Micro-Spec: an ultracompact, high-sensitivity spectrometer for far-infrared and submillimeter astronomy

Giuseppe Cataldo,* Wen-Ting Hsieh, Wei-Chung Huang, S. Harvey Moseley, Thomas R. Stevenson, and Edward J. Wollack

NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, Maryland 20771, USA

*Corresponding author: Giuseppe.Cataldo@nasa.gov

Received 22 July 2013; revised 17 December 2013; accepted 7 January 2014; posted 16 January 2014 (Doc. ID 194299); published 13 February 2014

High-performance, integrated spectrometers operating in the far-infrared and submillimeter ranges promise to be powerful tools for the exploration of the epochs of reionization and initial galaxy formation. These devices, using high-efficiency superconducting transmission lines, can achieve the performance of a meter-scale grating spectrometer in an instrument implemented on a 4 inch silicon wafer. Such a device, when combined with a cryogenic telescope in space, provides an enabling capability for studies of the early universe. Here, the optical design process for Micro-Spec (μ-Spec) is presented, with particular attention given to its two-dimensional diffractive region, where the light of different wavelengths is focused on the different detectors. The method is based on the stigmatization and minimization of the light path function in this bounded region, which results in an optimized geometrical configuration. A point design with an efficiency of ~90% has been developed for initial demonstration and can serve as the basis for future instruments. Design variations on this implementation are also discussed, which can lead to lower efficiencies due to diffractive losses in the multimode region. © 2014 Optical Society of America

OCIS codes: (050.1940) Diffraction; (120.6200) Spectrometers and spectroscopic instrumentation; (220.2740) Geometric optical design; (300.6340) Spectroscopy, infrared.

1. Introduction

Far-infrared (IR) and submillimeter (15 μm–1 mm) spectroscopy provides a powerful tool to probe a wide range of environments in the universe. In the past 30 years, many space-based observatories, such as Spitzer, Herschel, and the Infrared Space Observatory (ISO) [1–3], have opened the far-IR window to the universe, revealing rich line and continuum spectra from objects ranging from interplanetary dust particles to major galactic mergers and young galaxies in the early universe. Discoveries made by these observatories have provided unique insights into physical processes leading to the evolution of the universe and its contents. However, many questions remain unanswered regarding the very first moments of the universe, as well as galactic, stellar, and planetary formation.

Micro-Spec (μ-Spec) is proposed as a technology concept to enable new discoveries in the far-IR spectral range. μ-Spec will be a high-sensitivity, direct-detection spectrometer operating in the 450–1000 μm regime. The light from the telescope is coupled to the spectrometer chip via a broadband slot antenna (Fig. 1) and directed to the power splitters and delay network realized in microstrip transmission lines. These microstrip elements are connected to a planar waveguide multimode region via transitions, which also serve as internal feed apertures. The planar waveguide multimode region has two internal antenna arrays, one for transmitting and one for receiving the spectrum as a function of wavelength. Absorber structures lining the
multimode region are used to control the power emitted into large angles or reflected from the receiver antenna array. Finally, each receive antenna in the multimode region is coupled to a bank of order-sorting filters terminated in microwave kinetic inductance detectors (MKIDs) for readout. The instrument, with a resolution $R \sim 1200$, will be integrated on a 4 inch diameter (100 mm) silicon chip. This reduction in size is largely realized through the use of single-mode microstrip delay lines, which can be relatively compactly routed on the silicon wafer. Relative to counterparts realized in free space, the required transmission line lengths are more compact by a factor of the medium’s effective index. For all these reasons, μ-Spec can become an important instrument capability under the low background conditions realized in balloon and space-borne platforms such as the Space Infrared Telescope for Cosmology and Astrophysics (SPICA) [4]. The spectral range of the implementation studied here is set by the materials employed for the transmission line and detectors. Application of this approach is presently confined to wavelengths $\lambda > 250 \, \mu m$ by the gap frequency of available low-loss superconductors. These include niobium (Nb) and niobium-titanium nitride (NbTiN) for the transmission line structures and molybdenum nitride (MoN) for the detectors.

Two distinct electromagnetic modes are used in the μ-Spec transmission spectrometer: microstrip and parallel-plate waveguide [5] (see Fig. 4, right). Microstrip is a single-mode transmission line structure that is utilized in its dominant quasi-transverse electromagnetic mode (TEM) and in this context is analogous to single-mode fiber optic. It is used to realize the electrical delay line network. The parallel-plate waveguide has a TEM symmetry, and the region serves as a multimode spatial power combiner. Its function can be compared to a grating spectrometer [6], in which a plane wave is reflected from the grating and the phase of each partial wave scattered from the rulings is a linear function of position across the grating. μ-Spec thus differs from similar technologies by the order of processing of the light in the spectrometer. For instance, in a Rowland spectrometer the required phase retardation is generated by reflection from the grating grooves [6–10], whereas in Z-Spec, which is an example of planar Rowland grating architecture, propagation occurs in parallel-plate waveguides [11–14]. A last comparison can be made with one-dimensional bootlace lenses found in microwave practice [15–18], which μ-Spec builds upon for sub-millimeter wave applications.

In this work, the detailed optical design of the μ-Spec transmission spectrometer is presented. We begin in Section 2 with an outline of the multimode power-combiner region optical design, then describe the internal two-dimensional feed structures used to couple from microstrip lines to this multimode region in Section 3, and finish by bringing these elements together to compute the modeled response of the entire antenna array in Section 4.

2. Detailed Spectrometer Design

The performance of the spectrometer is optimized by the selection of the $x$ and $y$ coordinates of each radiator and a differential phase shift on each delay line. In the case of the Rowland spectrometer, two points on the focal surface can be chosen and wavelengths assigned to them [11,12]. The $x$ and $y$ positions of the grooves in the grating can be selected to provide zero phase errors at each of these stigmatic points. A similar constructive approach can be used in the case of μ-Spec, where each emitter has an additional degree of freedom, the phase shift at the radiator, which allows generating a third stigmatic point. In the case of the Rowland grating, the selection of wavelengths at the two stigmatic points determines the dispersion of the grating.

In μ-Spec, the selection of the wavelengths of three stigmatic points determines both the dispersion and the linearity of the dispersion of the system. A constraint on the selection is required if linear dispersion is desired. In the case presented here, the stigmatic points and their associated wavelength...
were chosen to provide relatively linear dispersion and good imaging over the wavelength range of interest. The positions and wavelengths of the stigmatic points are free parameters and can be selected to optimize aspects of the instrument performance. A reasonable choice, which is adopted here, is to select points that minimize the overall root-mean-square (RMS) phase error for all spectrometer outputs (Fig. 3).

In the prototype version of the instrument described here, in order to fit four spectrometers in a 100 mm silicon wafer, each in a 25 mm × 25 mm region, the maximum radius allowed is R = 1.25 cm. This choice is not optimal but enables several designs to be experimentally studied within a single mask set and the development of the required fabrication process. The overall design was therefore carried out with the parameters shown in Table 1, where values for an initial design (A) and its optimized version (B) here adopted as the baseline are shown.

With these parameters, the design procedure described was implemented in MATLAB through the goal attainment method of Gembicki [19]. This method was chosen because its implementation enables the resolution of vectorial equations subject to nonlinear constraints that arise in the case presented in this paper. The following set of equations describes the propagation of light through the spectrometer,

\[ F_m(x) = F_m^e, \quad m = 1, \ldots, 3N_e \]  

(1)

where \( F_m^e \) represents the electrical delay that must be equal to \( F_m \), an integer multiple of the wavelengths, at the three stigmatic points. \( N_e \) represents the number of radiators, and 3\( N_e \) equations are needed to calculate the electrical delay of all radiators at each of the three stigmatic points. The vector of the unknowns, \( x \), consists of the coordinates \( x_i \) and \( y_i \) of the \( N_e \) emitters’ centers and the electrical path lengths in silicon, \( R_i^e \), for each feed horn delay line.

Constraints are necessary to limit the search of the solution in the correct parameter space. In fact, each of the system’s equations is of order 2 and the search of the solution was restricted within \( \lambda_1/40 \) from the geometric optics solution for the grating solution (the 2R dotted arc on the left side of Fig. 2). Using a numerical as opposed to an analytical solution thus allows us to analyze and incorporate the channel isolation and higher-order diffractive response into the optimization (see Section 3). Indeed, while the feed phase centers reside on the position of the analytical solution, it will be shown that the array response is not uniform in angle.

All quantities in Eq. (1) are defined as follows:

\[ F(x) = \begin{bmatrix} R_i^e + R_{1i}^e - R_1^e \\ R_i^e + R_{2i}^e - R_2^e \\ R_i^e + R_{Bi}^e - R_B^e \end{bmatrix} \quad i = 1, \ldots, N_e \]

is the light path function matrix \((3N_e \times 1)\), whose first 2\( N_e \) rows are associated with stigmatic points 1 and 2 respectively, and the last \( N_e \) rows with the blaze point B. All quantities in the \( F(x) \) matrix represent electrical lengths in the transmission medium. For example, the terms in the first \( N_e \) rows associated with stigmatic point 1 are indicated below:

- \( R_i^e \) is the electrical path length of the \( i \)-th silicon microstrip delay line;
- \( R_{1i}^e = \sqrt{(R \sin \beta_1 - x_i)^2 + (R \cos \beta_1 - y_i)^2} \times \sqrt{\mu_r} \) is the electrical path length of the \( i \)-th feed horn from stigmatic point 1 in the silicon parallel-plate waveguide multimode region;

![Fig. 2. Simplified representation of the grating geometry. On the left side three radiators can be seen, which point to the blaze point, B. The path from each radiator’s phase center to the first and second stigmatic points are also indicated by solid lines for the array’s \( i \)-th and center reference feed.](image-url)
A particular solution is found when placing the blaze point on the free-space waveguide lengths of the three stigmatic points. Here, the problem reduces to solving an ordinary differential equation for the pair of symmetric off-axis grating points, with the radiation lengths of the three stigmatic points. Here, the problem reduces to solving an ordinary differential equation for the pair of symmetric off-axis grating points, with the radiation lengths of the three stigmatic points.

Having selected the phase centers of the grating, we proceed to the radiators near the center of the radiating array, and their wave-plates (section 4). As mentioned in the beginning of this section, the antenna equations can be found by restricting the solutions to the antenna positions. In the general case, a change of 13 db can be achieved by various methods. For example, a design example is given by Rotman and Turner [15], where the free-space waveguide lengths of the three stigmatic points. Here, the problem reduces to solving an ordinary differential equation for the pair of symmetric off-axis grating points, with the radiation lengths of the three stigmatic points.

Now, we choose an adjacent radiator and place the blaze point on the free-space waveguide lengths of the three stigmatic points. Here, the problem reduces to solving an ordinary differential equation for the pair of symmetric off-axis grating points, with the radiation lengths of the three stigmatic points. Here, the problem reduces to solving an ordinary differential equation for the pair of symmetric off-axis grating points, with the radiation lengths of the three stigmatic points.

The optical axis. This provides a partial diagonalization of the problem, as in [15]. Under these hypothesis, the last N equations of Eq. (1), which refer to the blaze point, can be explicitly rewritten as shown below:

\[
\begin{align*}
\sqrt{R^2 - \lambda^2} &\leq \lambda_j/40 \\
\left(\sqrt{R^2 - \lambda^2} - (2R + 2)\right) &\leq \lambda_j/40,
\end{align*}
\]

A 0.45 mm thick monocrystalline silicon substrate is employed on the chip. The electrical length of the central feed horn (12) is equal to 2R. Also, \( \alpha = \frac{\lambda}{2} \) is the linewidth of the grating circle needed in order to satisfy the stigmatic condition, which is a linear relation between the electric length in silicon, \( R_1 \), and the free-space waveguide lengths. The geometric parameter is the electric length in silicon, \( R_1 \), and the free-space waveguide lengths. The geometric parameter is the electric length in silicon, \( R_1 \), and the free-space waveguide lengths.
perturb its position and differential phase, which results in arriving at each of the three stigmatic points. This procedure is continued for the radiators on each of the delay lines. This results in zero phase error at the stigmatic points by construction, low RMS phase errors between the stigmatic points, and diffraction-limited performance in the multimode region over the design spectral range. The RMS phase error, \( \varphi_{\text{RMS}} \), is illustrated in Fig. 3 for design A, and it was calculated by assuming a uniform weight for each radiator at each point on the focal plane:

\[
\varphi_{\text{RMS}}(\theta_j) = \left[ \frac{1}{N_e} \sum_{i=1}^{N_e} (\varphi_{ij} - \langle \varphi(\theta_j) \rangle)^2 \right]^{1/2}.
\]  

Here, \( \varphi_{ij} \) is the relative phase of each transmitter, \( \langle \varphi(\theta_j) \rangle = 0 \) is the relative phase of the central transmitter (this is zero by construction as the central radiator is used as a reference, i.e., \( F_i^{(1)}(N_e + 1/2) \equiv 0 \)), \( N_e \) is the number of transmitters, and \( \theta \) represents the angle corresponding to each of the points in which the focal plane was discretized. The phase between the \( i \)-th transmitter and the \( j \)-th point on the receiver focal plane is defined as \( \varphi_{ij} = 2\pi \mu_y/h_{ij} \times F_i \) and \( h_{ij} \) is the free-space wavelength.

3. Antenna Feed Response

Conceptually, each microstrip feed structure is a two-dimensional analog of an adiabatic feed horn. To achieve the desired performance, the feed needs to match impedance, modal shape, and angular resolution over the design frequency band. An adiabatic impedance taper [21] for the input microstrip line width to the output is achieved via a Hecken taper [22], which is a near-optimum matching section in which the phase is zero by design as the central radiator is used as a reference, i.e., \( F_i^{(1)}(N_e + 1/2) \equiv 0 \), and \( N_e \) is the number of transmitters. The influence of higher-order diffraction is mitigated through the feed apodization, the coupling geometry of arrays, and the use of absorbing boundaries on the multimode region walls.

4. Antenna Array Response

An efficient and accurate expression is desired for the evaluation of the structure’s response during numerical optimization. To evaluate the antenna array’s response, a classical phased array analysis approach is used. The emitters and receivers in the multimode region can be considered as a pair of opposing two-dimensional antenna arrays. By considering the Poynting vector as a function of spatial dimension, \( d \), the following expression for the electric field is adopted for the computation of the array’s response:

\[
E(r, \theta) = \sum_{i=1}^{N_e} \frac{E_i(\theta)}{1 + \left( \frac{\|r\|}{w} \right)^2(d-1/4)} \cdot e^{(ikr-\varphi_{ij})}.
\]
illustrates the power distribution at Eq. (8) were validated by high-frequency structure simulation (HFSS) finite-element simulations of the structure's response with distance. These simulations were used to derive the feed aperture's effective width, \( w \), and to determine the location of the phase center for radiation emitted by the structure.

The power is then computed from

\[
P(r, \theta) = \frac{1}{2\eta} |E|^2 \cdot E(r, \theta), \quad \eta = \left(\frac{\mu_r}{\varepsilon_r}\right)^{1/2} \eta_0, \quad (9)
\]

where \( \eta_0 = 377\Omega \) is the wave impedance of free space. Figure 5 illustrates the power distribution at 450 GHz, as was calculated from Eqs. (9) and (8).

On the left side, it can be seen that the power peaks at the first stigmatic point as required, and then it quickly drops to zero along the rest of the focal plane. At separations from the feed elements greater than \( \sim 2 \text{ mm} \), the field transitions to the far field and the electric field amplitude is observed to asymptotically fall as \( |r|^{-1/2} \). Simulations show the power is conserved throughout the multimode region. In design example A, the ratio of the power emitted by the feed horns to the power received by the antennas is equal to \( \sim 30\% \). The feed response that is not intercepted by the antenna arrays is terminated by the absorber structures on the sides of the multimode region. This also enables any residual reflected power to be terminated in a controlled manner. In addition, a second-order diffraction peak, visible in the lower lefthand quadrant near the emitters, is responsible for part of this loss.

Improved performance can be achieved with different trade-offs between the antennas and the array geometry, which increase the throughput and reduce the effects of diffraction (see Table 1, design B). If the two stigmatic points are moved away from the blaze...
point on the optical axis along the Rowland circle by 4.2%, the instrument reaches its highest efficiency of 87% when the following applies: The radius is decreased by 16% \((R = 1.05 \text{ cm})\), the number of receivers is increased by 18 \((N_r = 65)\), and the emitter pitch, \(p\), is set to 161 μm. The resulting increase in the receiver array acceptance angle is better matched to capture the transmitted beams given the apodization function of the radiators. This is the primary reason for the multimode coupling-efficiency increase observed in this example. As an aside, this change in array geometry reduces the magnitude of the second-order diffraction peak and moves it out of the range of angles of interest. For additional detail, see Fig. 5, rightmost inset.

Figure 6(a) shows the emitter antenna array’s normalized power response evaluated at the three stigmatic wavelengths for the optimized multimode
region design. Table 2 summarizes the coupling efficiency between receiver and transmitter antenna arrays for the two design variations considered in this section. The interchannel power isolation is presented in Fig. 6(b) for the center channel of the optimized design. The observed finite isolation, <0.48, arises from the design goal to fully sample the spectrum and maximize the structure’s optical throughput. The resulting signal correlation can be addressed in calibration of the spectrometer response. Finally, the RMS phase error now reaches a maximum of 0.011 at β = 75.8° and 104.3° and vanishes at the three stigmatic points (β = 67.8°, 90° and 112.2°).

5. Conclusions

A theory and design methodology has been described for μ-Spec, a superconducting microstrip-based direct-detection spectrometer, in which the phase retardation is generated by propagation through transmission lines of different length. A proof-of-concept design was realized to demonstrate the feasibility of the instrument. The location of the three stigmatic points was chosen to minimize the overall RMS phase error for all spectrometer outputs, whereas the location of both radiators and receivers was found by setting the light path function equal to integer multiples of the three wavelengths associated with the three stigmatic points. The structure’s response was evaluated through an E-field model validated by finite-element simulations. The results show that, through a geometric optimization of the instrument, a maximum throughput of ~90% can be achieved in the multimode region. The attributes of this design approach, namely reduced spectrometer envelope and high throughput, have the potential to enable a new class of space-borne instrumentation that will reveal new details of the early universe.

Financial support received from the NASA ROSES/APRA program is gratefully acknowledged by the authors. GC would also like to acknowledge the support provided by Universities Space Research Association (USRA) in the administration of his appointment at the NASA Goddard Space Flight Center.

References


