitational limitations alone, 8 m is the largest 'stiff' (steel) antenna that one can build with a net error less than 25 μm. This consideration and the need to keep the antenna inexpensive dictate the choice of 8-m diameter.

The thermal limit is more difficult to deal with. Some antennas, such as the SMT and the SEST telescopes, rely on carbon-fiber-reinforced plastic (CFRP) for all or part of the backing structure to solve this problem. Others, such as the IRAM 30-m, have active thermal-stabilization schemes, while a third method is

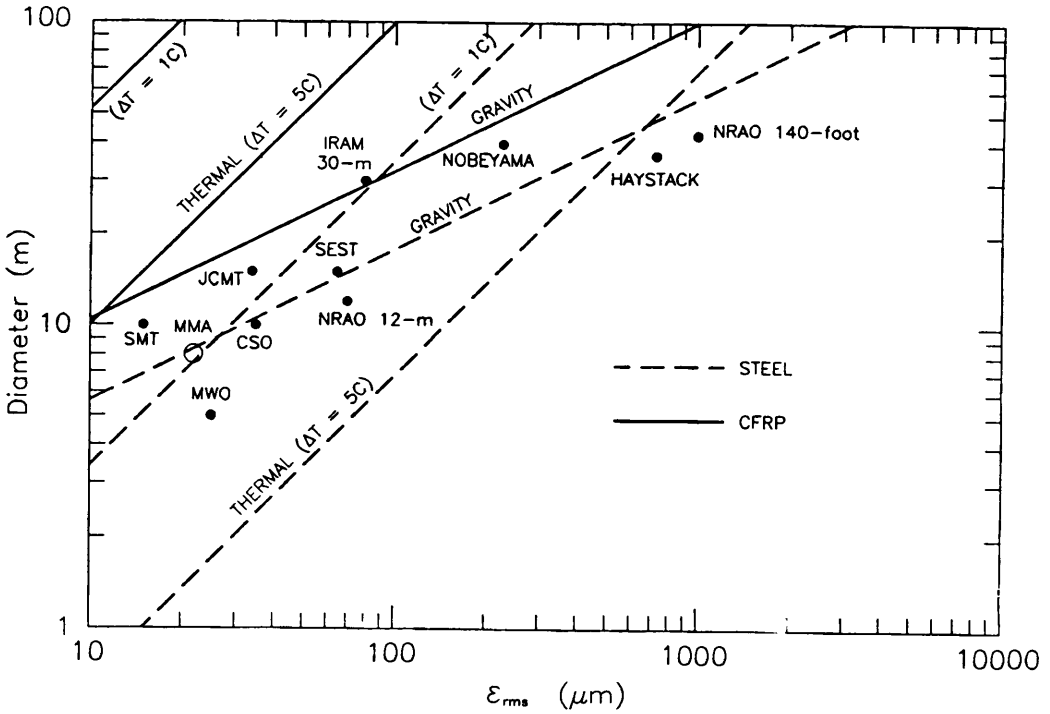


Figure VII-1. A plot of antenna diameter vs. rms surface error, showing physical limits to the precision that can be attained.

to use a protective enclosure (as at the JCMT and CSO). Note also that daytime operations of some of the telescopes are restricted by 'anomalous refraction', so that the daytime thermal properties may not have been well studied. It is probable that the MMA will be sited in a region where the topography does not produce this anomalous effect, and the antennas should therefore be designed for full precision during the daytime.

The precision requirements of the telescope may be met by using CFRP as the constructional material for the backing structure and panels. Its advantages are:

(1) Excellent thermal properties;
(2) Good mechanical properties, allowing use of a 'stiff' structure to meet the precision requirements;
(3) Light weight permits easy assembly of antennas;
(4) High resonant frequencies result in better servo performance;
(5) Capital-cost savings, since no enclosure system or thermal stabilization system is required;
(6) Operating costs are reduced since there is no maintenance of temperature-stabilization equipment or enclosure; and
(7) The design of the transporter is simplified, since the structure itself is lighter and there is no additional weight of an enclosure.

The disadvantages are:

(1) The material cost is higher than for steel (though this is offset somewhat, as described above); and

(2) There is little experience with CFRP antennas in the U.S., though there is some in Europe.

A primary concern for the MMA antennas is that they be mechanically simple. Reliability is an utmost concern with an array of forty antennas. The 'tricks' one might consider to maintain the figure of a single antenna—capacitive edge-sensors guiding actuators, differential thermal screws or strips, and even complicated heaters and air handling/circulation systems—impose too great a compromise between array reliability and operations expense. A simple, passive mechanical design is needed. As a working design for the MMA antennas, we have chosen to pursue the approach used for the backing structure of the IRAM 15-m antennas, a combination of CFRP (lateral) and carbon-steel (diagonal) members. The surface panels either can be CFRP bonded to aluminum honeycomb and coated with a conducting surface or they can be (probably smaller) machined aluminum panels. The latter approach, employed successfully with the new BIMA antennas, has the advantage of providing durable and simple panels, with the residual tooling marks being invisible at millimeter-wavelengths while at the same time scattering both visible and infrared light: the Sun can be observed with these panels.

Drawing on the cost experience of two existing precision telescopes, the IRAM 15-m antennas and the 10-m SMT, we consider individually the costs of the constituent parts of the structures, analyze these costs in terms of the technical specifications, and then scale to the needs of the MMA. The results are shown in Table VII-1. Note that these panel cost estimates are for the case of CFRP panels (as used in both the IRAM 15-m and the SMT). If we can use machined aluminum panels instead, the costs will be lower.

Table VII-1. Estimated Costs of CFRP MMA Antennas

Component	Formula†	Total (k$)
Mount	$250 + 10D^{3/2}p^{-1}$	475
CFRP Panels	$3.9D^2$	250
Backing Structure	$0.88D^3$	450
Tooling/Servo/Alignment		25
Site Assembly		50
TOTAL		1250

†D is the antenna diameter in meters; p is the pointing accuracy in arcseconds.

No attempt has been made here to incorporate savings that would accrue from procurement of forty antennas all at one time. This is a 'one-of' estimate, and as such the total of the antenna costs will be less than forty times 1.25 M$. Nevertheless, the antennas will remain the dominant component in the overall construction cost of the project. During the lifetime of the Millimeter Array, receivers and computers are likely to be replaced as newer technology

becomes available, but the antennas themselves are not likely to be replaced. This suggests that the basic antenna design may provide the ultimate limitation in performance of the MMA. It would be unfortunate to compromise at the start with too marginal a design.

3. Preliminary Design

A preliminary design we have developed shows the feasibility of satisfying our basic antenna requirements. Figure VII-2 shows a sketch of the antenna. The specifications are given in Table VII-2 and illustrated in Figure VII-3.

In order to meet the field-of-view specifications, the primary and secondary mirrors are not shaped for maximum gain but are a conventional paraboloid/hyperboloid combination. It is, in principle, possible to shape the beam close to the secondary focal plane to give a higher aperture efficiency, at least for the single on-axis beams (and possibly for off-axis beams). Some work on improved-efficiency optics is being carried out for receivers on the NRAO 12-Meter Telescope.

The receivers, which are located in a room within the base of the antenna, are stationary, so that access to them is relatively easy. The weight is not on the moving part of the antenna, so no rebalancing is required as receivers are removed or added. There are no cable or cryogenic lines which could cause drag on the antenna and potentially could limit the pointing accuracy. Additionally, these connections to the receivers are not subject to constant flexing which would make them more prone to failure.

An arrangement of plane mirrors brings the beam to the receivers, and rotation of one of the mirrors is sufficient to select any desired receiver (a dichroic reflector would allow two receivers to be used simultaneously). This system will transmit beams for receivers down to about 40 GHz with good efficiency, and to 30 GHz with reduced efficiency. Sufficient clearance is provided for off-axis imaging for future array receivers at 200 GHz and higher frequencies. The need for imaging with array feeds requires a hole in the primary which is larger than the secondary mirror (giving a geometric area blockage of 1.6% rather than the 1% for the secondary alone). It may be possible in the future to reduce the hole size (and possibly the secondary diameter) by using focusing mirrors in the optics to reduce the beam size at the hole. At present the performance of such a system for off-axis beams is not sufficiently well understood to incorporate it in a preliminary design. The central blockage for single-beam receivers could be reduced by having a ring which could fit in the central hole to reduce the clearance to the minimum.

3.1. Primary Mirror. It is assumed that the diameter D of the primary reflector will be 8.0 m. The primary will be a symmetric paraboloid with a focal length f of 3.2 m, yielding a primary f-ratio (f/D) of 0.4. This value gives a compact antenna. It could be reduced further (to $f/D = 0.35$, say), but the tolerances in locating the secondary then would be more stringent (both for pointing and for beam quality).

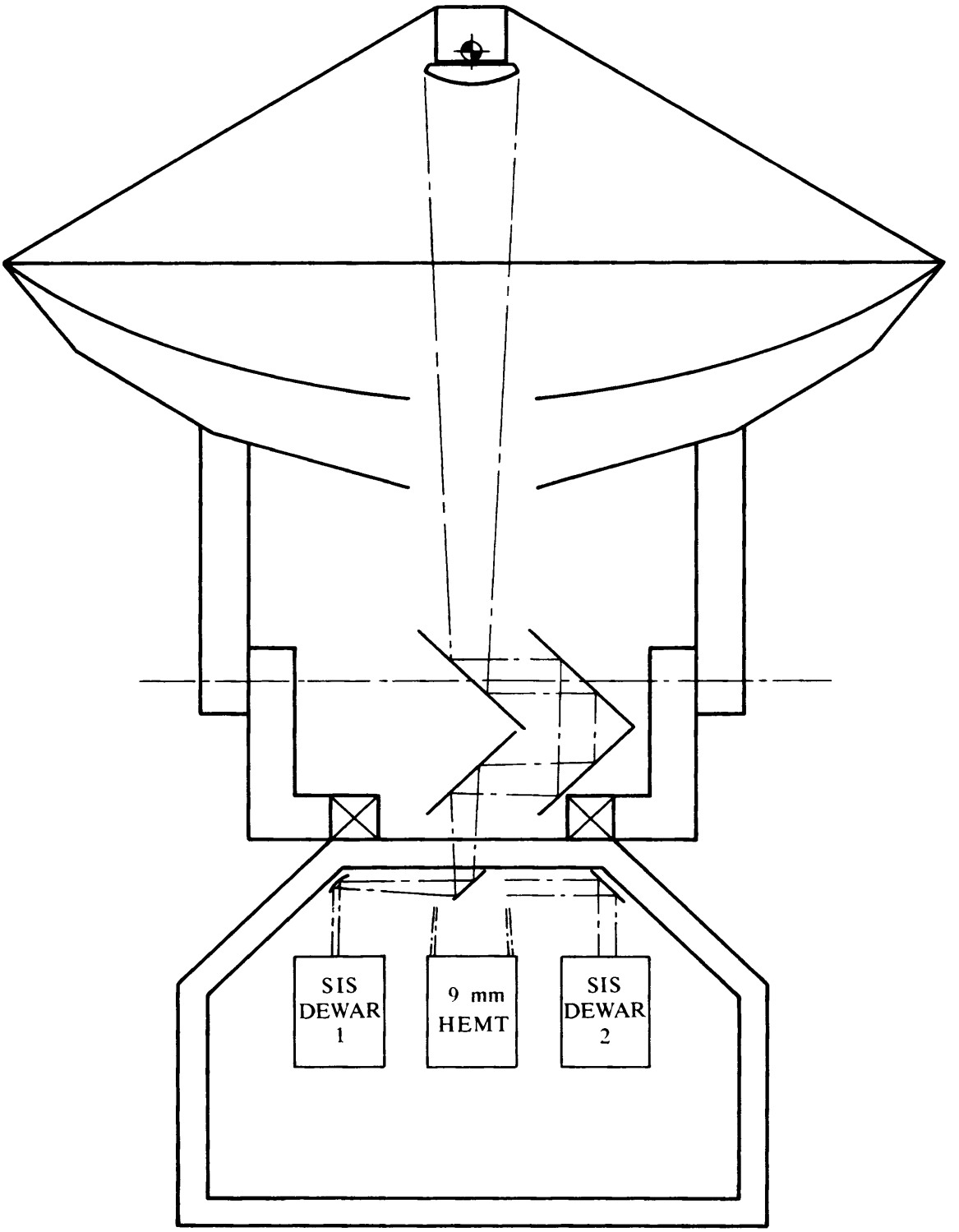


Figure VII-2. Sketch of the preliminary antenna design. The receivers are selected by rotating the central mirror.

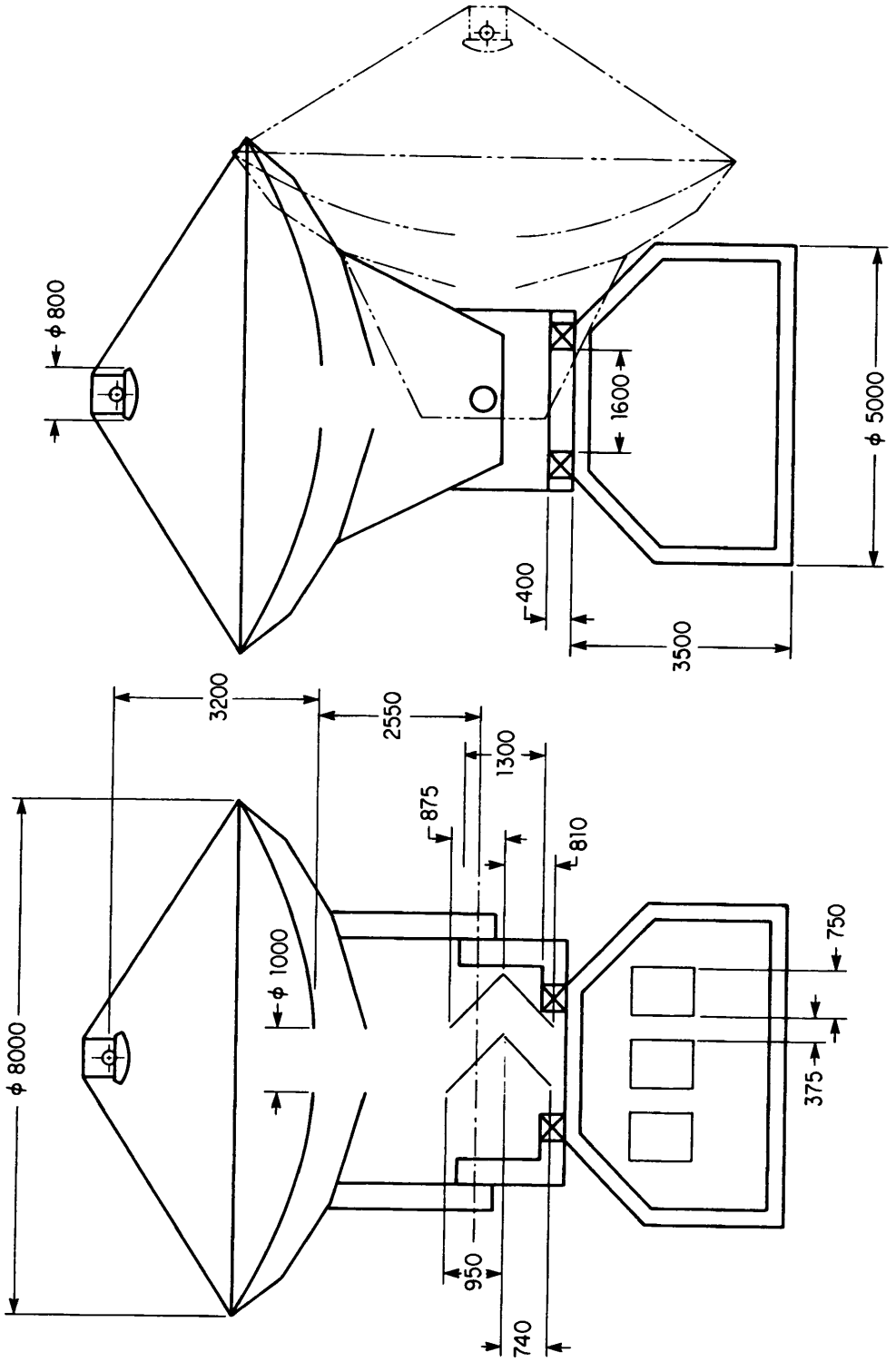


Figure VII-3. Sketch of an antenna, with dimensions illustrated (in units of millimeters).

Table VII-2. Specifications of Prototype Antenna

General		
Effective surface accuracy	ϵ	22 μm
Pointing accuracy	p	$1''$
Primary Mirror		
Shape		Paraboloidal
Diameter	D	8.00 m
Focal length	f	3.20 m
Focal ratio	f/D	0.4
Edge angle	θ_p	64°01
Diameter of central hole	d_h	1.00 m
Secondary Mirror		
Shape		Hyperboloidal
Diameter	d_s	0.8 m
Distance between foci	f_s	9.7908 m
Additional path length	$2a$	9.18409 m
Eccentricity	e	1.06897
Magnification	M	30
Paraxial focal length	$f_a = a(e^2 - 1)/2$	0.655 m
Equivalent Paraboloid		
Focal length	$F = Mf$	96.0 m
Focal ratio	F/D	12.0
Edge angle	$\theta_s = 2\cot^{-1}(4F/D)$	2°39
Plate scale	$1/F$	$2''.15$ mm^{-1}

3.2. Secondary Mirror. The secondary mirror will be a segment of a symmetric hyperboloid. Its size is set principally by the field of view and the requirement to have the secondary focus in the receiver room. It should be as small as possible, within these constraints, in order to minimize blockage and to allow for the possibility of nutation for beam switching at some stage.

A diameter of 0.8 m and a distance between foci of 9.79 m gives an effective f-ratio at the secondary focus of 12, which is convenient for most quasi-optical systems. With some more optimization, the secondary diameter could be reduced but probably not to much less than 0.75 m. Similarly, the focal ratio could be changed, but it is fairly tightly constrained and will probably be in the range 11 to 13.

3.3. Tertiary Mirrors. Four mirrors are used to remove the rotations produced by azimuth and elevation movement and to produce a constant beam in the receiver cabin. One is fixed to the primary-mirror structure, and the other three to the rotating part of the azimuth bearing. As now envisaged, these are all planar mirrors, for the following reasons:

(1) They are cheaper to fabricate than shaped reflectors;
(2) They are simple to align;

(3) The imaging properties of shaped mirrors for off-axis beams are unknown.

The sizes of these mirrors vary from about 1.4 m × 0.94 m to 1.0 m × 0.74 m.

3.4. Receiver Room. The receiver room is located in the base of the antenna mount. Five receiver bays are planned (see Figs. VII-2 and VII-3). Of these, only three will be used initially; the other two can be conveniently provided and will allow for possible expansion or for installation of special-purpose, perhaps user-supplied, instrumentation. One receiver bay, which is at the center of the room and is intended for the 9-mm HEMT receiver, receives its signal directly through the azimuth bearing. For the other four bays, the beam is deflected by the selection mirror to one of the focusing mirrors. For the second and third bays, we plan two 4-K dewars containing receivers for two frequency bands each. If the central mirror is dichroic, then both receivers can be used simultaneously.

Note that the image of the sky rotates relative to a receiver's position, so that either the receivers have to be rotated or software corrections have to be applied. The electronics section, Chapter VIII, discusses the receiver configuration in more detail.

3.5. Beam-Switching Optics. Each of the MMA antennas must be capable of recording total-power measurements which can be merged with the cross-correlation information to provide a complete image of the sky brightness. To this end, each antenna must have an optical system that allows rapid total-power differences to be taken at adjacent sky positions; this can be achieved either by nutation of the secondary mirror or by means of a chopping tertiary mirror. The cost vs. performance tradeoffs will be evaluated as part of the antenna design.

4. Other Design Considerations

Initial considerations show that it will be possible to construct antennas meeting the required specifications, but many details still need to be determined. It has been assumed that the antennas will operate without enclosures, mainly so that they can be brought together to form the close-packed array, but also to reduce the weight that needs to be transported.

Clearly, the precision of the surface can be maintained without an enclosure, but it is possible that survival may dictate some kind of protection. This may be simply rear cladding, as used on the IRAM 15-m antennas to prevent the build up of ice loads on the antennas. Such questions can be better answered once the site is chosen and well characterized.

The question of how to achieve the required pointing accuracy will receive emphasis in the next three years of the MMA design phase, in the restricted sense it which it applies to the MMA. In this regard, we expect to profit from experience gained in the design of the Green Bank Telescope, experience with the IRAM 30-m telescope, and the experiments with optical pointing systems at the NRAO 12-Meter Telescope and at BIMA. Our plans are described in further detail in Appendix A.

Panel design will be critically important with regard to the overall antenna precision. CFRP panels may be a good choice, but there is then the problem of

applying a good reflective surface. A thin aluminized layer may be subject to significant corrosion if used in the open, and a dielectric protective layer (such as used on the IRAM 15-m antennas) may be difficult to produce for very-high-frequency operation. Machined aluminum panels on a CFRP backing structure would have a distinct advantage.

Methods of transporting the antennas still need to be considered. A specialized vehicle operating on a road surface rather than on rails appears to be a very feasible solution.

5. Conclusions

A design has been found which will meet the specifications for the antennas of the MMA. This design is intended to be as simple as possible, to minimize maintenance and initial setup requirements. Several areas, such as the panel design, need more study to arrive at a satisfactory solution. Other areas have options which must be studied in more detail, such as minimizing central blockage by using a focusing beam-waveguide system. Future investigations will emphasize detailed structural calculations of the effects of gravity, temperature, and wind-loading on the surface and pointing accuracies.

VIII. ELECTRONICS

1. Introduction

Much practical experience in the design of electronics for the Millimeter Array has accrued from the design of the VLA, the VLBA, single-antenna millimeter-wave receivers, and the millimeter-wave interferometers, particularly those at Owens Valley, Hat Creek, Nobeyama, and the IRAM facility in France. Much of the electronics is already feasible, but it is clear that extensive development work will have to be carried out over the next few years to achieve the high standards of performance and reliability that the MMA requires.

In this chapter the main electronics areas are discussed, with particular attention to the key parameters that must match the astronomical requirements, and with cost estimates for construction and operation. The areas discussed are receivers, the local-oscillator system, data transmission, and the correlator.

2. Receivers

2.1. General. The four frequency bands given greatest emphasis at the 1985, 1987, and 1989 MMA Science Workshops and Advisory Committee Meetings are centered in the atmospheric windows at 9, 3, 1.3, and 0.85 mm. In each band a large tuning range is desirable, so that the receiver characteristics do not impose artificial limitations on the science. Ideally, one would like to be able to observe anywhere within the entire frequency range of each of the atmospheric windows, but, especially in the broader windows, this is likely to require multiple receivers within each band. Note that, even with a single receiver for each atmospheric window, there would be a very large number (320) of receivers in the MMA (4 bands $\times$ 2 polarizations $\times$ 40 antennas).

It is clear that only SIS receivers will be able to achieve the sensitivity and bandwidth required for the MMA in all but the 9-mm band, where HEMT amplifiers can achieve receiver noise temperatures $\lesssim$ 50 K. About ten observatories worldwide routinely use SIS receivers from 40 GHz to 360 GHz. However, despite their widespread use, many SIS receivers are little or no more sensitive over their full tuning ranges than the best Schottky mixer receivers, and very few have approached the ultimate sensitivity limit imposed by the uncertainty principle, even allowing for the inevitable losses in input optics. The SIS mixers now in use tend not to be widely tunable, and their dynamic ranges can be a limitation even in astronomical applications. We believe these present shortcomings are a reflection of the limited engineering effort that has gone into SIS mixer design, and the lack of facilities for fabricating high-quality SIS junctions and multi-level superconducting integrated circuits. Most SIS mixers still are modifications of whisker-contacted Schottky-diode mixer designs, and shadow-mask Pb-alloy junctions remain widely used despite the existence of superior

technology. Recent developments at NRAO[1] in collaboration with IBM[2] and Hypres[3] suggest that with sufficient engineering development receivers suitable for the MMA will indeed be possible within five years.

2.2. Single-Sideband vs. Double-Sideband Receivers. SIS mixers appear best-suited to operation with the upper and lower sidebands identically terminated—i.e., in DSB mode. There are three ways to operate DSB SIS receivers on the MMA: (i) with equal response to both sidebands; (ii) with an image-separation mixer—this has separate output ports for the upper and lower sidebands; and (iii) with a sideband diplexer to terminate the image in a cold load.

With simple DSB receivers on an interferometer it is possible to separate the upper and lower sidebands using Walsh-function phase-switching of the local oscillators. This is the approach adopted for the MMA; the sidebands will be separated at the correlator, providing the astronomer access to both sidebands. However, atmospheric noise, which is not correlated between different antennas of the array, degrades the sensitivity of observations in both sidebands. This can be overcome if image-separation mixers are used on each antenna. Image-separation mixers consist essentially of two parallel DSB mixers with their local oscillators phased in quadrature. The upper and lower sideband components at the IF are separated by combining, with appropriate phasing, the outputs of the two mixers. Because of the complexity of the latter approach, we consider it not as a viable option for the first round of MMA receivers, but as a future option if the considerable engineering development—probably of an integrated image-separating SIS mixer—can be supported.

The third receiver configuration—using a sideband diplexer to terminate the image sideband in a cold load—also prevents degradation of the array sensitivity by atmospheric noise in the image sideband. In practice, however, the loss in the diplexer and the finite temperature of the cold termination are likely to reduce the benefit of this type of image rejection. An example discussed at the 1989 MMA Workshop showed that the overall SSB system noise temperature (including the atmosphere) of a simple DSB receiver at 230 GHz might be improved by only $\sim$ 3% by inserting an image diplexer and a cold image termination. (Cooling the diplexer could increase the improvement to $\sim$ 20%.)

We believe that the additional complexity of sideband separation in the individual receivers is not justified for the initial MMA receiver complement. Work on image-separating mixers may eventually provide a compact solution to this problem with little or no degradation of receiver performance and without requiring additional mechanical tuners to adjust quasi-optical diplexers. Therefore, the

[1] A. R. Kerr and S.-K. Pan, "Some recent developments in SIS receiver technology", presented at the *First International Symposium on Space Terahertz Technology*, March 1990.

[2] S.-K. Pan, A. R. Kerr, M. J. Feldman, A. Kleinsasser, J. Stasiak, R. L. Sandstrom, and W. J. Gallagher, "A 85–116 GHz SIS receiver using inductively shunted edge-junctions", *IEEE Trans. Microwave Theory Tech.*, **MTT-37** (1989) 580–592.

[3] A. R. Kerr, S.-K. Pan, S. Whiteley, M. Radparvar, and S. Faris, "A fully integrated SIS mixer for 75–110 GHz", *IEEE International Microwave Symposium, Digest of Technical Papers*, May 1990, to appear.

MMA will use DSB receivers and will use Walsh-function phase-switching of the LO at the correlator to separate the sidebands.

2.3. Polarization. Virtually all millimeter-wave receivers are inherently linearly polarized. Waveguide circular-polarizers, commonly used at centimeter wavelengths, are not practical for the MMA receivers because of their extremely small size and inherently narrow bandwidths. Although conversion of a dual linearly polarized receiver to dual circular polarization can be done quasi-optically (e.g., using a Martin–Puplett polarizing interferometer), the instantaneous bandwidth is limited, the loss in the diplexer is non-negligible, and mechanical tuning is required.

We propose initially to use linearly polarized receivers on the MMA and later to consider adding removable quasi-optical circular polarizers at selected frequencies if this option is recommended with high priority by the MMA Scientific Steering Committee.

2.4. Local Oscillators. It is likely that developing widely-tunable local oscillators (LOs) for the MMA will be almost as large a task as developing the SIS mixers. At present, the LO circuit in a typical millimeter-wavelength radio astronomy receiver consists of a phase-locked Gunn oscillator driving a frequency multiplier. Such multipliers, using whisker-contacted Schottky diodes, would probably be the least reliable components in the whole of the MMA electronics. A Gunn oscillator/multiplier combination may have as many as five mechanical tuners, and its tuning range is limited to $\sim$ 10% by the oscillator. As the desired frequency bands for the MMA have widths as large as 36%, it is clear that to build the MMA today would require multiple LOs to cover each band.

The LO power required by an SIS receiver is discussed in an NRAO internal report.[4] Depending on the RF bandwidth and the required dynamic range, up to 1 μW may be required by each mixer. Allowing for loss in the power splitter (to drive the two oppositely polarized mixers), LO-leveler loss (SIS receivers need leveled LOs), and loss in the LO diplexer and other components, a safe margin should be assured with 100 μW to 1 mW available from the source, depending on the type of LO diplexer used.

Some recent developments may ultimately lead to better solutions to the LO problem:

(1) *RTD Oscillators.* Recent experiments with Resonant-Tunneling-Diode (quantum-well) oscillators have given 20 μW at 370 GHz,[5] and future results are sure to exceed this. A tuning range of 10% to 20% is possible by adjusting the bias voltage. Inherent suppression of shot-noise gives noise levels comparable to those of a Gunn oscillator.

[4] A. R. Kerr, M. J. Feldman, and S.-K. Pan, "SIS mixer design by frequency scaling", Electronics Division Internal Report No. 267, NRAO, April 1987.

[5] E. R. Brown, "Submillimeter-wave resonant-tunneling oscillators", presented at the *First International Symposium on Space Terahertz Technology*, March 1990.

(2) *TUNNETT Oscillators.* Tunnel-Transit-Time Oscillators are expected to give 200 mW at 100 GHz and 20 mW at 500 GHz.[6]

(3) *Transistor (HEMT) Oscillators.* These may ultimately compete with Gunn oscillators up to $\sim$ 100 GHz, and they could have a wider electronic tuning range.[7]

(4) *GaAs Varactor Multipliers.* Recent work on whiskerless varactor diodes and monolithic integrated varactor multipliers may result in multipliers without mechanical tuners but with substantial bandwidth.[8]

(5) *Quantum-Well Multipliers.* Quantum-well double-barrier varactor or varistor diode multipliers are potentially good odd-order multipliers. At 200 GHz, 240 μW has been measured using such a tripler.[9] A tuning range of 10% to 20% is anticipated.

(6) *Oscillator Grids and Multiplier Grids.* The use of planar grids of oscillators or multipliers can be an effective method of power-combining. A multiplier array has produced 1 W at 100 GHz.[10]

(7) *Josephson Oscillators.* Large, series arrays of Nb/Al–Al_2O_3/Nb Josephson junctions have delivered 7 μW into 50 Ω at 400 GHz and are tunable over more than two octaves.[11] Powers of 100 μW are expected in the future. These oscillators must operate at $\sim$ 4.2 K.

2.5. Receiver Configuration. The receivers will be mounted in a fixed room in the pedestal of the antenna. As mentioned earlier (in Ch. VII) frequency selection will be by a rotating mirror located in the center of the room, under a ceiling opening.

The receivers for the frequency bands proposed will be installed in three dewar packages. There will be adequate space for the addition of future receivers.

The three packages will be as follows:

Package	Frequency Range	Physical Temperature
1	30–50 GHz	20 K
2	68–115 GHz	4 K
	130–183 GHz	4 K
3	195–314 GHz	4 K
	330–366 GHz	4 K

[6]G. Haddad, J. East, C. C. Chen, K. Hashim, "Tunnel transit-time (TUNNETT) devices for terahertz sources", presented at the *First International Symposium on Space Terahertz Technology*, March 1990.

[7]M. Pospieszalski, NRAO, private communication.

[8]R. J. Mattauch, W. L. Bishop, and A. W. Lichtenberger, "Recent results on: surface-channel Schottky, InGaAs Schottky, and Nb based SIS mixer element research", presented at the *First International Symposium on Space Terahertz Technology*, March 1990.

[9]M. A. Frerking, "Quantum well multipliers: triplers and quintuplers", presented at the *First International Symposium on Space Terahertz Technology*, March 1990.

[10]D. B. Rutledge, Z. B. Popovic, R. M. Weikle, M. Kim, K. A. Potter, R. C. Compton, and R. A. York, "Quasi-optical power-combining arrays", in *IEEE International Microwave Symposium, Digest of Technical Papers*, May 1990, to appear.

[11]J. Lukens, State University of New York, private communication, Sept. 1989.

Package 1: This will be a dual-channel (linearly polarized) receiver using two 30–50 GHz HEMT amplifiers. This receiver will probably be very similar to the VLBA 43-GHz receiver that is currently being prototyped. *Packages 2 and 3:* These receivers will probably be similar to the closed-cycle-refrigerator receiver currently being commissioned on the NRAO 12-Meter Telescope. There will be four inserts in each of the receiver dewars, each insert covering two frequencies and two linear polarizations.

The proposed grouping is desirable so that the same quasi-optical diplexer for LO injection can be used for the paired frequency bands.

2.6. Dual-Frequency Operation. Dual-frequency operation would be possible using a dichroic reflector plate. All possible frequency pairs might not be provided, however, because of restrictions imposed by the optics.

2.7. Cryogenics. We propose using cryo-cooler systems similar to those which have been used by various organizations, including the Jet Propulsion Laboratory, the NRAO, and Caltech. The latest version in use at the NRAO incorporates a CTI Model 1020 refrigerator, and the heat exchangers are sized for a 1-watt capacity at 4 K. A separate compressor for the Joule–Thompson circuit is used to aid in installation on the antenna and also to help minimize the possibility of contamination, since the CTI refrigerator runs on a completely separate gas circuit.

2.8. Receiver Development. While all the individual components of the Millimeter Array could, in principle, be constructed using present-day technology, it is clear that substantial technical development will be needed in several areas before we are in a position to build a usable instrument of the scale and versatility proposed here. In the years *preceding* the construction of the MMA, work should commence in the following areas:

(1) *SIS Mixers:* Designs for reproducible SIS mixers should be developed for each of the frequency bands. In particular, wide-band tunerless designs should be investigated to determine how many receivers will be needed to cover each atmospheric window. Crucial to this work is access to one or more facilities able to fabricate superconducting microcircuits.

(2) *Wideband Local Oscillators:* While existing Gunn oscillators and whisker-contacted Schottky-diode frequency multipliers can deliver the local-oscillator powers needed, the tuning ranges of the Gunns and the reliability and reproducibility of the varactor multipliers need improvement for MMA receivers. It is clear that resources should be allocated to development, at NRAO or elsewhere, of suitable LOs. NRAO should support the development of planar (whiskerless) varactor-diode multipliers and, when appropriate, one of the alternative LO sources mentioned above.

(3) *Reliable 4-K Cryogenic Refrigerators:* In view of the large number of 4-K cryostats on the MMA, it is important to have proven refrigeration systems that can be economically maintained. The new

Joule–Thompson refrigerators now being installed on the 12-Meter Telescope should be regarded as prototypes of MMA refrigerators, and their design should be improved and refined as much as possible over the next few years.

3. Signal Transmission

The Millimeter Array will require the transmission of wideband (2-GHz) receiver outputs, local-oscillator reference signals, and control and monitor signals. The key parameters are the bandwidths for the receiver outputs and the phase stability for the LO signals, along with the lengths of the transmission paths.

Optical-fiber links are the most likely approach for all signal transmission. Possible alternatives are free-space radio (with severe bandwidth and interference problems), TE_{01}-mode circular waveguide (used in the VLA but no longer commercially available), and coaxial cable (high loss and restricted bandwidth, relative to fibers). We consider each of the requirements separately.

For the LO reference, we would like to achieve a phase stability of 0.3 radian rms at 345 GHz, which corresponds to a time stability of 0.23 psec. Over the maximum distance of 3 km (10 μsec in free space, 20–30 μsec in cable or fiber), variations several thousand times this limit will occur, because of temperature variations of a few degrees. In addition, flexing of the transmission line with antenna motion will cause changes in the effective length. It is therefore clear that a round-trip-correction system will be needed and that precise reciprocity must be maintained. Such systems employing UHF or microwave carriers on coaxial cable are highly developed, and the requirement is within the state of the art. Thus, we know that a stable LO system can be built for the MMA. On the other hand, it would be a great convenience and a significant cost saving if the LO reference could be sent on a single optical fiber of a multi-fiber cable. Until recently, fibers had higher temperature coefficients of effective length than cables and possessed an intrinsic non-reciprocity. But the technology has been developing rapidly, and new results indicate that a sufficiently stable fiber-based system may now be feasible.[12,13,14] We will continue to monitor these developments before making the final choice, but the cost estimates below assume the use of optical fibers.

Transmission of the receiver output signals will be easy to accomplish with optical fibers. Transmitters and receivers with more than 6-GHz bandwidth are available, and single-mode fibers can support even greater bandwidth. Capacity substantially exceeding our requirement of 2 GHz per antenna can be installed at negligible extra cost, providing for future expansion. An important design choice is whether to digitize the signals before or after transmission; i.e.,

[12] J. Coe, "Interferometer analog optical links", NRAO Electronics Div. Tech. Note No. 149, Jan. 1989.

[13] N. Sarma, "Optical fibre link for reference frequency distribution", Australia Telescope Memo AT/23.2/020, Feb. 1988.

[14] L. Primas *et al.*, "Stabilized fiber optic frequency distribution system", Jet Propulsion Laboratory preprint, Aug. 1989.

whether to transmit the signals in digital or analog form. Analog transmission is somewhat simpler, and it allows some of the tuning and filtering to be done in the central building, where it is more conveniently accomplished. Digital transmission avoids the nonlinearities of modulators and demodulators, avoids any effects of delay variations in the fibers, and facilitates multiplexing channels onto each fiber (the bandwidth will be divided into at least four, and probably eight channels, for reasons explained in Sec. 4 below). Both methods are feasible, and a final choice will depend on cost/performance tradeoffs that can be determined only after detailed design work.

Data transmission requirements for monitor and control tasks are extremely modest compared with those just discussed. Digital links of 10 kbaud or less in each direction will suffice, and can easily be provided on a multi-mode fiber in the same cable as the other signals.

A serious problem with the type of single-mode fiber that supports wideband transmission is the lack of suitably rugged connectors designed for it. To reconfigure the array, antennas will regularly have to be disconnected, moved, and reconnected. The connectors will generally be outdoors, at least those at unoccupied stations. Connectors for low-loss coupling of single-mode fibers must be very precisely manufactured, and those that are now available are designed for benign environments. This situation has caused the Australia Telescope to select multi-mode fiber, for which military-style connectors are available. But this leads to difficulties with modal noise and dispersion, effectively limiting the bandwidth, and thus makes this type of fiber unsuitable for our LO-transmission requirements. We will continue close monitoring of new developments in this area. Unlike the AT project, the MMA project will have time to develop a special connector if a suitable commercial product does not become available.

Suitable optical-fiber cables are available now. For example, an armored cable suitable for direct burial, containing four single-mode fibers with less than $\sim$ 1 dB/km loss and four multi-mode fibers capable of 400 MHz-km each, costs less than \$5 per meter. To this must be added the installation cost, estimated at \$6 per meter for burial at one-half-meter depth. The system also requires terminal electronics at the central building and at each antenna (estimated at \$30,000 per antenna); connectors (at \$150 per fiber-end, installed); and development and engineering (estimated at \$400,000).

Since there will be nearly four times as many antenna stations as antennas (to allow reconfiguring), we need to consider whether to install one cable per *station* or one per *antenna.* The latter would minimize the cable cost but require the routing of signals through unused stations by patch cables, which would increase connector costs and perhaps compromise reliability. Since most stations will be close to the array center, the cable cost is dominated by the outermost stations' requirements, so the extra cable required at one per station is only about twice, not four times, that for one per antenna. Furthermore, since the cables to many stations can share a common routing over much of their lengths, the installation costs will be nearly the same for either approach. We therefore select the 'cable-per-station' option. This leads to the following cost equation:

$$T = (5\,\text{k\$/km})D_1 S + (6\,\text{k\$/km})D_2 N + (31\,\text{k\$})N + 400\,\text{k\$}\,, \qquad (1)$$

where T is the transmission system cost, D_1 is the average distance to all antenna stations, D_2 is the average distance to the antennas stations of the largest configuration, S is the number of antenna stations, and N is the number of antennas. The terms cover cable, installation, terminal electronics and connectors, and development, respectively.

4. Correlator

Much of the versatility of the MMA as a research instrument is contingent on the flexibility designed into the correlator. The need for sensitivity implies that the correlator should be capable of analyzing broad continuum bandwidths. Spectroscopy of galaxies and dark clouds imposes quite different constraints on spectral resolution. Rapid imaging—particularly, rapid comparative imaging requiring simultaneous observations of two or more spectral lines—further increases the complexity of the correlator but provides a vital observational capability. Drawing together these demands, the resulting specifications of the MMA correlator are summarized in Table VIII-1.

Table VIII-1. Correlator Requirements

Number of antennas	$N = 40$
Maximum total bandwidth per antenna (arranged as $J \geq 4$ channels)	$B = 2$ GHz
Spectral channel bandwidth	$b = 2.0$ MHz†
Polarization: cross-polarized products available if B, b, or N^2 is degraded by 2×	
Special features: sideband separation	
enhanced spectral resolution or bandwidth	

†Can be degraded in steps of a factor of two by a reduction in the total bandwidth analyzed. Several forms of this tradeoff are available (see text).

As one would expect, for any reasonable configuration the cost is proportional to the total input bandwidth $NB = 80$ GHz, where N is the number of antennas and B is the maximum analyzed bandwidth. This 'figure of merit' for the MMA correlator is much larger than for any correlator so far constructed (e.g., for the VLA correlator $NB = 27 \times 200$ MHz $= 5.4$ GHz and for the VLBA $NB = 20 \times 128$ MHz $= 2.56$ GHz). Therefore, the minimization of cost must be carefully considered in the correlator design.

Optical correlation approaches have been considered (see MMA Memoranda Nos. 7 and 23), but, because of problems of flexibility and dynamic range, this is not currently our favored technique. Rather, we consider digital- and hybrid-(digital/analog) electronic approaches. We begin by investigating the impact of the primary specifications on the overall size and cost; details of implementation that affect important operational features are considered separately.

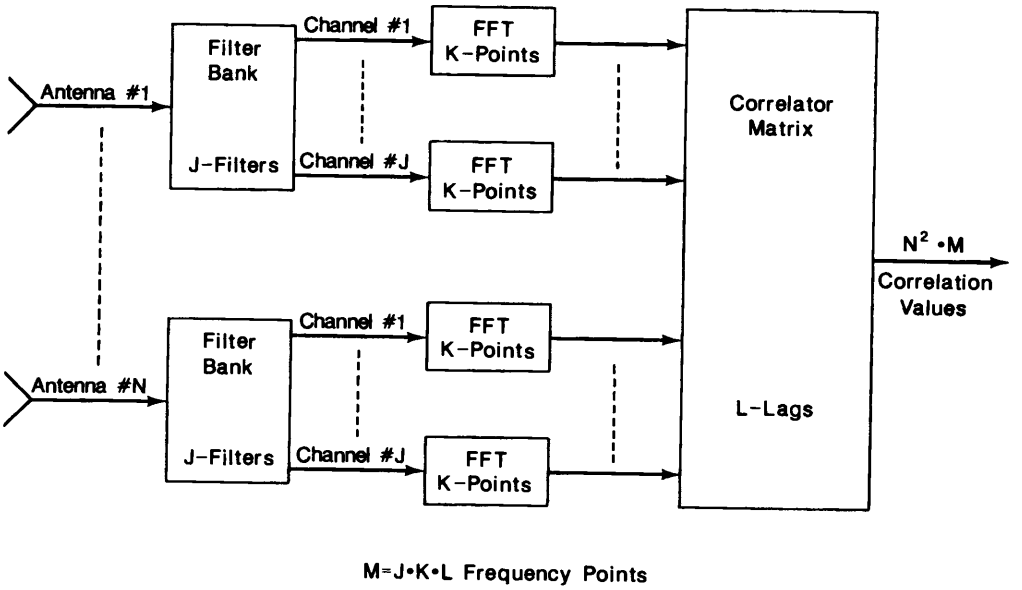


Figure VIII-1. General configuration of a correlator in which spectral analysis is accomplished by a combination of analog filters, FFT operations, and cross-correlation. The minimum-cost configuration is described in the text.

4.1. The Cost Equation. The required frequency resolution b may be obtained in several ways: by using (analog) filter banks, (digital) pre-correlation Fourier transformation, or (digital) time-domain cross-correlation (with post-correlation transformation). In principle a hybrid correlator could use all three techniques, as illustrated in Figure VIII-1. We now investigate the optimal distribution of filters, FFTs, and cross-correlators for minimization of cost. For this purpose, we assume that each component has a cost coefficient that depends on technology but is independent of the system parameters; these coefficients are: C_F, the cost of one analog filter; C_T, the cost of hardware capable of performing one FFT-radix-2 butterfly per second; and C_X, the same for one cross-correlation (1-sample delay, multiply, and accumulate) per second. The total cost is then

$$C_{\text{tot}} = C_F N J + C_T N B \log_2 K + C_X 2 N^2 B L , \tag{2}$$

where J is the number of filters per antenna, K is the length (number of complex samples) of the pre-correlation FFTs, and L is the number of lags per cross-correlator. It has been assumed that each signal is digitized by sampling at the Nyquist rate $2B/J$. The quantization coarseness will have some effect on the cost, and this is assumed to be included in the coefficients. The cross-correlator also computes the autocorrelations. Under these assumptions, Equation 2 is exact.

The total number of frequencies resolved is then JKL, and the frequency resolution is

$$b = \frac{B}{JKL} . \tag{3}$$

Taking B and b to be fixed by the specifications, Equation 3 constrains JKL; solving for K and substituting into Equation 2 allows us to minimize the cost over all J and L. The formal result of a linear optimization is

$$L_{\rm opt} = 0.721\frac{C_T}{NC_X} \tag{4}$$

and

$$J_{\rm opt} = 1.443\frac{C_T B}{C_F}\,. \tag{5}$$

There are a few problems with the linear optimization that yields the latter two equations. First, the parameters L, J, and K are constrained to be positive integers. If we find $L_{\rm opt} \ll 1$ formally, as happens in practical cases, then rounding up to $L = 1$ will cause $C_{\rm tot}$ greatly to exceed its formal minimum. Next, for convenient FFT implementation, K should be highly composite, preferably a power of two, and must not be too small; $K \geq 32$ is a reasonable constraint. Finally, the bandwidth may be filtered into several channels for reasons independent of this optimization (e.g., dual-band operation, bandwidth synthesis, polarimetry), forcing additional constraints on J. If we minimize $C_{\rm tot}$ under the constraint $K = 1$ (an 'XF' correlator), we find that

$$J_{\rm opt}({\rm XF}) = B\sqrt{2NC_X/bC_F}\,, \tag{6}$$

and if we constrain L to unity (an 'FX' correlator) we find that

$$J_{\rm opt}({\rm FX}) = 1.443\frac{C_T B}{C_F} = J_{\rm opt}\,. \tag{7}$$

The cross-correlation cost rate C_X deserves special attention, in that it depends strongly on the dynamic range and the digital representation of the data to be correlated. In the case of an 'XF' correlator, the data are coarsely quantized (to one or two bits), and each cross-multiplier can always be associated with the same integration register; furthermore, the data are real, simplifying the multiplier organization (although the real correlation rate per frequency channel is the same as for a complex correlator). Otherwise (for an 'FX' correlator), the potential dynamic range increases in proportion to $\log K$; successive products must be accumulated in separate registers (because they represent different frequencies or baselines); and the numbers are complex. The net result is that C_X is much larger in the latter case.

To make use of these results, we must know the cost coefficients C_F, C_T, and C_X. For C_F, the cost of an analog filter, we take \$900, based on experience with the hybrid correlator built for the NRAO 12-Meter Telescope.[15] This cost should be fairly stable. The costs of digital circuitry are harder to estimate

[15] S. Weinreb, "Analog-filter digital-correlator hybrid spectrometer", *IEEE Trans. Instr. & Meas.*, **IM-34** (1985) 670.

because of the rapid evolution of digital technology. It is clear from recent experience in building large correlators that the most economical approach is to base the design on one or more application-specific VLSI circuits (ASICs), in which case the overall cost can be extrapolated from the cost of these chips. Table VIII-2 is a compilation of data on ASICs that have been used in other large correlators. Each chip can be characterized by a figure of merit equal to the equivalent number n_x of real cross-correlations per clock cycle multiplied by the clock speed f_c; our cost factor when this chip is used in a cross-correlator is then $C_X = c_c/n_x f_x$, where c_c is the total cost of the chip and its support circuits (smaller ICs, circuit board, power supplies, etc.). A similar approach is used for FFT chips to obtain C_T, but in that case we have only the VLBA chip as an example. At the end of the table, an attempt has been made to project the technology a few years into the future. It assumes the use of a CMOS gate array at larger die size and clock speed, both of which are feasible now and will almost surely be in routine production within several years. Other details of the estimation procedure used are given in footnotes to the table.

Some conclusions are immediately apparent from Table VIII-2. The cost ratio C_X for XF correlators has been reduced with each new chip development, through higher levels of integration, higher clock speeds, or both. The hardware for a correlator equivalent to the VLA's would now be five times more expensive than in 1977. Also, the FX architecture results in an increase in the unit cost of cross-multiplication of four to eight times, for the reasons discussed above.

Applying these results to the problem at hand leads to Table VIII-3. Results based on the VLBA-correlator parameters are given for comparison. It is clear that construction of the MMA correlator with present technology would be very expensive, with an estimated hardware cost of 12 M$ to 14 M$. Perhaps surprisingly, the XF architecture is somewhat less expensive, even though it is drastically more expensive in the VLBA case. This is primarily because of the large difference in spectral resolution requirements. Fortunately, the situation looks much better using next-generation technology. Here the FX architecture looks much better, largely because we project only a factor-of-four cost penalty in C_X for FX architecture, vs. a factor-of-eight for presently available chips. It can be seen that the optimizations often yield numbers of filters or lags much less than unity, whereas the costs have been calculated for integral values. Further, we choose $J = 8$ input channels as a practical value from flexibility considerations (see below), so the cost for this number of filters is given when the optimum is fewer; it can be seen that the cost penalty is small. The selected configuration is therefore the FX architecture with $J = 8$ input channels, which has an estimated hardware cost of 1.3 M$. This cost should be regarded as a lower limit. Including inflation, ancillary hardware, and assembly, the hardware cost estimate becomes 3.2 M$.

To the hardware costs just estimated must be added the development costs and the cost of a computer to handle control and output archiving. These are fairly independent of the technology and are estimated at 2.5 M$ and 0.5 M$, respectively, based on the VLBA experience. Thus, the total cost estimate for the correlator is 6.2 M$ (1990).

Table VIII-2. Digital Correlator Integrated Circuits

Project Name	Year of Design	f_c MHz	n_x	— ASIC Costs — NRE/run, c_1	Production, c_2, \$	— Support — Chips	\$	Total Cost $c_c = \sum c_i$, \$	— Cost Ratios — C_X(FX) \$/GHz	C_T \$/GHz	C_X(XF) \$/GHz
VLA	1977	100	1.0	63 k\$/18000	5.50	2.1	55	64			640
NFRA*	1986	64	16	35 k\$/6400	56	3.7	155	216			211
OVRO†	1988	256	8	100 k\$/5000	100	3.5	148	268			130
VLBA‡	1989	32	4	95 k\$/5000	89	3.3	142	250	1950	1950	—
Future XF	1995	120	64	100 k\$/5000	90	3.5	148	288			37.5
Future FX	1995	120	16	100 k\$/5000	90	3.5	148	288	150	150	

Notes:

f_c = maximum clock frequency

n_x = equivalent number of real correlations per clock period

$C_X = \frac{c_1+c_2+c_3}{n_x f_c}$

'Support' includes non-ASIC chips, circuit boards, power supplies, wiring, and hardware. Based on actual cost and chip count for VLA and VLBA (excluding tape drives and their interfaces); estimated for others.

*A. Bos, "The NFRA correlator chip", Internal Tech. Report No. 176, Dwingeloo Radio Obs., Aug. 1986; R. Escoffier, "VLBA correlator cost estimate", VLBA Correlator Memo No. 86, June 1987.

†M. Ewing, "High speed correlator chip, version 1, preliminary design", OVRO Correlator Memo No. 19, April 1987.

‡R. Escoffier, "VLBA correlator ASIC chip specification", VLBA Correlator Memo No. 87, July 1987.

Table VIII-3. Cost Estimates For Various Cases

	VLBA	MMA
Number of antennas, N	20	40
Total bandwidth, B	128	2,000 MHz
Spectral resolution, b	.031	2 MHz
Formal Optimization:		
$C_X = C_T = 1950\$/\text{GHz}$		
L_{opt}	.04	.02
J_{opt}	.28	6.25
C_{opt} (formal)	105	2,496 k$
Present Technology:		
$C_X = C_T = 1950\$/\text{GHz}$		
J_{opt}(FX)	.53	6.25
C_{tot}(FX), $J = 1,6$	278	13,780 k$
$J = 8$	384	13,860 k$
Actual cost, $J = 8$	660*	—
$C_X = 145\$/\text{GHz}$		
J_{opt}(XF)	214	160
C_{tot}(XF) ($J = 128$)	2,779	12,032 k$
Future Technology:		
$C_X = C_T = 150\$/\text{GHz}$		
J_{opt}(FX)		2.89
C_{tot}(FX), $J = 3$		1,169 k$
C_{tot}(FX), $J = 8$		1,332 k$
$C_X = 32\$/\text{GHz}$		
J_{opt}(XF)		75
C_{tot}(XF), $J = 64$		6,144 k$

*VLBA actual cost for hardware only, excluding tape drives and their interfaces.

4.2. Operational Features and Flexibility. There is considerable scientific interest in maintaining a high level of flexibility in the correlator and in assuring that certain desired operational features are included, with particular emphasis on being able to separate the sidebands of a DSB front-end and on being able to observe more than one spectral line at a time. This section addresses such issues.

We assume that:

(a) The 2 GHz bandwidth is supplied in $J \geq 4$ separate IF channels, and these can be processed independently by the correlator. The channels may be from different front ends, from different parts of a wide IF in one front end, from different polarizations, or from some combination of these.

(b) Fringe stopping has been done in the LOs. This means that the fringes have been stopped at the total LO frequency for each IF channel or at the center frequency of the channel.

(c) The front ends use double-sideband mixers, but accurate phase switching of the first LO in increments of $\pi/2$ is possible, and this can be synchronized with correlator integrations.

4.2.1. Separation of Sidebands. Assuming that accurate phase switching of the first LO is available, visibilities in each sideband can be measured separately. If the switching interval is not too short, then this requires doubling the size of only the long-term accumulator in the correlator. The long-term accumulator cost is usually less than 10% of the cost of the cross-correlation part of the correlator. If the short-term dump time of the correlator is 100 msec, then for forty antennas the total switching cycle time is 6.4 sec (assuming Walsh-function switching patterns), which becomes the quantum of integration time for sideband-separated observations.

The sideband separation does not, in itself, impose any penalty on the signal-to-noise ratio (SNR); that is, a line appearing in only one sideband will be detected with the same SNR whether or not the sideband separation procedure is applied. But the noise of both sidebands is present in each result, so the SNR is worse—typically by a factor of $\sqrt{2}$—than would be the case if the front ends were single sideband.

4.2.2. Polarization Measurements. For these measurements, the correlator can be reorganized to produce the necessary cross-polarized correlations, provided that only half the maximum bandwidth is processed. That is, if half of the input channels are unused, then the corresponding sections of the correlator can be devoted to the additional correlations. This is possible in both the FX and the XF architectures. The cost lies in providing the appropriate interconnections between correlator sections; in accord with our experience in designing the VLA and VLBA correlators, the incremental cost can be made negligible by careful design. On the other hand, if full polarization measurements must be made at the full bandwidth, then the cost of the *cross-correlation part* of the correlator must be doubled. With present technology, it appears that the cross-correlation part dominates the cost in any architecture, in which case the total cost also doubles.

In the XF case, it is also possible to obtain polarization measurements by degrading the spectral resolution while continuing to process the full bandwidth. This generally does not work with an FX correlator.

4.2.3. Multiple Observing Bands or Spectral Lines. This feature has a bigger effect on the design of the front ends and IF processing than on the correlator. The question is to what extent the correlator input channels are separately tunable. We have assumed that a minimum of four channels is available, and for various reasons eight or sixteen channels may be more appropriate. It is straightforward to arrange that each channel may be tuned to any frequency within the bandwidth of the first IF channel; in addition, any available front end can be connected to any channel.

It may be desirable to operate different channels at different resolution or bandwidth. This will be possible if the correlator is organized as J quasi-independent correlators, each handling one channel of all N antennas. Such an organization also facilitates a future increase in J (and hence total bandwidth) by adding identical hardware, but it makes difficult any future increase in N.

4.2.4. High Resolution. The maximum bandwidth of each channel is B/J, where B is the total bandwidth (assumed to be 2 GHz) and J is the number of channels. If the channel bandwidth is reduced by filtering, then the resolution b can be decreased proportionally just by slowing down the correlator's clock. The resolution can be further improved by recirculation or by not using some input channels, while keeping the correlator working at its full capacity. In this case, the resolution becomes b/α^2, where α is the bandwidth reduction factor.

This applies to both the FX and the XF architecture, but in the FX case additional memory must be built into every stage of the FFT engine and the number of stages must be increased by $\log_R \alpha$, where R is the FFT radix. The FX correlator is at a disadvantage because it must be built for a particular maximum FFT length. Although the FFTs of unused channels may be utilized to increase the number of stages and hence the resolution of the remaining channels, each stage must have sufficient memory to support the longer transform. It would be cost-effective to make this memory a factor of two or four larger than could be utilized at full bandwidth in order to support higher resolution at reduced total bandwidth.

4.2.5. Trading Baselines for Bandwidth or Resolution. If each antenna actually provides more channels than the correlator can process, then by not using some antennas and connecting the extra channels from the remaining antennas to the now-free correlator inputs, larger bandwidths can be accepted. This would be useful only for special purposes, such as observing more lines simultaneously when SNR and u-v coverage on any one line are not limitations.

Another such trade is conceivable. While retaining all antennas at the full bandwidth, some baselines might be of little interest because of redundant u-v coverage. In principle, the cross-correlation hardware for these baselines might be re-allocated to improve the spectral resolution on the other baselines. However, this increases the complexity of correlator organization to such an extent that it should not be implemented unless this feature is clearly shown to be essential.

IX. OPERATIONS

With the Millimeter Array located on a high-altitude site in Arizona or New Mexico, operation of the MMA will necessarily have two foci. First, the actual telescope functions—moving the antennas, connecting power and signal-transmission systems, identifying and replacing defective components, etc.—and the data-taking function, both must occur at the MMA site. Second, the remainder of the extensive activity concerned with the operation of a scientific user-oriented facility can occur at a more convenient location. The repair of telescope modules and all user-interaction with the MMA and with the data reduction software are in this latter category. Fortunately, because the three potential MMA sites are close to existing NRAO facilities in Socorro, NM and Tucson, AZ, it is possible to share the present infrastructure and realize a significant operational cost savings.

In estimating operating costs we have specifically estimated the incremental cost of operating the instrument together with the on-going NRAO operations. Cost of a stand-alone operation, of course, would be higher. In addition, we have assumed that a small operational crew would work at the MMA site. Depending on the actual site chosen, these individuals would be based either in the area of Socorro/Magdalena, NM, or in Springerville, AZ.

One of the advantages in splitting the operation in this fashion is that it leads to a straightforward generalization that may be a convenience to remote users. In real time the data from the array correlator will be sent via optical-fiber cable to the NRAO office, there to be resident for imaging and analysis by the astronomer. However, the data could also be sent on to the user's home institution or computational facility. Given that the MMA scheduling must be flexible enough to meld scientific needs with changing weather conditions, it may be impossible or undesirable to fix the schedule long in advance. Remote observing and, for some observers, remote imaging may be the optimal way to use the MMA.

Personnel needs and cost estimates for MMA operation as outlined above are derived from experience with the Tucson 12-Meter Telescope and with the VLA. As part of an operation incremental to NRAO operations, the MMA staff will be smaller than that of the present VLA. The distribution of skills will also be somewhat different. Table IX-1 summarizes by job category the personnel needs of the MMA, and, for comparison, the present VLA staffing levels are also shown.

A few of the differences in this table deserve discussion. First, the MMA site and antenna crew is so much smaller than the corresponding VLA crew primarily because the MMA site is so much smaller than the enormous VLA site. The MMA has no railway system to maintain. Antenna reconfiguration is always over short distances. There is no cafeteria planned for the MMA site, and only a modest residence hall. Second, the MMA electronics group is larger

Table IX-1. MMA and VLA Personnel

	VLA	MMA
Site and Antennas	40	12
Electronics	26	38
Array Operations	9	8
Computer	13	12
Business	12	8
Scientific Services/Management	12	12
TOTAL	112	90

than the VLA electronics staff. This reflects both the complexity of millimeter-wavelength equipment relative to centimeter-wavelength instrumentation and the fact that its maintenance is such demanding work. Note that the total numbers of receivers on the two arrays are about the same: 400 polarization/frequencies on the MMA, and 336 presently on the VLA. Finally, although the computer staffing levels of the two arrays are comparable, we imagine most of the MMA individuals to be concerned with algorithm development pertinent to imaging techniques, while several of the existing staff are involved in system management and maintenance. This area is a good example of the operational savings that accrue to a shared operation.

Approximation of the total operations budget for the MMA, including support of the 90 individuals noted above, in 1990 dollars, is shown in Table IX-2.

Table IX-2. MMA Operations Costs

Personnel	
Salary	\$ 2750 k
Benefits	780 k
Communications and Utilities	900 k
Materials, Services, and Supplies	950 k
Travel	125 k
Subtotal	\$ 5505 k
Equipment	1000 k
TOTAL	\$ 6505 k

These costs are estimates. The precise costs of electric power and communications, for example, will depend on the site chosen.

The equipment line in the operating budget subsumes two categories of expense. The first is maintenance and replacement of defective instrumentation, and the second includes installation of new equipment. The annual equipment budget suggested here is one percent of the capital cost of the MMA, or \$25 k per antenna per year. It should be higher.

X. CONSTRUCTION COST AND SCHEDULE

1. Construction Cost

In the previous chapters the costs of the major components of the Millimeter Array were estimated and discussed. Here we summarize those costs.

1.1. Antennas. The design of the 8-m transportable antenna incorporates a conventional steel mount together with a back structure employing both carbon-fiber-reinforced-plastic lateral members and carbon-steel diagonal members and nodes. The central hub is carbon steel very much like that used in the IRAM 15-m antenna. The panels will be either CFRP or machined aluminum; the cost estimate shown below assumes CFRP panels. Our cost estimates for the mount and the servo drive are made in cognizance of the need for 1-arcsecond rms pointing accuracy.

Each of the antennas is equipped to provide total-power measurements of the sky. An optical chopping system—incorporating a nutating secondary mirror or a tertiary mirror used as a chopper—is estimated separately at \$100 k per antenna. An additional \$800 k is budgeted for special-purpose vehicles to transport the antennas between antenna stations.

1.2. Receivers.

1.2.1. 9-mm HEMT. The 9-mm receiver will be adapted from the VLBA design and packaged in a 20-K dewar identical to that used for the VLBA. There is little development cost required for the MMA application. The costs of the dewar and dual-polarization 9-mm mixers are:

Materials cost	\$40,000
Assembly	20,000
TOTAL, per antenna	\$60,000

1.2.2. SIS Receivers. As discussed in Chapter VIII, each antenna will have two 4-K cryostats, each containing four SIS mixers (orthogonal polarizations in each of two frequency bands). The SIS receivers are selectable with a rotatable central mirror, or they may be used simultaneously if that mirror is a dichroic mirror. Since the cost of the 4-K cryostat is far greater than the mixer cost, packaging several mixers in one dewar results in a significant savings.

A prototype of the MMA receiver package has recently been installed on the NRAO 12-Meter Telescope. A single dewar cooled by one 4-K refrigerator may contain eight independent SIS mixers and receiver inserts, including the IF amplifiers (4 frequency bands with 2 polarizations each). We have extrapolated the costs of MMA receivers from the experience of constructing this prototype receiver. Each MMA receiver package is somewhat simpler than the 12-Meter

prototype—four mixers in each MMA package initially—and the costs break down as follows:

4-K refrigerator (incl. JT and compressor)	$40,000
Dewar	6,000
Optics (incl. diplexer, LO leveler)	8,000
SIS mixers	20,000
Feed horns and lenses	3,500
IF amplifiers	12,000
Receiver inserts	24,000
Tuning mechanism	4,500
Electronics (SIS and LO bias, 2nd LO)	35,000
Frame, assembly, test	7,000
TOTAL, per receiver (2 frequencies)	$160,000

Meeting the requirements of the Millimeter Array for SIS receivers will demand a substantial advance in the state of the art. Receiver reliability is, of course, a paramount issue for the MMA, as is broad bandwidth and ease of tuning: it is unrealistic to contemplate two or more mechanical tuners running reliably on each of the MMA's 320 SIS mixers. For the MMA application we have worked to develop broadband tunerless SIS mixers. A prototype design exists for 115 GHz, but considerable development work is needed to produce full waveguide band versions for the four frequency bands that use SIS mixers. The SIS development costs shown for the MMA are addressed to this task. We estimate that a four-year development program will be required, involving approximately $250 k per year for mask design and production, most of the expense going under contract to a university-based or industrial device supplier. An amount approximately equal to this will support a staff of four individuals working in mixer fabrication and development. Thus, the total development cost will be $2 M.

This activity should begin in 1991 and would position us to commence receiver production for the MMA by 1995.

1.3. Local Oscillators. In principle, development of the local-oscillator system is straightforward. However, it is made challenging by the need for broadband performance. Fortunately, considerable work is in progress at various universities and commercial organizations directed toward development of new devices for generation of high frequencies with sufficient bandwidth and power to be usable for the MMA local oscillators (see Chapter VIII). We plan to exploit the successes of others in this area, as appropriate. The LO cost shown below is the estimate for a multiplied Gunn source phase-locked to a master distributed LO, broken down as follows for each receiver package:

Gunn oscillators	$12,000
Multipliers	28,000
Harmonic mixers & phase-lock system	5,000
TOTAL, per receiver (2 frequencies)	$45,000

1.4. Correlator. A thorough discussion of the cost tradeoffs associated with the design and construction of the MMA correlator is given in Chapter VIII. A correlator capable of analyzing a total bandwidth of 2 GHz from forty antennas and with the analog flexibility to divide its input into as many as eight selectable bands (which may be chosen from among many combinations of frequencies and polarizations), built with today's hardware, would cost approximately \$12 M. However, we project that the favorable cost/performance trend in digital hardware will lead to a correlator of significantly less cost in two to five years, as discussed in Section 4 of Chapter VIII.

1.5. Transmission System. The cost of the optical-fiber transmission system is analyzed in Chapter VIII. The cost equation is

$$T(\text{k\$}) = 5D_1S + 6D_2N + 31N + 400\,,$$

where D_1 (km) is the average cable run to a station, S is the number of stations, D_2 (km) is the average cable run to stations in the most extended configuration, and N is the number of antennas. Using $D_1 = 0.5$, $S = 110$, $D_2 = 1.5$, and $N = 40$ leads to $T = \$2.3$ M.

1.6. Computers. The computing needs of the MMA, including both array control and image processing, are estimated in Chapter V. In this case we have estimated the magnitude of the MMA computational burden and projected the cost of its relief as \$9.0 M using 1995 technology. To place this figure in 1990 dollars, the estimate has been de-escalated by five percent per year, leading to \$7.2 M in 1990. As with the correlator, should the cost/performance ratio continue to decline, the expense in this category could be adjusted or a more capable computational facility could be assembled at the same cost. This is an item which it would be well to defer to the end of the construction project.

Development of the algorithms necessary for fast and efficient imaging with the MMA will require an additional \$1 M.

1.7. Communications. Since the MMA can image effectively in real time, it is possible for the astronomer to make decisions while observing. To this end, provision is made to connect the MMA correlator directly to other NRAO facilities and from there to the user via existing networks.

1.8. Site and Buildings. The costs for antenna stations, roads, and support facilities are based on the VLA and VLBA experience. They depend, in some measure, on the site chosen for the MMA; the figures shown here are predicated on the selection of any one of the three sites discussed in Chapter VI.

2. Construction Schedule

The construction schedule of the Millimeter Array is dependent on (1) site acquisition, (2) SIS-receiver design and prototyping, and (3) antenna design. Site testing has been a major effort for four years. We expect to select a site in the summer of 1991 and to obtain full approval for use of the site early enough to allow construction to begin no later than 1994. In Chapter VIII the

Table X-1. MMA Construction Costs (in 1990 Dollars)*

Antennas:		
	40 8-m, 25 micron rms	$50.0 M
	Optics	4.0
	Transporters	0.8
Front Ends:		
	9-mm	2.4
	SIS Dewar 1	6.4
	SIS Dewar 2	6.4
	SIS Development	2.5
Local Oscillator:		3.6
Correlator:		
	Development	2.5
	Hardware and assembly	3.2
	Computer	0.5
Transmission System:		2.3
Computers:		
	Hardware	7.2
	Algorithm Development	1.0
Communications:		1.0
Site:		4.0
Buildings:		4.5
Project Management:		2.0
	Subtotal	$104.3 M
Contingency (15%):		15.7
	TOTAL	$120.0 M

*Applicable state and local taxes would be incremental to the costs shown here and have not been included in these estimates.

need for approximately four years of SIS design work was described: the MMA receivers—broadband, tunerless, and reliable—represent a significant advance in the state of the art. Beginning in FY 1991, at a modest level, this activity also could accommodate an FY 1994 start of major construction. Finally, the antenna design could begin as early as 1992, when the NRAO design team will have completed their work on the Green Bank Telescope. A 1994 construction start is consistent with the time required for the detailed antenna design.

Given these constraints, the MMA project should begin in FY 1991, with preliminary design funding leading to five years of major construction beginning in 1994. Useful observations could begin with an interim MMA in 1996, and Array completion would occur in 1998.

Table X-2. MMA Project Schedule

1991	Site Evaluation Site Selection Electronics Development	$ 1.0 M
1992	Site Configuration Layout Antenna Design Electronics Development Algorithm Development	2.0 M
1993	Electronics Development Antenna Design Algorithm Development	5.0 M
1994	Site Configuration Construction Electronics Design Algorithm Development Construction	30.0 M
1995	Construction	30.0 M
1996	Construction Interim Operations	30.0 M
1997	Construction Interim Operations	22.0 M
1998	Full Operation	

APPENDIX A.
TECHNOLOGY DEVELOPMENT FOR THE MMA

The Millimeter Array is a unique instrument. It is the only synthesis telescope conceived and designed as a complete imaging instrument capable of measuring accurately all spatial frequency components, from zero to the maximum array baseline. Partially, the motivation for such a complete instrument comes from the spatial complexity of the sky at millimeter wavelengths and the interrelationship of astrophysical phenomena on a wide range of spatial scales. (For example, do sites of nascent star formation—cold cores—within extended GMCs evolve from, or initiate, specific local changes in the cloud chemistry?) And partially, the motivation derives from the experience of imaging with existing synthesis-array telescopes, especially that gained from the WSRT, the VLA, and the BIMA and OVRO millimeter-wave arrays, all of which have identifiable limitations. The design of the Millimeter Array incorporates remedies for the deficiencies recognized in existing synthesis arrays, and it projects capabilities to satisfy the needs unique to millimeter wavelength astronomy in the next decade and beyond.

The burden of designing a unique instrument, however capable, is that it necessarily demands an extension of existing technology. This is as true for the MMA with its densely packed configuration, total-power instrumentation, and broadband receivers as it was for the VLA (waveguide IF transmission, multiband up-converter receivers), the Keck telescope (segmented optical-quality primary mirrors), the Columbus and Magellan telescopes (spin-cast mirrors), the Green Bank Telescope (unblocked 100-m aperture, real-time surface metrology), the IRAM telescopes and the SMT (application of CFRP technology for extreme thermal stability). Application of significant technological advances is the *sine qua non* of the design of a forefront scientific instrument.

The MMA design calls for technological progress to be made in nearly all areas, three examples of which are noted below. Work in each of these areas has begun, albeit at a very modest level, and will continue with the guidance and involvement of the community of MMA users in the years preceding MMA construction.

1. Antennas

1.1. Surface Accuracy. The MMA specifications call for an 8-meter antenna which, in 1995, is built to achieve an rms surface accuracy of 25 μm, is transportable, and which points to $1''\lambda_{(\mathrm{mm})}$ rms accuracy. There are several other considerations that influence the design adopted for the MMA antennas as noted in Chapter VII, particularly the need for coudé optics and a beam-switching capability on the antenna, but for the purpose of identifying the technological forefront let us focus on the surface accuracy and the pointing.

The state of the art in the construction of precision antennas is represented by the Submillimeter Telescope (SMT), a joint project of the Max-Planck-Institut für Radioastronomie and the University of Arizona's Steward

Observatory.[1] This 10-m telescope will achieve a surface accuracy of 15 μm, substantially better than that required of the MMA's antennas. The SMT uses CFRP surface panels and backing structure; it will therefore retain its precision under changing thermal conditions. When in use, the telescope is fully exposed to the environment. It can be sheltered from inclement weather in an enclosure. The SMT is not transportable.

The two submillimeter-wave antennas on Mauna Kea, Hawaii—the Caltech Submillimeter Observatory and the James Clerk Maxwell Telescope—both achieve surface accuracies close to the MMA specification, but both achieve control of the thermal environment by means of an enclosing astrodome.[2,3] Neither telescope, of course, is transportable.

Three ten-meter antennas designed by R. B. Leighton make up the OVRO millimeter-wave interferometer. These antennas are transportable. They are also very stiff against gravitational deformations, and in a benign thermal environment (a calm, thermally stable night) the total surface error over the zenith-angle range 15° to 70°, as determined from holographic measurements, can be made less than 20 μm rms.[4] This is better performance than required for the smaller MMA antennas.

The MMA needs to achieve the expected surface accuracy of the transportable Leighton antennas, but, moreover, it needs to do so over the full range of thermal conditions, including full sun on the antenna, faced by the SMT. In this particular area, the MMA antennas need not break new technological ground to reach the required accuracy; instead the design will meld the solutions developed to meet the needs of several observatories, particularly in the application of CFRP members in the backing structure to control thermal deformations. The principal antenna-design work will concentrate on meeting the required specifications at the lowest cost.

1.2. Antenna Pointing Accuracy. A paramount consideration, one which poses greater difficulties at the higher observing frequencies, is that of maintaining the pointing accuracy to a small fraction of the size of the antenna primary beam. Moreover, the problem is exacerbated as the telescope's sensitivity improves (owing to better receivers, for example), for then telescope users can construct larger maps in the same observing time and they are faced with the difficulty of accurately placing observations on a spatial grid of the sky. Thus, imprecise pointing translates directly into errors in the resulting image. In the past, even at millimeter wavelengths (e.g., CO $J = 1$–0 on the NRAO 12-Meter Telescope), pointing errors as large as five to ten arcseconds were still less than one-sixth of the telescope half-power beam width. This is no longer true at

[1] J. W. M. Baars and R. N. Martin (1989), in *Physics and Chemistry of Interstellar Clouds*, G. Winnewisser and T. Armstrong, Eds., (Springer-Verlag: Berlin).
[2] E. Serabyn, C. R. Masson, and T. G. Phillips (1990), in *Submillimeter Astronomy*, G. D. Watt and A. S. Webster, Eds., (Kluwer: Dordrecht), p. 41.
[3] A. S. Webster, A. R. P. Russel, H. E. Matthews, G. D. Watt, S. S. Hayashi, I. M. Coulson, R. Genzel, A. I. Harris, T. Stutzki, U. U. Graf, and R. Padman (1990), *ibid.*, p. 213.
[4] D. Woody (1990), in *Submillimeter Astronomy*, G. D. Watt and A. S. Webster, Eds., (Kluwer: Dordrecht), p. 43.

high frequencies: As we push telescopes to frequencies well beyond their design limits, the errors inherent in the motors, servomechanisms, encoders, bearings, etc. (all of which were specified corresponding to the original design-frequency limit) collectively result in an insurmountable limit to the pointing accuracy.

In order to mosaic most effectively, the MMA antennas, in the compact configuration, need individually to track to within $1''\lambda_{(mm)}$ rms for the time it takes to construct the mosaic. Because the array is fast, a small mosaic may take a few tens of minutes, while the longest observation, in a single day, will take no more than four hours (viz., between plus and minus two hours of meridian transit, in order to avoid the penalty in system temperature that accrues to observations made at increasing air mass). Thus, the requirement for $1''\lambda_{(mm)}$ rms antenna pointing refers to observations of a particular region of the sky (it is not a 'blind' pointing requirement), over a limited range in hour angle, around a limited range in azimuth centered on meridian transit, and for a period no longer than four hours.

The state of the art in pointing large, exposed, millimeter-wave antennas is most advanced at IRAM, where the 11″ beam of the 30-m telescope at 1.3 mm introduces a requirement for 1″ rms pointing. While it is more difficult to point a 30-m antenna than an 8-m antenna, the same considerations apply to both, especially the need to include provisions for such precision pointing in the specification of the basic antenna design. The IRAM 30-m design, the SMT design, and the MMA design all do this; at the 30-m telescope the technique is the following:

> Prior to beginning an observation for which precision pointing is important, the observer establishes fiducial ('local') pointing corrections appropriate to an object of known position located near the region of interest. This is not necessarily a continuum source. One then observes the program source, returning periodically to the pointing source to refine the telescope pointing offsets. If pointing is done once an hour in this manner, the rms antenna pointing precision can be maintained to 2″ rms.[5] Reducing this error to 1″ requires more frequent pointing—that is, pointing more than once an hour. The point-source-detection observations done on the 30-m at 1.3 mm and 0.87 mm require this pointing accuracy and maintain it.[6,7,8]

Thus, there are two ingredients to precision pointing. First, the antenna must be designed and built, in all its component parts (drives, servomechanisms, motors, bearings, encoders) to the astronomical specifications required; and second, the pointing may need to be refined once or more per hour during

[5]See, e.g., Truong-Bach, D. Morris, N.-Q-Rieu, and S. Deguchi (1990), *Astron. & Astrophys.*, **230**, 431.

[6]W. J. Altenhoff, W. K. Huchtmeier, E. Kreysa, J. Schmidt, J. B. Schraml, and C. Thum (1989), *Astron. & Astrophys.*, **222**, 323.

[7]R. Chini, P. L. Biermann, E. Kreysa, and H.-P. Gemünd (1989), *Astron. & Astrophys.*, **221**, L3.

[8]R. Chini, E. Krugel, and E. Kreysa (1990), *Astron. & Astrophys.*, **227**, L5.

the observations. The former consideration accounts for the relatively poor performance of older antennas ('poor' compared only to today's desires, *not* poor compared to the specifications to which the antennas were designed and built) and brings with it a cost premium to the MMA antennas. The latter dictates operational considerations for the mosaicing mode of the MMA in its compact configuration. In practice, the need for frequent pointing observations might best be addressed by equipping the antennas with co-aligned optical/IR detectors and video cameras, so that stellar pointing sources could be used. These techniques are under investigation at the NRAO and elsewhere.

The cost premium needed to achieve precision pointing as defined above for the MMA is explicitly budgeted in Table VII-1. Note that the cost of the antenna mount alone, which includes the drive system, is nearly 40% of the total cost of each antenna and that this cost depends on the pointing accuracy which one seeks to achieve according to

$$\text{Antenna Mount Cost (k\$)} = 250 + \frac{10D^{3/2}}{p}.$$

Here $D = 8$ m is the antenna diameter, and p is the required pointing accuracy, measured in arcseconds. This empirical formula is derived from the costs of the IRAM antennas and the SMT.

2. Receivers

2.1. Broadband, Tunerless SIS Mixers. Perhaps the area of greatest technological advance to the MMA, and indeed to all of millimeter-wave astronomy, is the development of broadband, tunerless SIS receivers. The MMA has as one of its central design goals complete frequency coverage across all of the atmospheric windows between 9 mm and 0.85 mm. The need for complete coverage is a lesson learned from the VLA, where redshifted galaxies and new lines—even enormously strong lines such as the 12-GHz methanol masers—are inaccessible, owing to the limitations of the instrument. For the MMA, with its concentration on the very rich millimeter-wave spectrum of molecular lines, we cannot tolerate such an artificial limitation. The alternative, however, is a need for broadband, sensitive receivers covering a decade of frequency.

In the years preceding construction of the MMA considerable effort will be devoted to developing SIS mixers which are (1) broadband; (2) of low noise figure; (3) easily reproducible; and (4) free of any mechanical tuners. Progress in this area will be paced by funding for design, experimentation, and evaluation, and by access to SIS device fabrication facilities.

2.2. Broadband Local Oscillators. A second area of technological development required by the MMA is that of broadband local oscillators. This issue is discussed in Chapter VIII where it is noted that although a solution exists in the form of phase-locked Gunn oscillators driving frequency multipliers, the achievable bandwidth is a limiting factor, and even this bandwidth comes at the cost of several mechanical tuning devices. The MMA needs for reliable, broadband

LOs go considerably beyond such an approach and will therefore require several years of research and development, which will be undertaken in cooperation with interested university and industrial groups. The resulting technological innovations will be broadly applicable to astronomical instruments other than the MMA.

2.3. Multi-Band 4-K Cryostats. The cost of MMA SIS receivers shown in Section 1.2.2 of Chapter X is dominated by the cost of the 4-K cryostat. As a way of minimizing this cost for a wideband receiving system such as that planned for the MMA, typically several receivers are packaged together, along with their separate feed horns and lenses, in a single cryostat. In the case of the MMA receiver package, eight individual SIS receivers share a common cryostat; the receivers are selectable quasi-optically in orthogonally polarized pairs. A prototype MMA multiband receiver has been installed on the NRAO 12-Meter Telescope and will be tested in astronomical observations beginning in the fall of 1990.

3. Algorithm Development

3.1. Linearly Mosaiced Images. The exceptionally good instantaneous u-v coverage of the MMA in its compact configuration means not only that it will be possible to mosaic very rapidly, but also that most final images will require no deconvolution. No other synthesis telescope, existing or planned, has this capability. This presents both a problem and an opportunity. The problem is computational: how best to acquire the data and present the mosaic image, quickly, so that the astronomer can make judgements and modify the observational program while it is in progress. The opportunity is to use that information to optimize the approach, or to experiment with the system. Development of the software to facilitate this (the computational task is not overwhelming) is a significant challenge that will benefit by the participation of the community of interested users.

3.2. Preview Observations and Presentation. For those observations in which the astronomer is searching to understand local phenomena in the context of a more extended astrophysical environment, it may be desirable to 'preview' the celestial region of interest by using the forty MMA antennas as individual, single antennas. The total-power hardware exists on each antenna to make that possible; the software needed to acquire the data and present it to the user in a form suitable for judgements to be made is a second software task which will receive our attention in the years before actual MMA construction begins.

APPENDIX B.
MMA 'SINGLE-DISH' OBSERVING MODES

As an instrument specifically designed to recover information at all spatial frequencies, from zero (total power) to the maximum interferometer baseline, the Millimeter Array necessarily includes a variety of capabilities which find application in specialized circumstances. Among these is the capability for astronomers to use the array antennas as individual elements to make total-power measurements, either independently or by summing incoherently all of the total-power signals. This capability is unique to the MMA and derives from each of the MMA antennas being a complete instrument in itself—each is equipped with total-power instrumentation. A few of the MMA 'single-dish' modes are summarized below.

Earlier discussion of the principal MMA low-resolution observing mode—mosaicing in the compact configuration—will not be repeated here. Instead the reader is referred to Chapter IV for this information. However, it is worth emphasizing that the MMA in the mosaicing compact configuration provides such complete instantaneous u-v coverage that the synthesized beam is in every sense equivalent to that of a single antenna, seventy meters in diameter, with a somewhat-reduced aperture efficiency (because the MMA equivalent aperture is $\sim$ 50% filled). The MMA response as a function of distance from the array center is very similar to the corresponding response of a conventional 70-m filled-aperture single antenna (see Fig. IV-5).

1. Preview Mode. The desire to obtain a low-resolution image of an extended region prior to making synthesis observations (often in order to establish the positions of features of special interest or to map the extent of an area to be imaged by the MMA as a mosaic) has been identified previously in Appendix A as requiring 'quick look' software development. A total-power spatial preview image, or spatial/spectral preview image, can be made rapidly by individually pointing the MMA antennas to differing, but adjacent, sky positions. Observing in this manner, a region of the sky $3\farcm3\lambda_{(\mathrm{mm})}$ square can be imaged completely with Nyquist sampling (half-power sampling for an 8-meter antenna) in a single integration period. A very large region could be imaged quickly at low resolution by combining a suitable number of such observations. The real task is to efficiently provide the astronomer with the resultant information so that the requisite judgements can be made to optimize follow-up synthesis observations.

Solar observations may especially benefit from the MMA 'single-antenna' preview mode. By pointing each of the MMA antennas to adjacent positions on the solar disk, separated by full beam spacings at 90 GHz for example, most of the solar disk can be imaged instantaneously and monitored continuously. When a flare erupts or an active region begins evolving, the entire array can be pointed to the region of interest and the subsequent evolution followed with high-angular-resolution synthesis observations. This is a good example of the flexibility inherent in the design of the MMA.

2. Deep Spectroscopic Searches in Extended Sources. Subtle questions in astrochemistry and not-so-subtle questions associated with the excitation or abundance of molecular species are often best answered by searches for weak confirming lines or isotopic lines of low optical depth. Sought for in spatially extended regions such as in the Orion molecular cloud, toward Sgr B2, or in the Taurus or Ophiuchus clouds, the task is to detect the lines and establish their brightness temperatures. For such research (1) the $31\lambda_{(mm)}$-arcsecond half-power beam width of the individual MMA antennas is not a significant limitation, and (2) there is no gain in observing with a larger antenna. Indeed, the only way to gain in sensitivity to *extended* brightness temperature is to accumulate integration time, and here the forty MMA antennas help dramatically. Pointing each of the MMA antennas toward the same sky position and incoherently summing their spectral output acts effectively as a forty-fold increase in integration time.

For detection of the brightness of weak spectral lines in extended sources the MMA is $\sqrt{40} \approx 6.3$ times faster than a single antenna of any size.

3. Detection of Faint Compact Sources: The Role of Bolometers. Cross-correlation interferometry is an extremely sensitive technique for the detection of weak signals from compact radiating sources owing to the fact that the astronomical signal is correlated in the two (or more) antennas, whereas all the other sources of noise are not so correlated—one seeks the signal in the presence of a level of noise equivalent to the difference between the powers received by the two antennas. In contrast, in using a total-power instrument such as a single antenna the astronomer seeks the signal in the presence of the total-power sum. The consequence of this latter fact represents a considerable challenge to single-antenna observations, which can be alleviated to some extent by taking rapid differences of the total power, alternately including and excluding the astronomical source (by nutating a subreflector, for example). The effectiveness of the technique depends on the balance that can be established, and maintained, between the 'on' and 'off' phases of this alternating cycle.

At centimeter wavelengths interferometry is far the more powerful technique: with a collecting area equivalent to a 125-m diameter single-antenna facility, the VLA routinely detects sources two orders of magnitude fainter than does the 100-m telescope of the Max-Planck-Institut für Radioastronomie (MPIfR) or than did the 91-m Green Bank antenna. To be sure, observations of the centimeter-wave sky suffer from confusion due to the random distribution of background sources which contaminate the single-antenna total-power measurements and account for part of this difference. What can we expect at millimeter wavelengths?

Point-source-detection observations using the MMA can be done in two modes, both of which provide the same sensitivity. In the case where the position of the source is known, the MMA can be used either as an adding interferometer or in the 'normal' fashion as a multiplying interferometer. For a source with poorly known position, one would, of course, use the multiplying mode. In either case the sensitivity of the array is that corresponding to the entire collecting area of the MMA, 2000 m^2, and the angular resolution is that corresponding to the

longest baselines in the array. The MMA analyzes a total continuum bandwidth of 2 GHz using double-sideband receivers. The resultant point-source sensitivity of the MMA in the continuum band at 1.3 mm is 1 mJy/$\Delta t^{1/2}_{(\mathrm{min})}$ for a zenith opacity corresponding to 2 mm of precipitable water vapor. In the 0.87-mm continuum band the sensitivity degrades to approximately 3 mJy/$\Delta t^{1/2}_{(\mathrm{min})}$ because of the greater sky contribution and the higher expected receiver temperature at the shorter wavelength.

The fact that the 1.3-mm continuum atmospheric window is more than 50 GHz wide (see Fig. II-4) provides an opportunity for great continuum sensitivity that is being exploited by equipping large millimeter-wave telescopes with incoherent bolometric detectors. The state of the art in bolometric receivers is the superb MPIfR instrument[1] which, on the IRAM 30-m telescope, provides a sensitivity of 70 mJy/$\Delta t^{1/2}_{(\mathrm{sec})}$ at 1.3 mm and 450 mJy/$\Delta t^{1/2}_{(\mathrm{sec})}$ at 0.87 mm, under the same sky as described above for the MMA, viz., 2 mm of precipitable water vapor.[2,3,4,5] The achieved sensitivity is limited by sky noise, that is, by imprecise sky subtraction, not by detector noise; in the absence of sky noise the detector sensitivity is 20 mJy/$\Delta t^{1/2}_{(\mathrm{sec})}$.[6] Thus, even if the detector noise could be reduced to zero the 1.3-mm sensitivity on the 30-m telescope would be no better than 50 mJy/$\Delta t^{1/2}_{(\mathrm{sec})}$, owing to the sky-noise contribution.

In Figure B-1 the achieved rms sensitivity of the 1.3-mm MPIfR bolometer on the 30-meter IRAM telescope is plotted as a function of integration time for every recent reference in which integration time and resulting rms sensitivity are quoted by the authors. The MMA point-source sensitivity is plotted for comparison.

This figure illustrates several points. First, the bolometer/30-meter combination detects point sources of tens to hundreds of milli-Janskys quickly, with just a few minutes' integration. Second, the sensitivity of the system is dominated by sky noise rather than detector noise.[7] Third, the sky noise provides the ultimate limitation to the achievable point-source sensitivity. In the case of this particular instrument on the 30-m telescope, that limiting rms point-source sensitivity is approximately 2 mJy, reached after an integration time of 30 to 60 minutes. Longer integrations fail to reach greater sensitivity because the total-power sky is not, or cannot be, subtracted to greater precision.

---

[1] E. Kreysa (1985), in *Proceedings of the URSI International Symposium on Millimeter- and Submillimeter-Wave Radio Astronomy*, (Granada, Spain), p. 153.

[2] R. Chini, E. Kreysa, and P. L. Biermann (1989), *Astron. & Astrophys.*, **219**,87.

[3] R. Chini, P. L. Biermann, E. Kreysa, and H.-P. Gemünd (1989), *Astron. & Astrophys.*, **221**, L3.

[4] C. J. Salter, R. Chini, C. G. T. Haslam, W. Junor, E. Kreysa, P. G. Mezger, R. E. Spencer, J. E. Wink, and R. Zylka (1989), *Astron. & Astrophys.*, **220**, 42.

[5] R. Chini, E. Krugel, and E. Kreysa (1990), *Astron. & Astrophys.*, **227**, L5.

[6] R. Chini, E. Kreysa, and P. L. Biermann (1989), *op. cit.*

[7] A further limitation, not illustrated here, is that the effective beam width of the bolometric detector varies as the sky fluctuates. The typical variation of the beam width for observations such as those shown is 20%, or from 10″ to 12″ (*cf.* Chini *et al.* (1989), *op. cit.*).

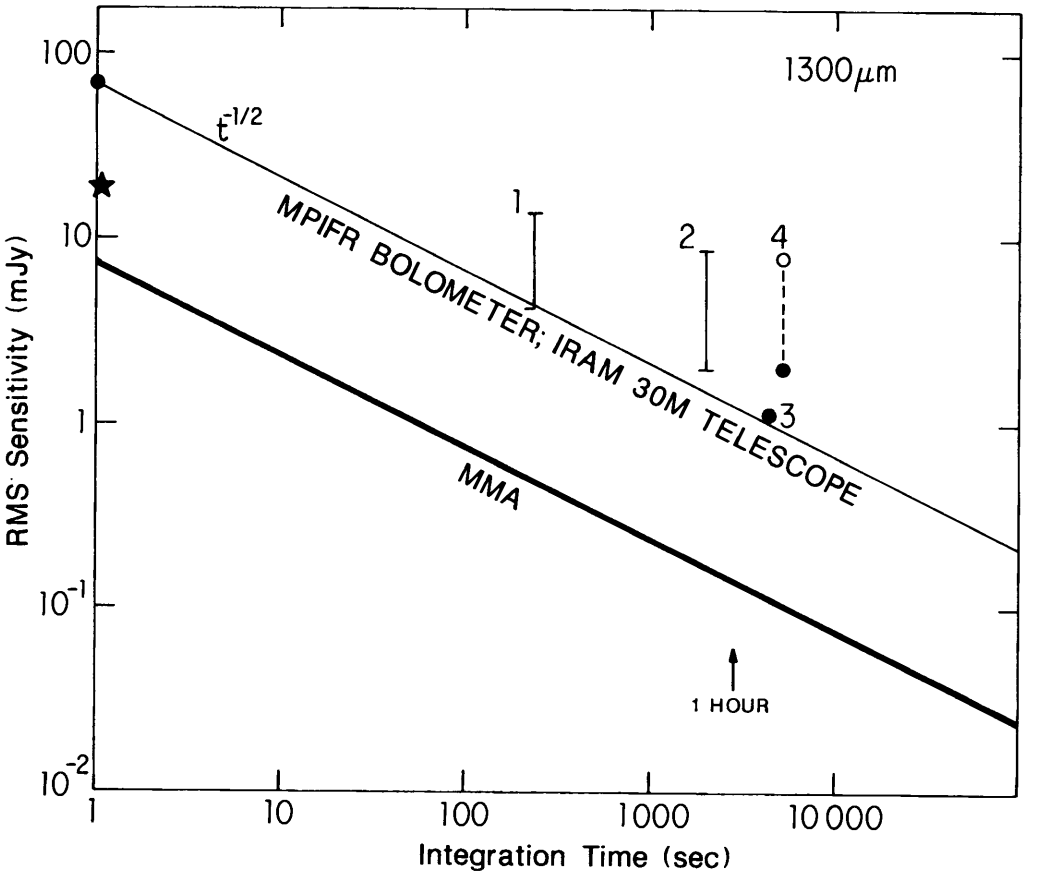


Figure B-1. The rms sensitivity achieved with the MPIfR 1.3-mm bolometer [Kreysa *et al.* (1985) *op. cit.*] on the 30-meter IRAM telescope, as a function of integration time. The data are reported by (1) Altenhoff *et al.* (1989) *op. cit.*; (2) E. Krugel, R. Chini, E. Kreysa, and W. A. Sherwood (1988), *Astron. & Astrophys.*, **190**, 47; (3) R. Chini, E. Kreysa, and C. J. Salter (1987), *Astron. & Astrophys.*, **182**, L63; and (4) P. L. Biermann, R. Chini, A. Greybe-Götz, G. Haslam, E. Kreysa, and P. G. Mezger (1990), *Astron. & Astrophys.*, **227**, L21. Two points are plotted for point (4): these data were taken on the SEST 15-m telescope, and the rms reported is the open circle; correcting this point to the sensitivity of the larger aperture of the 30-m IRAM telescope leads to the rms shown by the filled circle. The asterisk on the ordinate indicates the sensitivity reported for the MPIfR bolometer with one-second integration in the absence of sky noise.

The ability to reach an rms sensitivity of 2 mJy on the 30-m telescope with a total-power detector requires the ability to difference the total-power sky contribution to an accuracy of approximately one part in 10^5. This result follows from the following calculation: The sensitivity of the 30-m telescope is approximately 0.20 K/Jy. A 2-mJy rms therefore corresponds to a brightness-temperature fluctuation of 2×10^{-4} K. The temperature of the atmosphere at the zenith for $\tau \approx 0.10$, as reported for the measurements in Figure B-1, is ~ 30 K. Thus the 4×10^{-4} K rms fluctuation is achieved by subtracting the sky to a fractional precision of $(4 \times 10^{-4})/30 \approx 10^{-5}$.

Until techniques are developed to subtract the sky contribution to better than one part in 10^5, the bolometer sensitivity, even on the 30-m telescope, will

be limited to the milli-Jansky level. However, it is at this same milli-Jansky sensitivity level that the MMA continuum sensitivity begins! The MMA 1.3-mm continuum sensitivity is 1 mJy/$\Delta t_{(\mathrm{min})}^{1/2}$. For detection of point sources the MMA is more than 60 times faster than the MPIfR bolometer on the 30-meter telescope, it is 150 times faster at 0.87 mm and, with a day's integration, the MMA can reach a sensitivity two orders of magnitude better than is possible with the total-power techniques.

Examining Figure B-1 and noticing that the bolometer sensitivity reaches an asymptotic value with increasing integration time, one cannot avoid asking whether some of this behavior could be the result of fluctuations in the cosmic background rather than in the terrestrial atmosphere. One can think of two possible sources. The first is simply small-angular-scale fluctuations in the CMB, remnants of the spectrum of gravitational fluctuations at the epoch of recombination. No work which has been done with high angular resolution at millimeter wavelengths bears directly on this possibility. Second, classical radio confusion may be a contributing factor just as it is at centimeter wavelengths. If an appreciable fraction of the 150 L^* galaxies per square arcminute that Tyson and his collaborators see out to $z \approx 7$ (see Section 1 of Chapter III) are luminous sources of heated dust emission, then the IRAM 30-m 1.3-mm beam will include, on average, four such galaxies. Could $\sqrt{N}$-fluctuations of this number contribute to the asymptotic behavior evident in Figure B-1? The MMA will answer this question.

4. Practicalities of MMA 'Single-Dish' Observing. As a developing science, millimeter-wavelength 'single-dish' astronomy has been well served over the past two decades by the following:

- The capability to complete very-large-scale spatial and spectral surveys of the sky;
- Rapid application of new technology to the next generation of observing equipment (e.g., superconducting mixers, 0.3-K bolometers, and array detectors); and
- The operational ease with which new observing techniques could be explored and developed.

As a more mature science, millimeter-wave astronomy needs to realize the capabilities of the Millimeter Array without undue compromise to the experience of the past. The MMA design represents an attempt to achieve a reasonable balance by incorporating as much flexibility as practically possible in each of the components of the array. The natural limitations to synthesis observations on a high-altitude site provide, in themselves, some resources for 'single-dish' observations (see below).

Survey observations require not only sensitive hardware (to achieve adequate speed) but also the commitment of large blocks of telescope time. The forty MMA antennas, each of which is a fully instrumented 'single dish', can be used in one of the modes described above to speed survey observations. In the case of large sky surveys, pointing each of the antennas independently to different sky positions will, of course, increase one's mapping speed by a factor

of forty over that achievable with a single-beam single antenna. But will there be time available on the MMA for such work? Certainly. Because the MMA relies on specific, standard configurations for its imaging fidelity, the times when the array is being reconfigured are inappropriate for synthesis observations but can be used effectively for 'single-dish' work. Similarly, at those times when the MMA is in its largest configuration but the atmospheric stability compromises synthesis work—even though the atmospheric transparency may remain excellent—'single-dish' observational programs will be (dynamically) scheduled. From the site-testing data being gathered by the tipping radiometers, one can estimate the fraction of time likely to be available only to 'single-dish' observations. For each of the sites under active consideration this appears to be a very useful fraction of time.

New technology is notoriously difficult to incorporate into existing synthesis telescopes because, usually, the technology must be replicated many times for an array. In the specific case of the MMA there is greater flexibility. Because each of the antennas may be used independently and each will have two unused receiver bays (Ch. VII), a wide variety of special-purpose user-supplied and Observatory equipment may be installed and tested simultaneously and independently. For development work, equipmental and procedural tests, and many astronomical purposes (especially survey work, bolometry, and array-feed observations), one can anticipate that each of the MMA antennas will have its own unique instrumentation in the two 'spare' receiver bays.

The flexibility to instrument and operate each of the MMA antennas independently and uniquely should serve to encourage and broaden community participation in development of the technology and the techniques that will constitute the foundations of future millimeter-wave astronomy.

APPENDIX C. ACKNOWLEDGMENTS

Many individuals have contributed, in various ways, to the Millimeter Array effort. Below is a list of participants at the series of MMA scientific and technical design workshops which began in the autumn of 1985. We wish to acknowledge their contributions, both past—and future—to the realization and the ultimate success of the Millimeter Array.

We wish also to express particular gratitude to Ms. Patricia J. Smiley and Mr. George L. Kessler, of the NRAO Graphics Department, for their contributions of skillfull drafting and art work and for able assistance in the design and preparation of this proposal and of the MMA Design Study Volumes.

J. T. Armstrong *Naval Research Laboratory*
M. Balister *NRAO*
J. Bally *AT&T Bell Laboratories*
T. M. Bania *Boston University*
T. S. Bastian *NRAO*
G. L. Berge *California Institute of Technology*
J. H. Bieging *University of California, Berkeley*
M. Birkinshaw *Harvard College Observatory*
G. Blake *California Institute of Technology*
L. Blitz *University of Maryland*
E. Brinks *NRAO*
R. L. Brown *NRAO*
J. J. Condon *NRAO*
T. J. Cornwell *NRAO*
W. D. Cotton *NRAO*
P. C. Crane *NRAO*
I. de Pater *University of California, Berkeley*
H. R. Dickel *University of Illinois*
G. A. Dulk *University of Colorado*
N. Duric *University of New Mexico*
J. A. Eilek *New Mexico Institute of Technology*
D. T. Emerson *NRAO*
N. J. Evans *University of Texas*
J. A. Gallagher *Association of Universities for Research in Astronomy*
D. E. Gary *California Institute of Technology*
I. Gatley *National Optical Astronomy Observatory*
F. C. Gillett *Kitt Peak National Observatory*
P. F. Goldsmith *University of Massachusetts*
W. M. Goss *NRAO*
S. T. Gottesman *University of Florida*
M. Guélin *Institut de Radio Astronomie Millimétrique*
W. J. Hack *New Mexico State University*

R. J. Havlen *NRAO*
D. S. Heeschen *NRAO*
E. Herbst *Duke University*
R. H. Hildebrand *University of Chicago*
R. M. Hjellming *NRAO*
P. T. P. Ho *Harvard College*
D. E. Hogg *NRAO*
W. M. Irvine *University of Massachusetts*
M. Ishiguro *Nobeyama Radio Observatory*
P. R. Jewell *NRAO*
A. R. Kerr *NRAO*
G. R. Knapp *Princeton University*
M. R. Kundu *University of Maryland*
M. L. Kutner *Rensselaer Polytechnic Institute*
P. J. Leahy *Jodrell Bank*
H. S. Liszt *NRAO*
K. Y. Lo *University of Illinois*
R. B. Loren *Unaffiliated*
R. J. Maddalena *NRAO*
J. G. Mangum *NRAO/University of Virginia*
H. M. Martin *Steward Observatory*
R. N. Martin *Steward Observatory*
C. R. Masson *Harvard-Smithsonian Center for Astrophysics*
J. M. Moran *Harvard-Smithsonian Center for Astrophysics*
D. Muhleman *California Institute of Technology*
L. G. Mundy *University of Maryland*
P. C. Myers *Smithsonian Astrophysical Observatory*
P. J. Napier *NRAO*
C. A. Norman *Space Telescope Science Institute*
F. N. Owen *NRAO*
P. Palmer *University of Chicago*
R. B. Partridge *Haverford College*
R. A. Perley *NRAO*
W. L. Peters *Steward Observatory*
T. G. Phillips *California Institute of Technology*
R. L. Plambeck *University of California, Berkeley*
S. P. Reynolds *North Carolina State University*
L. J. Rickard *Naval Research Laboratories*
G. Rieke *Steward Observatory*
D. H. Roberts *Brandeis University*
L. R. Rudnick *University of Minnesota*
A. I. Sargent *California Institute of Technology*
F. P. Schloerb *University of Massachusetts*
F. R. Schwab *NRAO*
P. R. Schwartz *Naval Research Laboratories*
N. Z. Scoville *California Institute of Technology*

G. A. Seielstad *NRAO*
E. Silverberg *Smithsonian Astrophysical Observatory*
L. E. Snyder *University of Illinois*
P. Solomon *State University of New York*
R. A. Sramek *NRAO*
A. A. Stark *AT&T Bell Laboratories*
H. Tananbaum *Harvard-Smithsonian Center for Astrophysics*
G. Taylor *NRAO*
B. E. Turner *NRAO*
N. Ukita *Nobeyama Radio Observatory*
J. M. Uson *NRAO*
P. A. Vanden Bout *NRAO*
J. H. van Gorkom *Columbia University*
S. N. Vogel *University of Maryland*
C. M. Wade *NRAO*
P. G. Wannier *Jet Propulsion Laboratory*
A. Webster *James Clerk Maxwell Telescope*
S. Weinreb *Martin-Marietta Laboratories*
W. J. Welch *University of California, Berkeley*
D. J. Westpfahl *New Mexico Institute of Technology*
R. W. Wilson *AT&T Bell Laboratories*
D. O. S. Wood *Harvard-Smithsonian Center for Astrophysics*
D. Woody *California Institute of Technology*
H. A. Wootten *NRAO*
M. C. H. Wright *University of California, Berkeley*
J. M. Wrobel *NRAO*
A. Zensus *NRAO*