Micro-Spec: an integrated direct-detection spectrometer for far-infrared space telescopes

Giuseppe Cataldo\textsuperscript{a,\textasteriskcentered}, Wen-Ting Hsieh\textsuperscript{a}, Wei-Chung Huang\textsuperscript{a}, S. Harvey Moseley\textsuperscript{a}, Thomas R. Stevenson\textsuperscript{a}, and Edward J. Wollack\textsuperscript{a}

\textsuperscript{a} NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, USA;
\textsuperscript{b} Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, USA

ABSTRACT

The far-infrared and submillimeter portions of the electromagnetic spectrum provide a unique view of the astrophysical processes present in the early universe. Our ability to fully explore this rich spectral region has been limited, however, by the size and cost of the cryogenic spectrometers required to carry out such measurements. Micro-Spec (\(\mu\)-Spec) is a high-sensitivity, direct-detection spectrometer concept working in the 450–1000 \(\mu\)m wavelength range which will enable a wide range of flight missions that would otherwise be challenging due to the large size of current instruments with the required spectral resolution and sensitivity. The spectrometer design utilizes two internal antenna arrays, one for transmitting and one for receiving, superconducting microstrip transmission lines for power division and phase delay, and an array of microwave kinetic inductance detectors (MKIDs) to achieve these goals. The instrument will be integrated on a \(\sim\) 10 cm\(^2\) silicon chip and can therefore become an important capability under the low background conditions accessible via space and high-altitude borne platforms. In this paper, an optical design methodology for \(\mu\)-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 maximization of the instrument resolving power and minimization of the RMS phase error on the instrument focal plane. This two-step optimization can generate geometrical configurations given specific requirements on spectrometer size, operating spectral range and performance. Two point designs with resolving power of 260 and 520 and an RMS phase error less than \(\sim\) 0.004 radians were developed for initial demonstration and will be the basis of future instruments with resolving power up to about 1200.

Keywords: Far-infrared, submillimeter, spectroscopy

1. INTRODUCTION

Far-infrared (IR) and submillimeter (15 \(\mu\)m-1 mm) spectroscopy provides a powerful tool to probe a wide range of environments in the universe. In the past thirty years, discoveries made by several space-based observatories have provided unique insights into physical processes leading to the evolution of the universe and its contents. This information is encoded in a wide range of molecular lines and fine structure lines. Observations of such spectral lines enable the exploration of galaxies at high redshifts. The fine structure lines of abundant elements (C, N, and O) allow tracing the obscured star formation and AGN activity into the high-redshift universe. One can measure galaxy redshifts and determine their elemental abundances and physical conditions out to redshifts of \(z > 5\). In spite of this, a number of questions remain unanswered regarding the very early steps of the universe as well as galactic, stellar and planetary formation. The ability to fully explore this rich spectral region has been limited by the size and cost of the cryogenic spectrometers required to carry out these measurements. The work proposed here specifically addresses the need for integrated spectrometers and ultra-low-noise, far-IR, direct detectors, whose specific requirements are shown in Table 1\textsuperscript{1} and compared against the current state of the art.

To contribute to realizing the goals outlined in Table 1 and fill in the gaps in the current state of the art, a high-performance integrated spectrometer module, Micro-Spec (\(\mu\)-Spec), operating in the 450–1000 \(\mu\)m (300–650 GHz) range is proposed. With \(\mu\)-Spec (Fig. 1), the incoming radiation collected by the telescope is coupled to the spectrometer via a broadband slot antenna and directed to a series of power splitters and a delay network.

Further author information: (Send correspondence to G.C.)
G.C.: E-mail: Giuseppe.Cataldo@nasa.gov, Telephone: +1 301 286 7497
Table 1: Summary of far-IR spectrometer and detector array requirements and comparison with current state of the art.

<table>
<thead>
<tr>
<th>Technology</th>
<th>Metric</th>
<th>Symbol</th>
<th>State of the art</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Far-IR detector arrays</td>
<td>Wavelength range</td>
<td>$\lambda$</td>
<td>250 – 700 $\mu$m</td>
<td>220 – 2000 $\mu$m</td>
</tr>
<tr>
<td></td>
<td>Sensitivity</td>
<td>NEP</td>
<td>$10^{-19}$ W/Hz</td>
<td>$3 \times 10^{-21}$ W/Hz</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
<td>$R$</td>
<td>$\geq 100$</td>
<td>$\geq 1200$</td>
</tr>
<tr>
<td></td>
<td>Detective Quantum Efficiency</td>
<td>DQE</td>
<td>$\sim 15%$</td>
<td>$&gt; 90%$</td>
</tr>
<tr>
<td></td>
<td>Format</td>
<td></td>
<td>30 x 30, Component test</td>
<td>1500 spectral elements</td>
</tr>
<tr>
<td></td>
<td>Temperature</td>
<td>$T$</td>
<td>$&lt; 1$ K</td>
<td>$\sim 0.05$ K</td>
</tr>
<tr>
<td></td>
<td>Time constant</td>
<td>$\tau$</td>
<td>100 ms</td>
<td>&lt; 10 ms</td>
</tr>
</tbody>
</table>

made of superconducting microstrip transmission lines. The delay network creates a retardation across the input to a planar waveguide multimode region, which has two internal antenna arrays, one for transmitting and one for receiving the radiation as a function of wavelength. Absorber structures lining the multimode region control the power emitted into large angles or reflected from the receiver antenna array. An array of feed structures is employed to couple the radiation to the multimode region and concentrates the power along the focal surface with different wavelengths at different locations. The outputs are connected to a bank of order-sorting filters which terminate the power in an array of microwave kinetic inductance detectors (MKIDs) for detection and readout. The entire instrument is integrated on a 100-mm-diameter silicon chip. This size in reduction is accomplished by using single-mode microstrip delay lines, which can compactly be folded on the silicon wafer by a factor of the medium’s effective refractive index.

![Figure 1: Layout of the $\mu$-Spec module. The radiation is coupled into the instrument through a broadband antenna and is then transmitted through a superconducting transmission line to a divider and a phase delay network. The spectrum enters the multimode region through an array of feeds which concentrates the power along the focal surface with different wavelengths at different locations. The receivers are connected to a bank of order-sorting filters and MKID detectors.](image)

It is worth noting that the frequency range of the implementation presented here is limited to wavelengths $\lambda > 250$ $\mu$m by the gap frequency of currently 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.

Finally, $\mu$-Spec can be compared to a grating spectrometer, 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. However, it 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, whereas in Z-Spec, which is an example of planar Rowland grating architecture, propagation occurs in parallel-plate waveguides. A last comparison can be made with one-dimensional bootlace lenses.
found in microwave practice,\textsuperscript{12-15} which \(\mu\)-Spec builds upon for submillimeter wave applications. This paper will describe the design process of the multimode region and illustrate the results in terms of geometry, imaging quality and efficiency.

2. MULTIMODE REGION DESIGN IN HIGHER ORDER

A prototype version with a resolving power \( R = 65 \) in first order was designed\textsuperscript{2} and built, and is currently under evaluation at NASA’s Goddard Space Flight Center. This design was specified by requiring the phase error to vanish at three specific stigmatic points on the focal plane. This can be seen as a generalization of the approach described in Ref. 8 with two stigmatic points. Here, we present designs for \( R = 260 \) (Design A) and \( R = 520 \) (Design B) in higher order with zero phase error at three preselected stigmatic points employing a multi-step optimization process described in the following section.

2.1 Problem formulation

As explained in Ref. 2, the design variables are the \( x \) and \( y \) coordinates of the \( N_e \) emitters’ centers and the electrical path lengths in silicon, \( R_e^i \), for each feed’s electrical delay. The resolving power is defined as

\[
R = M \cdot N_e, \tag{1}
\]

where \( M \) is the order of the spectrometer and \( N_e \) is some power of 2, given the structure of the power divider network (see Fig. 1). The first step of the design consists of finding the maximum achievable resolving power, \( R_{\text{max}} \), as a function of \( M \) and the relative emitter pitch, \( \eta = p/\lambda_B \) (\( p = \) emitter pitch, \( \lambda_B = \) wavelength associated with the central stigmatic point on the \( x \) axis - see Fig. 2), given specific requirements on spectrometer radius, \( R \), and operating spectral range as well as certain constraints on performance.

Figure 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 central reference feed.
The formulation of this mixed integer non-linear problem is as follows:

\[
\max \mathcal{R}_{\text{max}}(M, \eta) = \frac{M}{\eta} \cdot \frac{R}{\lambda_B} 
\]

subject to

\[
H_e(M, \eta) \leq R \quad \text{(width of emitter array \leq radius, so that aberration is OK)} \quad (3)
\]

\[
H_r(M, \eta) \leq R \quad \text{(width of receiver array \leq radius, so that aberration is OK)} \quad (4)
\]

\[
\mathcal{R}_{\text{max}}(M, \eta) > 64 \quad \text{(maximum resolving power > 64)} \quad (5)
\]

\[
p(M, \eta) = s(M, \eta) \quad \text{(emitter pitch equal to receiver pitch)} \quad (6)
\]

\[
\eta > 0 \quad (7)
\]

\[
1 \leq M \leq 10, \quad M \text{ integer} \quad (8)
\]

Here we note that constraint (6) was used to simplify the positioning of the receivers on the focal plane and may be removed in future designs. The objective spaces as a function of \( M \) and \( \eta \) for the two cases, \( \mathcal{R} = 260 \) and \( \mathcal{R} = 520 \), are shown in Figs. 3-4, respectively. On the left (Figs. 3a-4a), it is possible to visualize the feasible objective space of the optimization problem described above for \( \mathcal{R}_{\text{max}} \) along with the active constraints, Eq. (4) (blue area) and Eq. (6) (black lines). The optimal solution to Problem (2) is thus given by the intersection of the contour plot representing the feasible values of \( \mathcal{R}_{\text{max}} \) and constraint (6) (lower black line). Table 2 shows the requirements on spectrometer size and spectral range for the two cases. The minimum and maximum frequencies are associated with stigmatic points 2 and 1 respectively, whereas the average frequency, computed as their geometric mean, is associated with the blaze point, \( B \) (Fig. 2). Table 3 shows the values of the design variables associated with the optimal solutions as well as the values of the constraints, which are all satisfied.

Once this problem solved, it is then possible to compute \( N_e = \mathcal{R}_{\text{max}}/M \) and round this down to a power of 2. According to Eq. (1), this causes the actual resolving power, calculated with this updated value of \( N_e \), to be lower than \( \mathcal{R}_{\text{max}} \). The plots in Figs. 3b-4b show the values of the actual resolving power \( \mathcal{R} \) (red line) and the values of \( M \) and \( \eta \) that would make such realizations possible.

### Table 2: Requirements on spectrometer size and spectral range

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Unit</th>
<th>Value (Design A)</th>
<th>Value (Design B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multimode region radius</td>
<td>( R )</td>
<td>cm</td>
<td>1.25</td>
<td>1.25</td>
</tr>
<tr>
<td>Minimum frequency</td>
<td>( f_{\text{min}} )</td>
<td>GHz</td>
<td>570.0 (in order ( M' = 1 ))</td>
<td>605.0 (in order ( M' = 4 ))</td>
</tr>
<tr>
<td>Maximum frequency</td>
<td>( f_{\text{max}} )</td>
<td>GHz</td>
<td>650.0 (in order ( M' = 1 ))</td>
<td>650.0 (in order ( M' = 4 ))</td>
</tr>
<tr>
<td>Average frequency</td>
<td>( f_{\text{avg}} )</td>
<td>GHz</td>
<td>608.7 (in order ( M' = 1 ))</td>
<td>627.1 (in order ( M' = 4 ))</td>
</tr>
<tr>
<td>Minimum frequency</td>
<td>( f_{\text{min}} )</td>
<td>GHz</td>
<td>142.5 (in order ( M' = 4 ))</td>
<td>302.5 (in order ( M' = 8 ))</td>
</tr>
<tr>
<td>Maximum frequency</td>
<td>( f_{\text{max}} )</td>
<td>GHz</td>
<td>162.5 (in order ( M' = 4 ))</td>
<td>325.0 (in order ( M' = 8 ))</td>
</tr>
<tr>
<td>Average frequency</td>
<td>( f_{\text{avg}} )</td>
<td>GHz</td>
<td>152.2 (in order ( M' = 4 ))</td>
<td>313.5 (in order ( M' = 8 ))</td>
</tr>
</tbody>
</table>

### Table 3: Optimal solutions of Problem (2) for designs A and B

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Unit</th>
<th>Value (Design A)</th>
<th>Value (Design B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrometer order</td>
<td>( M )</td>
<td>-</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Emitter diameter</td>
<td>( \eta )</td>
<td>-</td>
<td>1.1634</td>
<td>1.1809</td>
</tr>
<tr>
<td>Maximum resolving power</td>
<td>( \mathcal{R}_{\text{max}} )</td>
<td>-</td>
<td>297.83</td>
<td>604.60</td>
</tr>
<tr>
<td>Emitter array width</td>
<td>( H_e )</td>
<td>cm</td>
<td>1.0912</td>
<td>1.0751</td>
</tr>
<tr>
<td>Receiver array width</td>
<td>( H_r )</td>
<td>cm</td>
<td>1.1297</td>
<td>1.2154</td>
</tr>
<tr>
<td>Emitter/receiver pitch</td>
<td>( p = s )</td>
<td>cm</td>
<td>0.0165</td>
<td>0.0163</td>
</tr>
<tr>
<td>Number of emitters</td>
<td>( N_e )</td>
<td>-</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td>Number of receivers</td>
<td>( N_r )</td>
<td>-</td>
<td>87</td>
<td>91</td>
</tr>
<tr>
<td>Resolving power</td>
<td>( \mathcal{R} )</td>
<td>-</td>
<td>260</td>
<td>520</td>
</tr>
</tbody>
</table>
1 The contour plot represents all feasible values of $R_{\text{max}}$ and shows a solution exists for $M = 4$ that satisfies all constraints.

Figure 3: Objective spaces of Problem (2) for Design A. The active constraints are Eq. (4) (blue area) and Eq. (6) (black lines).

(b) The contour plot represents the feasible values of $R$ for all orders given as powers of 2. The $R = 260$ red curve identifies the optimal solution.

Figure 4: Objective spaces of Problem (2) for Design B. The active constraints are Eq. (4) (blue area) and Eq. (6) (black lines).
The second step consists in minimizing the overall RMS phase error, $\phi_{RMS}$, on the focal plane in order to determine the optimal solution in terms of the above-mentioned design variables.

$$
\min \int_{\theta_1}^{\theta_2} \phi_{RMS} \ d\theta = \int_{\theta_1}^{\theta_2} \sqrt{\frac{1}{N_e} \sum_{i=1}^{N_e} [\phi_{ij}(x_i, y_i, \tau_i, \theta_j) - \langle \phi(\theta_j) \rangle]^2} \ d\theta
$$

subject to

$$
\begin{cases}
\phi_{i1}(x_i, y_i, R_e, \theta_1) = 0 \\
\phi_{i2}(x_i, y_i, R_e, \theta_2) = 0, \quad i = 1, ..., N_e \\
\phi_{iB}(x_i, y_i, R_e, \theta_B) = 0
\end{cases}
$$

Here, $\phi_{ij}$ is the relative phase of each transmitter, $\langle \phi(\theta_j) \rangle = 0$ is the relative phase of the central transmitter (this is zero by construction as the central radiator is used as a reference) and $\theta_j$ represents the angle corresponding to each of the points in which the focal plane was discretized. $\theta_1$, $\theta_2$ and $\theta_B$ are associated with the three stigmatic points as depicted in Fig. 2.

### 2.2 Optimization results

The solutions to the system of equations (9)-(10) found for the two cases discussed in this paper can be seen in Fig. 5. The emitters' positions are indicated in red and present several characteristics which are worth discussing. First, they do not lie on the grating circle but on a curve that is tilted leftwards and intersects the grating circle at the central emitter before ending up inside the multimode region. It was verified that the shape of this curve only approximates a circle with a radius $\sim 2.2R$, is not symmetric and is caused by the imposition of zero RMS phase error on the blaze point, $B$ (last constraint in Eqs. (10)). It was verified that if such a constraint is removed, the emitters lie exactly on the grating circle, as shown in the literature.\textsuperscript{3, 8} In our previous work,\textsuperscript{2} this tilting effect was explicitly avoided by constraining all the emitters to lie within $\lambda_1/8$ from the $2R$ circle. Figure 5b shows also that it is less accentuated for the configuration in order $M = 8$.

The RMS phase error is shown in Fig. 6 for the two cases presented in this paper and compared with that computed for the optimized $R = 65$ version described in Ref. 2. The new designs represent an improvement of no less than $\sim 64\%$ over the nominal spectral range indicated in Table 2. While the RMS phase error vanishes at the three imposed stigmatic points, it is worth noting the presence of a fourth stigmatic point for both designs at $\sim 130^\circ$ and $\sim 155^\circ$. It can be shown that, in general, the RMS phase error as defined in Eq. (9), is a fourth-order function of the focal plane angle, $\theta$, and thus admits a maximum of 4 zeros. Here, equality constraints such as those in Eqs. (10) were only imposed on three points, leaving the fourth free because beyond the spectral
range of interest. As a consequence, this does not affect the overall quality of the instrument for the specified requirements. Exploration of a configuration with 4 stigmatic points spanning a desired frequency range and such to minimize the RMS phase error as defined in Problem (9) is left for future effort.

![RMS phase error figure](image)

**Figure 6**: RMS phase error for the two designs discussed in this paper and compared to the optimized $R=65$ version presented in Ref. 2.

### 2.3 Power efficiency

The power coupling in the new design configurations was computed with the model described in Ref. 2, Sec. 4. Figure 7 illustrates the values of the normalized power for the $R=260$ run in first, second and fourth order, as well as for the $R=520$ run through eighth order. When comparing the coupling efficiency of these designs in first order with that described in Ref. 2, it can be seen that the $R=260$ version outperforms the others, though still inferior to the optimized $R=65$ version. This low efficiency is the result of higher-order diffraction peaks that show up in the multimode region and are caused by $\eta$ being greater than 1/2, as required to avoid diffraction.\(^{16}\)

### 3. CONCLUSIONS

A design methodology has been developed for high-resolution configurations of the $\mu$-Spec multimode region. The design procedure first maximizes the resolving power subject to constraints on geometry, operating frequency range and performance, thereby determining the order of the spectrometer. This then allows the RMS phase error on the instrument focal plane to be minimized and vanish at three preselected stigmatic points associated with three specific wavelengths. The two designs explored here achieve resolutions of 260 and 520 in fourth and eighth order respectively and present a coupling efficiency approaching unity. The coupling efficiency was observed to decrease monotonically with the order. The RMS phase error between the three stigmatic points is less than 0.005 radians for both designs and the presence of a fourth stigmatic point was observed beyond the angular range in use. This feature can be used in future work to increase the number of spectrometer channels and resolving power.
Figure 7: Comparison of the efficiency of the two designs presented in this paper and the ones from Ref. 2.

ACKNOWLEDGMENTS

Financial support received from the NASA ROSES/APRA program and the Massachusetts Institute of Technology “Arthur Gelb” fellowship is gratefully acknowledged by the authors.

REFERENCES


