

Simulations of the spectral resolving power of a compact space-borne immersion-echelle spectrometer using mid-infrared wave tracing

Satoshi Itoh^{Ⓞ, a,*}, Daisuke Ishihara,^a Takehiko Wada^{Ⓞ, a},
Takao Nakagawa^{Ⓞ, a}, Shinki Oyabu^{Ⓞ, b}, Hidehiro Kaneda,^c
Yasuhiro Hirahara,^d and the SMI consortium

^aJapan Aerospace Exploration Agency, Institute of Space and Astronautical Science,
Department of Space Astronomy and Astrophysics, Sagamihara, Japan

^bTokushima University, Institute of Liberal Arts and Sciences, Tokushima, Japan

^cNagoya University, Graduate School of Science, Nagoya, Japan

^dNagoya University, Graduate School of Environmental Studies, Nagoya, Japan

Abstract. We performed wave-optics-based numerical simulations at mid-infrared wavelengths to investigate how the presence or absence of entrance slits and optical aberrations affect the spectral resolving power R of a compact, high-spectral-resolving-power spectrometer containing an immersion-echelle grating. We tested three cases of telescope aberration (aberration-free, astigmatism, and spherical aberration), assuming the aberration budget of the Space Infrared Telescope for Cosmology and Astrophysics, which has a $20\ \mu\text{m}$ wavelength diffraction limit. In cases with a slit, we found that the value of R at around 10 to $20\ \mu\text{m}$ is approximately independent of the assumed aberrations, which is significantly different from the prediction of geometrical optics. Our results also indicate that diffraction from the slit improves R by enlarging the effective illuminated area on the grating window and that this improvement decreases at short wavelengths. For the slit-less cases, we found that the impact of aberrations on R can be roughly estimated using the Strehl ratio. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JATIS.8.2.025004](https://doi.org/10.1117/1.JATIS.8.2.025004)]

Keywords: spectrograph; high-dispersion; space-borne; mid-infrared; spectral resolving power; wave optics.

Paper 21084 received Aug. 6, 2021; accepted for publication Jun. 3, 2022; published online Jun. 23, 2022.

1 Introduction

Space-borne spectroscopy with a spectral resolving power higher than tens of thousands at mid-infrared (MIR) wavelengths is an unexplored region of astronomy that has high scientific significance. For example, radial-velocity measurements with a precision of about 10 km/s can identify the position of the “snowline” in a protoplanetary disk, which is considered a key test of current models for planetary formation.^{1,2} The Space Infrared Telescope for Cosmology and Astrophysics (SPICA) includes the SPICA mid-infrared instrument (SMI),^{3,4} which contains a high-resolution (HR) spectrometer designed to be capable of such high-spectral-resolving-power spectroscopy.

In SPICA/SMI-HR, we planned to use an immersion echelle grating. An echelle grating is a grating that has a low groove density and is optimized for large-angle incident angles (i.e., large diffraction orders). In addition, in an immersion echelle grating, an echelle grating is in touch with a substrate through which the incident and diffracted beams pass. With a given incident angle, a large beam width incident to the grating provides us with a large spectral resolving power, because it brings a large optical-path-length differences on the grating. The optical-path-length difference is magnified by a refractive index of a substrate in an immersion echelle grating. Hence, a beam width required for a given spectral resolving power in an immersion echelle grating is smaller than one in usual echelle gratings by a factor of the refractive index

*Address all correspondence to Satoshi Itoh, sitoh@ir.isas.jaxa.jp

of the substrate. The size reduction of the spectrometer has a particular importance in space-borne instruments that are strictly limited in size and mass.⁵

As a material with a high refractive index, some semiconductors are promising in the infrared wavelength. Crystals of Si, which transmit the light at 1.2 to 7.0 μm , can be applied to immersion echelle gratings through the process of photolithography.^{6–8} The single-point diamond machining^{9,10} of softer materials including CdZnTe is a method to fabricate immersion echelle gratings, which have wavelength band pass different from Si crystals. SPICA/SMI-HR is designed with a substrate of CdZnTe ($n = 2.65$).⁹

Since the size of the spectrometer is limited by cost of space craft or the available ingot size¹⁰ of the substrate material for immersion grating, it is important to obtain detailed information on the spectral resolving power of such a size-limited spectrometer. An analytical evaluation, however, does not consider either diffraction from the spectrometer slit or telescope aberrations, both of which must be taken into account to properly evaluate the spectral resolving power. To include these effects, numerical simulations are needed to acquire information on the spectral resolving power in the MIR-wavelength band without performing difficult experiments. For the following reasons, the simulations must be based not on geometrical optics but on wave optics:

1. Diffraction from the entrance slit of a spectrometer is not negligible. Without a slit, the amplitude/phase distributions of the beam incident on the grating surface are effectively flat, as analysis with geometrical optics assumes. However, with a slit, diffraction produces beams with non-flat amplitude/phase distributions at the grating surface.¹¹
2. In the MIR, imaging performance tends to be close to the diffraction limit; therefore, it is inappropriate to use geometrical-optics-based concepts (e.g., the spot diagram radius) to estimate how the wavefront aberrations of a telescope affect the spectral resolving power.

In this paper, we describe the numerical simulations of the spectral resolving power of the SMI-HR optical model to investigate how the presence or absence of an entrance slit affects the spectral resolving power of a compact, high-spectral-resolving-power spectrometer. We performed the simulations using the wave-optics software Wyrowski VirtualLab Fusion (Second Generation Technology Update [Build 7.3.1.5]).¹² In Sec. 2, we review the principle of spectral resolving power for immersion-echelle spectrometers to clarify the goal of the simulations. In Sec. 3, we describe the assumptions of our simulations. In Sec. 4, we show our results and discuss the general characteristics of the simulated spectral resolving power achieved by a compact, immersion-echelle grating spectrometer. Sec. 5 summarizes the contents of this paper.

2 Review of Theory

The theory of the spectral resolving power of immersion-echelle spectrometers is no different from the usual theory for ordinary reflective gratings that can be found in many sources.¹³ We briefly review it here to clarify the aim of the simulations (stated in Sec. 1) and the symbol conventions used in this paper. Hereafter, we assume that the light source to be observed is an ideal point source.

Equation 1 is the fundamental equation for wavelength dispersion by a reflective grating; it is referred to as the “grating equation,” which is expressed as

$$u(\sin \alpha + \sin \beta) = \frac{m\lambda}{n}, \quad (1)$$

where u is the length of a period (the pitch) of the grating structure; α and β are the angle of incidence and diffraction of the plane wave, respectively; m is the diffraction order; λ is the wavelength of the light in vacuum; and n is the refractive index of the immersed grating substrate. Although the wavelength dispersion of the medium affects the spectral resolving power, in this paper, we assume the refractive index¹⁴ to be a real constant, independent of λ , to focus on the other principal effects. Note that Eq. (1) is derived by assuming the incident wavefront to be a perfect plane wave that spans the entire two-dimensional plane.

The relationship between $d\lambda$ and $d\beta$ is expressed as

$$\frac{nu \cos \beta}{m} d\beta = d\lambda. \quad (2)$$

The variation of $d\lambda/d\beta$ is negligible, compared to $d\lambda/d\beta$ itself, in the interval of the minimum resolvable exit-angle difference (denoted by $\Delta\beta$). Thus, we can interpret Eq. (2) as a sufficiently well-approximated relationship between $\Delta\beta$ and the minimum resolvable wavelength difference (denoted by $\Delta\lambda$) is given by

$$\frac{nu \cos \beta}{m} \Delta\beta = \Delta\lambda. \quad (3)$$

The spectral resolving power R thus becomes

$$R = \frac{\lambda}{\Delta\lambda} = \frac{\lambda m}{nu \cos \beta \Delta\beta} = \frac{\sin \alpha + \sin \beta}{\cos \beta \Delta\beta}. \quad (4)$$

In particular, when $\alpha = \beta$ (the Littrow configuration), Eq. (4) becomes

$$R = \frac{2 \tan \alpha}{\Delta\beta}. \quad (5)$$

Hereafter, we denote the physical size of the grating window along the dispersion direction by W and the geometrical beam width along the dispersion direction by w ; and we assume that $w \leq W$. For the development of SPICA/SMI-HR, the physical size W was limited mainly by the size of the available CdZnTe ingot (e.g., a column with a diameter of 135 mm and a thickness of 50 mm); see Ref. 10 for the detail of the size limitation.

As we are not considering diffraction from slits in this section, we can express $\Delta\beta$ as

$$\Delta\beta = \eta \frac{\lambda}{nw}, \quad (6)$$

where η is a correction factor that depends on the shape of the entrance window; more directly, η depends on how grating grooves are illuminated (e.g., rectangular or elliptical). Substituting Eq. (6) into Eq. (5) leads to the equation for diffraction-limited R , which is given by

$$R = \frac{2n \tan \alpha w}{\eta \lambda}. \quad (7)$$

In Eq. (7), we note that R is directly proportional to the quantity $2n \tan \alpha w$. This quantity can be interpreted as the round-trip optical-path difference between light rays from the opposite ends of the beam width at the incident (exit) window (Fig. 1).

For reference, we also consider the geometrically slit-limited case here. In this case, we can express $\Delta\beta$ using the telescope diameter D and the slit width in radians on the sky ϕ and the related equation is

$$\Delta\beta = \frac{1}{n} \frac{D}{w} \phi, \quad (8)$$

where D/w equals the magnification ratio of the optics. Hence, the geometrically slit-limited R is expressed as

$$R = \frac{2n \tan \alpha w}{D\phi}. \quad (9)$$

Equation (7) does not consider either diffraction from the slit or telescope aberrations, both of which are needed to properly evaluate R . In other words, Eq. (7) is based on the assumptions that the optical system (including the telescope) is aberration-free and that there is no entrance slit.

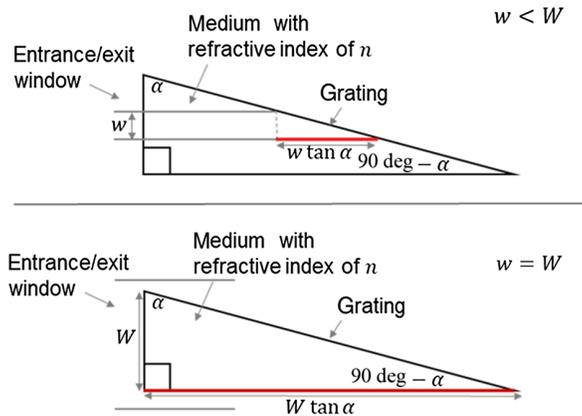


Fig. 1 Illustration of the calculation of the round-trip optical-path difference between light rays from the opposite ends of the beam on the incident (exit) window. The upper panel is for the case $w < W$, and the lower panel is for $w = W$. The lengths of the solid red lines show half of the round-trip path difference between light rays from the opposite ends of the beam on the incident (exit) window.

When we consider diffraction from the entrance slit, the amplitude/phase profile incident on the grating surface becomes far more complicated than the one that analysis with geometrical optics assumes. In addition, diffracted beams may be vignetted at the grating window. Furthermore, aberrations are additional factors that change the size of the point-spread function (PSF). Spot sizes determined by ray tracing are inappropriate for evaluating the sizes of the PSFs when wavefront errors are not sufficiently large compared to the wavelength. Hence, Eq. (7) is not satisfactory for evaluating the spectral resolving power of a realistic spectrometer with an entrance slit and/or optical aberrations.

To take these factors into account, we have performed simulations based on wave optics. The simulations assumed a simple high-spectral-resolving-power spectrograph of limited size ($W \approx w$). We aim to investigate how the presence or absence of an entrance slit and optical aberrations affect R in the 10 to 20 μm wavelength range.

3 Setup

3.1 Layout of the Optical Model

3.1.1 Simplification

The optical model used in the current study is based on the latest design of SMI-HR¹⁵ (Fig. 2), but we have made the following simplifications (Fig. 3):

1. The cross dispersing grating, which is used in the echelle spectrograph, is omitted.
2. We simplified some focusing/collimating systems. The SPICA/SMI is designed with focusing/collimating systems consisting of multiple high-order aspherical mirrors. These multiple higher-order aspherical mirrors well correct off-axis aberrations. This compensation is necessary to obtain good imaging performance over a wide field of view for the focusing at wavelengths other than the blaze wavelength of the Littrow configuration. In the simplified model, we omitted the multiple higher-order aspherical mirrors. The omission of these items does not affect the evaluation of the spectral resolving power at the blaze wavelength of the Littrow configuration.
3. The SMI-HR is designed with off-axis reflective optics, but our calculation model uses co-axial optics with spatial phase modulation to mimic a thin lens of the focal length f . With f being the focal length of the lens and r representing the radial coordinate on the plane of the thin lens, we can express phase modulation by multiplication with the factor $e^{-\frac{2\pi i}{\lambda}(\sqrt{f^2+r^2}-f)}$. We chose the spatial-phase-modulation pattern such that a spherical wave

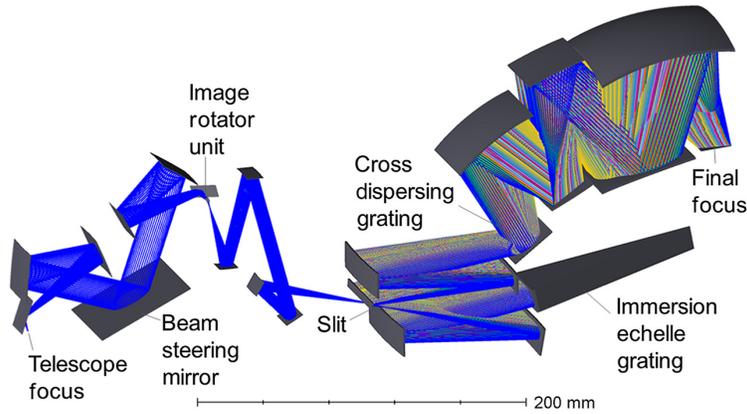


Fig. 2 The optical design of the SPICA/SMI-HR. The rays are colored by the diffraction orders of the diffraction by the immersion echelle grating. The HR spectrometer is one of four subunits that compose the SPICA/SMI.¹⁵

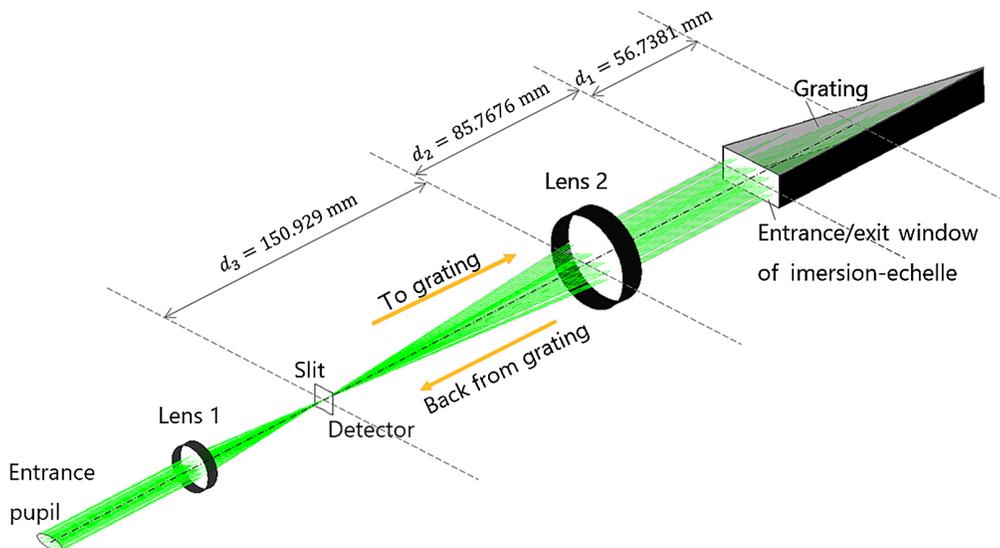


Fig. 3 Optical layout model used for the simulations. The thicknesses of the lenses are exaggerated in this figure to make them easy to see. In the actual simulations, we used spatial phase modulations to represent thin lenses with zero thickness. The surface-to-surface distances d_1 , d_2 , and d_3 shown here have the same definition used in Appendix B; see also Table 1.

from a point source at the front focal point of the thin lens is transformed into a plane wave by the modulation.

3.1.2 Main parameters

The main parameters of the simulation setup are compiled in Table 1. They reflect the design of the SPICA/SMI-HR, except for the simplifications described earlier.

3.2 Method of Calculation

3.2.1 Wavelength sampling

We performed calculations for the diffraction orders $m = 75, 85, 95, 105, 115, 125, 135,$ and 145 . For each sampled value of m , we selected the following wavelengths, for which $\alpha = \beta$ (the Littrow configuration) and the related equation is

Table 1 The main parameters for the setup of the simulations. These parameters come from the design of SPICA/SMI-HR. The reason why the geometry beam and grating window are elongated is to fabricate the longest possible grating from a CdZnTe ingot, which has a limited size. The width (100.000 μm) of the slit are designed through our consideration of the diffraction-limited PSF width at $\lambda = 10$ to $20 \mu\text{m}$ (Table 2).

Parameter	Value
Wavelength of light	$\lambda_m = \frac{2nu \sin \alpha}{m}$
Refractive index of the immersion-echelle grating	$n = 2.65000$
Pitch of the grating	$u = 282.700 \mu\text{m}$
Angle of incidence at the grating surface	$\alpha = 75 \text{ deg}$
Diffraction order	$m = 75, 85, 95, 105, 115, 125, 135, \text{ and } 145$
Numerical aperture (NA) in the dispersion direction	0.0939085
$dI/d\beta$ in Eq. (13)	399.962 mm
Entrance pupil diameter (ellipse)	14.000 mm \times 5.60000 mm
Distance between the entrance pupil and lens 1	80.3000 mm
Diameter of lens 1	20.0000 mm
Thickness of lens 1	0.00000 mm
Distance between lens 1 and slit	71.7112 mm
Aperture size of the slit	100.000 μm \times 720.000 μm
Distance between the slit (or detector) and lens 2	150.929 mm
Diameter of lens 2	40.0000 mm
Thickness of lens 2	0.00000 mm
Distance between lens 2 and the window of the immersion-echelle grating	85.7676 mm
Size of the window of the immersion-echelle grating	31.5000 mm \times 15.0000 mm
Size of the geometrical beam at the grating window	29.4000 mm \times 11.8000 mm
Distance between the window and the grating center	56.7381 mm
Size of the grating surface	117.479 mm \times 15.0000 mm

$$\lambda_m = \frac{2nu \sin \alpha}{m}. \quad (10)$$

These values of λ_m are listed in Table 2. For reference, we also show the slit efficiency E_m at each wavelength, calculated for the assumed slit size $100.000 \mu\text{m} \times 720.000 \mu\text{m}$.

3.2.2 Algorithms

We next describe the propagation algorithms we used. [We used two Intel(R) Xeon(R) processors, each with CPU E5-2620 v4 (2.10 GHz, 8 cores, 16 threads), in parallel for the simulations. The maximum amount of memory used during each simulation was less than 128 GB. The resolution of the simulation on each surface is 4096×4096 .]

1. A periodic dielectric media (such as a grating or lens) causes the propagated field to become periodic. Complex Fourier analysis shows that the propagated periodic field can

Table 2 The wavelengths λ_m used for the simulations and the slit efficiency E_m calculated for the assumed slit size at each wavelength. The wavelengths λ_m are the blaze wavelengths for the Littrow configuration.

m	145	135	125	115	105	95	85	75
λ_m (μm)	9.98108	10.72042	11.57805	12.58484	13.78339	15.23428	17.02654	19.29675
E_m (%)	87.1	86.2	84.9	83.0	80.5	77.2	72.9	67.7

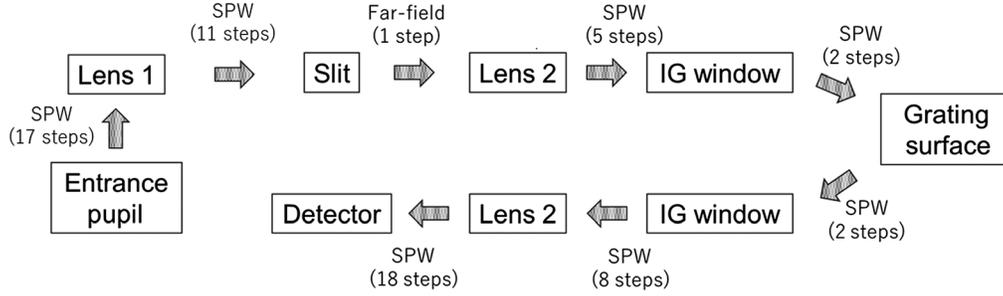


Fig. 4 Summary illustration of the method of calculation. The abbreviation SPW represents a method that uses a spectrum of plane waves, and IG is an abbreviation for immersion-echelle grating. We need the division of the propagation distances to reduce the amount of random-access memory that SPW requires. However, for the slit-less cases, we calculated the propagation between lenses 1 and 2 using SPW (1 step), which is different from this figure.

be represented as the sum of tilted plane waves known as the spectrum of plane waves (SPW). We used the method of the SPW^{16,17} to calculate all free-space propagation, except for the interval from the slit to lens 2 (Fig. 4). This exception is necessary because, with SPW, a large amount of memory is required to calculate large-angle diffraction after the slit. The SPW method also requires a large amount of memory to calculate the propagation between planes separated by a large distance. Hence, we split each interval between the optical elements into several smaller calculation steps. The number of such divisions is indicated in parentheses in Fig. 4.

- We used the far-field approximation to Rayleigh's diffraction formula of the first kind¹⁴ for the propagation from the slit to lens 2 (Fig. 4), which requires a huge amount of memory for large angle diffraction in the case of the SPW method. The operator $\mathcal{P}_{\Delta z}^{\text{FarField}}$ for the far-field approximation is expressed as

$$[\mathcal{P}_{\Delta z}^{\text{FarField}} U(x, y)](x', y') = \frac{\Delta z}{i\lambda r'} \frac{e^{\frac{2\pi i r'}{\lambda}}}{r'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy U(x, y) e^{\frac{2\pi i(x'x + y'y)}{\lambda r}}, \quad (11)$$

where (x, y) and (x', y') are Cartesian coordinate systems on the surfaces before and after the propagation, $U(x, y)$ is the original field before the propagation, Δz is the distance between the two planes along the optical axis, and

$$r = \sqrt{x^2 + y^2 + (\Delta z)^2}, \quad r' = \sqrt{x'^2 + y'^2 + (\Delta z)^2}. \quad (12)$$

- We modeled the grating as a linear phase-modulation pattern. This model can be interpreted as a grating that has an ideally selective diffraction efficiency; in other words, the diffraction efficiency of this grating model is 100% for an observed diffraction order and 0% for the other orders; see Appendix A for details.
- We used Eq. (5) to evaluate R . We evaluated $\Delta\beta$ in Eq. (5) using the width of the calculated monochromatic PSF, which we denote by Δl , through the equation expressed as

$$\Delta\beta = \left(\frac{dl}{d\beta}\right)^{-1} \Delta l. \quad (13)$$

The value of $dl/d\beta$ in Eq. (13) depends on the optical design (see Table 1 and Appendix B). Using the value of $dl/d\beta$ given in Table 1, together with Eqs. (5) and (13), we obtain the equation

$$R = \frac{2,985,356 \mu\text{m}}{\Delta l}. \quad (14)$$

The definition of Δl ¹⁸ used in this paper is the full width at half maximum (FWHM) along the direction of wavelength dispersion; we first integrated the PSFs along the direction perpendicular to the direction of wavelength dispersion and then used them to evaluate the FWHMs.

4 Results and Discussion

4.1 Results

The simulated spectral resolving powers for different wavelengths are shown in Fig. 5; for comparison, the results for slit-less cases are also shown. For cases both with and without a slit, we include the following three types of aberrations (added as wavefront errors at the entrance pupil

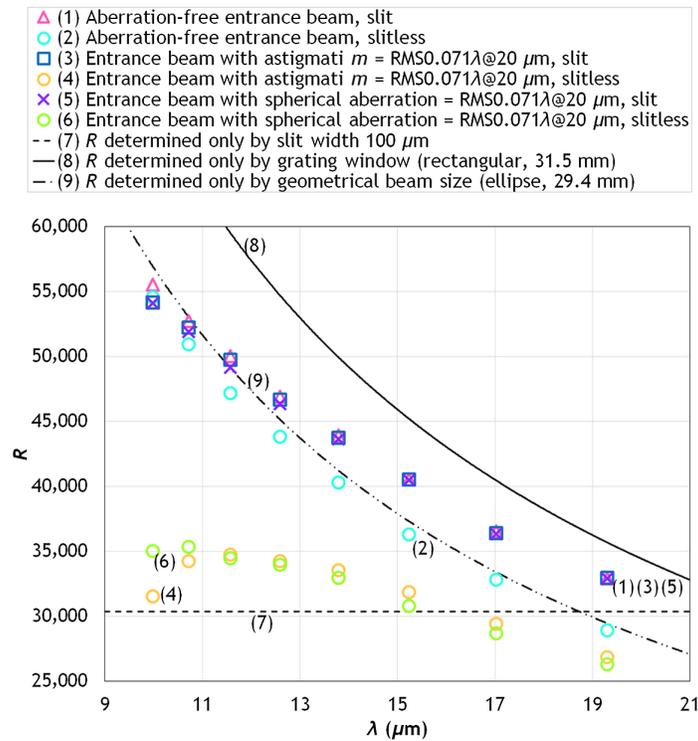


Fig. 5 The simulated spectral resolving power at different wavelengths. The symbols indicate the results of numerical simulations with SPW modeling, and the curves indicate the predictions of the analytical calculations. The magenta triangles (1) show simulated spectral resolving power for an aberration-free case with a $100 \mu\text{m}$ slit. The blue squares (3) and purple crosses (5) represent the results for the cases with a slit, but they include astigmatism and spherical aberration, respectively. In each case, the magnitude of the assumed wavefront aberration is $\text{RMS } 0.071\lambda_{\text{ref}}$ (where $\lambda_{\text{ref}} = 20 \mu\text{m}$). The circles indicate the results of the slit-less cases; the cyan (2), orange (4), and green (6) circles represent the aberration-free results, the results with astigmatism, and the results with spherical aberration, respectively. The assumed aberration is the same as that in the cases with a slit. The three black curves (7–9) are the results of some analytical estimates (see Sec. 4.2 in the main text).

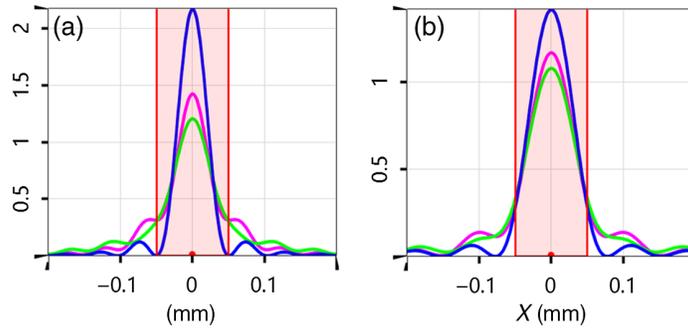


Fig. 6 The one-dimensional PSFs on the slit plane at (a) 9.98108 μm and at (b) 15.23428 μm of the wavelength. The vertical axis is the relative intensity. The blue, magenta, and green curves indicate the cases of aberration-free, astigmatism, and spherical aberration, respectively. The red filled area indicates the width of the slit aperture.

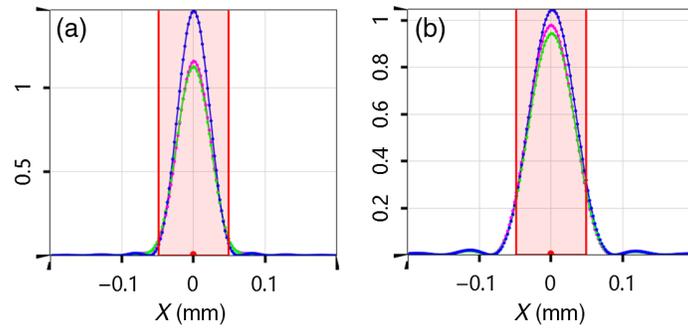


Fig. 7 The one-dimensional PSFs on the final focal plane at (a) 9.98108 μm and at (b) 15.23428 μm of the wavelength. The vertical axis is the relative intensity. The blue, magenta, and green curves indicate the cases of aberration-free, astigmatism, and spherical aberration, respectively. The red filled area indicates the width of the slit aperture.

of the simulation model): aberration-free [(1) and (2) in Fig. 5], astigmatism with root-mean-square (RMS) $0.071\lambda_{\text{ref}}$ [(3) and (4) in Fig. 5], and spherical aberration with RMS $0.071\lambda_{\text{ref}}$ [(5) and (6) in Fig. 5]. The symbols (1), (3), and (5) are for cases with a slit, whereas (2), (4), and (6) are for cases without a slit.

For reference, Fig. 6 shows the one-dimensional PSFs on the slit plane at the blaze wavelength of the diffraction orders $m = 145.95$ (Table 2); in addition, Fig. 7 shows the one-dimensional PSFs on the final focal plane at the same wavelengths as Fig. 6. We obtained these one-dimensional PSFs through integrating the two-dimensional amplitude distribution along the direction perpendicular to the direction of wavelength dispersion.

4.2 Discussion

In Fig. 5, the three curves [(7–9) in Fig. 5] represent some simple analytical estimates of R . In this section, we first explain those three curves and then discuss the results using them. In the discussion below, the distance between lens 2 and the detector is denoted by f ($= 150.929$ mm). We define two types of numerical apertures; one (NA_{GW}) is associated with the size W of the grating window along the dispersion direction, and the other NA is for the geometrical beam size w along the dispersion direction, which is given by

$$\text{NA}_{\text{GW}} = \sin\left(\arctan\left(\frac{W}{2f}\right)\right), \quad \text{NA} = \sin\left(\arctan\left(\frac{w}{2f}\right)\right). \quad (15)$$

The values of NA_{GW} and NA are 0.100550 and 0.0939085, respectively.

The horizontal dashed black line (7) represents the value of R determined only by the slit width, i.e., the value of R for the case in which Δl is the full slit width ($100 \mu\text{m}$). This line corresponds to a specific example of Eq. (9). This line provides a good estimate of R for the sufficiently diffuse sources in all the wavelength ranges discussed in the present paper. (Exceptions are extremely long wavelengths for which we cannot use Kirchhoff's boundary conditions.¹⁴) Next, the black solid curve (8) and the dash-dot-dot curve (9) represent R for cases in which Δl is evaluated from the following equations, respectively and they are expressed as

$$\Delta l = \Upsilon_{\text{rect}}\lambda/(2NA_{\text{GW}}), \quad (16)$$

and

$$\Delta l = \Upsilon_{\text{elli}}\lambda/(2NA), \quad (17)$$

where $\Upsilon_{\text{rect}} = 0.8859$ and $\Upsilon_{\text{elli}} = 1.002$. These lines correspond to specific examples of Eq. (7). Equations (16) and (17) are obtained with the calculation of the Fraunhofer diffraction integral for rectangular and elliptical aperture functions;¹⁴ these expressions are not used in the SPW simulations.

Curve (8) gives the value of R for the case in which the grating window is fully and homogeneously illuminated by a perfect plane wave; thus, this curve can be interpreted as the upper limit on R , which is realized approximately for a case with a sufficiently narrow slit width. In other words, when the collimated beam projection is less than the physical extent of the grating ($w < W$), inserting a narrow slit spreads the beam over the entire grating, reproducing the result where $w = W$ without a slit. Similarly, curve (9) shows the value of R for a case in which the geometrical beam size determines the illuminated region on the grating window. When we compare the symbols (1–6) and the lines (7–9) in Fig. 5, we find the followings:

1. Symbols (2) fall approximately on curve (9).
This means that the numerical and analytical estimates of Δl are approximately consistent with each other.
2. Symbols (1), (3), and (5) trace almost the same positions.
This can be interpreted as meaning that the slit suppresses the effect of the assumed aberrations on R . This is significantly different from the previous estimate of the SPICA SMI-HR spectral resolving power.¹⁵ The previous estimate is based on geometrical optics (Wada et al.¹⁵ estimated the spectral resolving power of SMI-HR assuming a telescope-PSF size diffraction-limited at $20 \mu\text{m}$ and geometrical optics for SMI-HR.) and R takes almost the same values as line (7). Experimental demonstration of the improvement of the spectral resolving power R for the cases with aberrations by the existence of a slit would be valuable.
3. Symbols (1), (3), and (5) lie in the region between curves (8) and (9) for $\lambda \gtrsim 12 \mu\text{m}$.

This is a result of diffraction from the slit spreading the illuminated region of the grating surface. Since the assumed width of the slit is not sufficiently larger than the widths of diffraction-limit PSFs, an electromagnetic wave that has passed through the slit has a different angular spectrum compared to the wave before passing through the slit. Hence, electromagnetic waves that have passed through the slit illuminate the grating surface in different ways from the assumption of curve (8).

Figure 8 shows the results for the simulations of the spectral resolving power for cases with slit widths of 50, 100, 150, 200, 250, 300, 350, and $400 \mu\text{m}$ ($m = 85$, $\lambda = 17.02654 \mu\text{m}$). For each slit width, we examined three types of aberrations (aberration-free; astigmatism with RMS $0.071\lambda_{\text{ref}}$, where $\lambda_{\text{ref}} = 20 \mu\text{m}$; and spherical aberration with RMS $0.071\lambda_{\text{ref}}$). Hereafter, we express the full width of the first-null point along the axis of the short-side direction of the slit ($221.20 \mu\text{m}$) as ε .

Figure 8 shows the following tendencies:

1. For all the aberrations, in the region where the slit width is narrower than ε , R declines approximately linearly as the slit width is increased. This indicates that narrow slits enhance the spectral resolving power even if the window of the immersion echelle grating is as compact as the geometrical beam width.

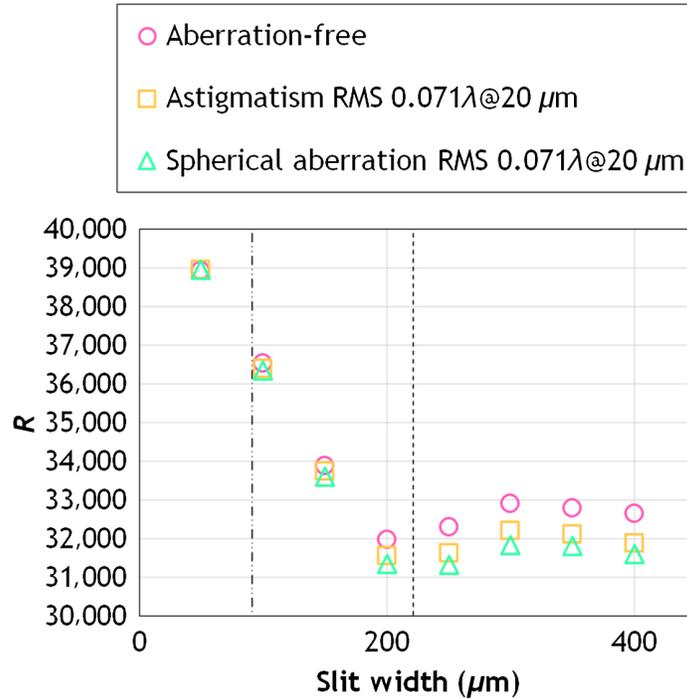


Fig. 8 Spectral resolving power in cases with different slit widths ($m = 85$, $\lambda = 17.02654 \mu\text{m}$). The magenta circles show the simulated spectral resolving power for the aberration-free case. The orange squares and green triangles represent the results for the cases with astigmatism and with spherical aberration, respectively. For each case, the magnitude of assumed wavefront aberration is $\text{RMS } 0.071\lambda_{\text{ref}}$, where $\lambda_{\text{ref}} = 20 \mu\text{m}$. The black dash-dot line represents the FWHM (for the short-side direction of the slit) of the PSF incident on the slit. The first-null point along the axis of the slit (in the aberration-free case) is located at a half width of $110.60 \mu\text{m}$; that is, at a full width of $221.20 \mu\text{m}$. The width is shown as a vertical dashed black line.

- As the slit width increased beyond ϵ , R increases again after first passing through a local minimum. To interpret this phenomenon, it is useful to consider the one-dimensional Fraunhofer diffraction model for the pupil function, which is expressed as

$$\Pi(q) = \begin{cases} 0 & (|q| > \frac{1}{2}) \\ \frac{1}{2} & (|q| = \frac{1}{2}); \\ 1 & (|q| < \frac{1}{2}) \end{cases} \quad (18)$$

that is,

$$\int_{-\infty}^{\infty} dq \Pi(q) e^{-2\pi i \xi q} = \frac{\sin(\pi \xi)}{\pi \xi}, \quad (19)$$

where q is the pupil coordinate normalized by the pupil diameter and ξ is the focal-plane coordinate normalized by $\lambda/(2\text{NA})$ ($= \epsilon/2$). Here, we assume that the slit aperture is $\Pi(\xi/p)$, where p is a parameter that represents the width of the slit (the full-width of the slit normalized by $\epsilon/2$). Then, by assuming that the inverse Fourier transform of $\Pi(\xi/p) \sin(\pi \xi)/(\pi \xi)$ gives the amplitude on the grating surface, amplitude can be expressed as (the finite size of the grating is ignored here)

$$\int_{-\infty}^{\infty} d\xi \Pi\left(\frac{\xi}{p}\right) \frac{\sin(\pi \xi)}{\pi \xi} e^{2\pi i \xi q} = \frac{\text{Si}\left(p\pi\left(\frac{1}{2} + q\right)\right) + \text{Si}\left(p\pi\left(\frac{1}{2} - q\right)\right)}{\pi}, \quad (20)$$

where $\text{Si}(z)$ is the sine integral:

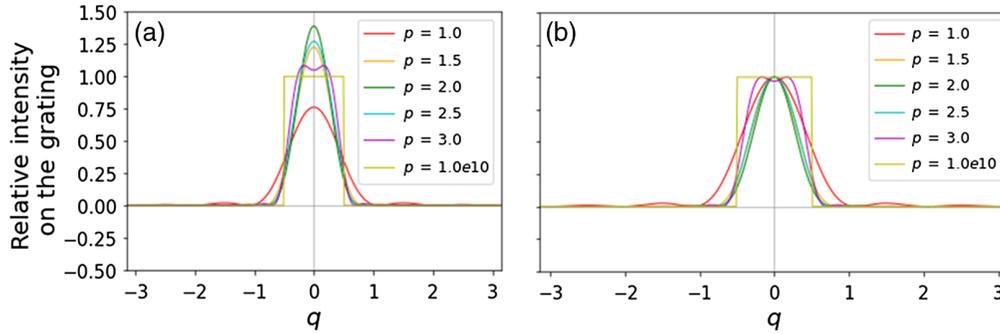


Fig. 9 (a) The functional form of the square of the right-hand side of Eq. (20) for different values of p ; in other words, the intensity profiles on the grating surface for different slit widths (for the one-dimensional model described in the main text). The horizontal axis is the coordinate on the grating surface normalized by the full width of the grating window. The vertical axis is the relative intensity of light incident on the grating. The parameter p is the full-width of the slit normalized by $\epsilon/2$. The curve for the case $p = 1.0 \times 10^{10}$ is shown to represent the slit-less case. (b) The same curves as the left panel that is normalized by their own peak values for clarity of their peak widths.

$$\text{Si}(z) = \int_0^z dt \frac{\sin t}{t}. \quad (21)$$

Figure 9 shows the functional forms of the right-hand side of Eq. (20) for different values of p ; in other words, these are the amplitude profiles on the grating surface for different slit widths; the local minimum in Fig. 8 corresponds approximately to the case $p = 2$. The curve for the case $p = 1.0 \times 10^{10}$ is shown to represent the slit-less case; i.e., this curve shows the amplitude profile of the original pupil aperture. Figure 9 shows that the amplitude of the case $p = 2.0$ is concentrated in a small region on the original pupil aperture, compared with the other cases. As p becomes smaller than 2.0, diffraction from the slit enlarges the effective illuminated regions, and R is thus increased. Conversely, as p becomes greater than 2.0, the amplitude profiles change from bell-shaped curves to top-hat-like shapes. This enlarges the effective illuminated region on the grating until it is as large as the original pupil aperture. Hence, R is increased in this case as well.

3. The difference among the three cases of aberration is not obvious in the region where the slit width is narrower than ϵ , whereas the difference is obvious when the slit is wider than ϵ . This occurs because a narrow slit works as a spatial filter that suppresses the impact of aberrations.
4. Symbols (1), (3), and (5) lie approximately on curve (9) for $\lambda \lesssim 12 \mu\text{m}$.
Since the sizes along the dispersion direction of beams incident on the slit are proportional to the wavelengths for the case without optical aberration, the extent of the slit-diffraction effect at short wavelengths is expected to be smaller than that at longer wavelengths.
5. Symbols (4) and (6) have spectral resolving powers lower than symbols (2).

This can be interpreted as the effect of the assumed aberration. Here, we estimate the effect of the assumed aberration on R as follows. The typical intensity $\langle I \rangle$ of the PSF is defined as

$$\langle I \rangle = \frac{P}{s(\Delta l)^2}, \quad (22)$$

where P is the power at the last focal plane (the two-dimensional integral of the intensity over the whole of the last focal plane) and s is a dimensionless constant, the actual value of which is arbitrary and does not affect the result. Since wavefront aberrations do not change P , Eq. (22) shows that the value of $\langle I \rangle (\Delta l)^2$ is constant, independent of wavefront aberrations. Therefore, by assuming as a working hypothesis that $\langle I \rangle$ is proportional to the Strehl ratio associated with the considered aberration (i.e., an assumption that the Strehl ratio is proportional to the typical PSF intensity $\langle I \rangle$), we obtain

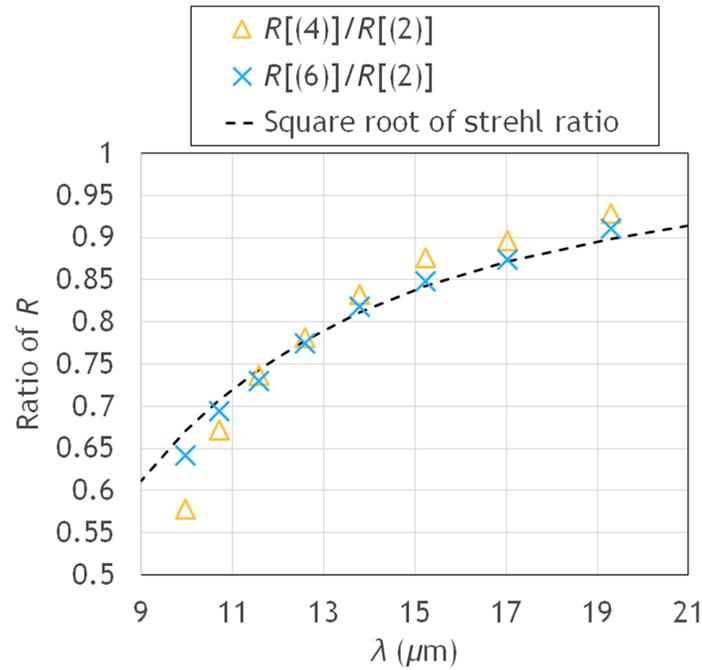


Fig. 10 The impact of aberrations on R in the cases without a slit. The numbers in the legend denote the model identification numbers in Fig. 5. The orange triangles represent the ratio of R for a case with astigmatism [(4) in Fig. 5] to the value of R for a case without aberration [(2) in Fig. 5]. Similarly, the blue crosses represent the ratio of R for a case with spherical aberration [(6) in Fig. 5] to the value of R for a case without aberration [(2) in Fig. 5]. The black dashed curve shows the square root of the Strehl ratio associated with the considered aberrations.

$$\Delta l \propto (\text{Strehl ratio})^{-\frac{1}{2}}. \quad (23)$$

Thus, since R is inversely proportional to Δl [Eqs. (5) and (13)], R (for the slit-less cases) is proportional to the square root of the Strehl ratio.

Figure 10 shows R for cases (4) and (6) in Fig. 5 divided by R for case (2), together with the square root of the Strehl ratio; note that the Strehl ratio for case (2) is approximately unity (diffraction limited). Here, we define the Strehl ratio as

$$(\text{Strehl ratio}) = e^{-(2\pi\Delta q)^2}, \quad (24)$$

where Δq is the RMS wavefront error in waves. Figure 10 shows that the square root of the Strehl ratio is valid as a rough estimate of the impact of aberrations on the value of R in the cases without a slit.

5 Conclusion

The commonly used theory for the spectral resolving power of an immersion-echelle grating does not take into account diffraction from the slit, optical aberration, or two-dimensional aperture shapes and beam profiles. In this study, we performed numerical simulations to take these factors into account to properly investigate the spectral resolving power of a compact, high-spectral-resolving-power MIR spectrometer. This simulation was fully based on wave optics, as computed using Wyrowski VirtualLab Fusion (Second Generation Technology Update [Build 7.3.1.5]).

Our main results are as follows:

1. As diffraction from the slit spreads the illuminated region of the grating surface, when there are no optical aberrations, R is larger for cases with a slit than for cases without a slit.
2. The spectral resolving power R reaches a minimum when the slit width is roughly at the width between the aberration-free-PSF's first null points along the axis of the short-side direction of the slit. As the slit widens more, R increases again slightly thanks to a more uniform illumination of the grating.
3. For cases with a slit (with the slit width assumed to be 94% of the FWHM of the aberration-free PSF at 20 μm), the impact of the assumed aberrations on R is small and can be ignored thanks to the spatial-filtering effect by the slit. This result is completely different from the prediction of geometrical optics, where R depends only on the assumed aberration and is independent of λ . In practical terms, it is important to consider the trade-off relationship between spatial filtering and a throughput since a narrower slit reduces the light throughput especially for aberrated PSFs.
4. The slit-diffraction effect on R is weaker at the short-wavelength region than at the long-wavelength region.
5. For cases without a slit, the impact of aberrations on R can be roughly estimated using the square root of the Strehl ratio associated with the considered aberration.

6 Appendix A: Model of a Reflective Diffraction Grating with Ideal Diffraction Efficiency

The $x - y$ plane on the grating surface is defined as follows:

1. x is the coordinate perpendicular to the grooves on the grating surface and 2. y is the coordinate parallel to the grooves on the grating surface.

For simplicity, assume that the diffraction grating is infinitely long in the y -direction. Also, assume that the diffraction grating modulates the incident plane wave on the grating surface as

$$g(x)\text{rect}(x/D), \quad (25)$$

where

$$g(x + u) = g(x), \quad (26)$$

u and D are the pitch and full width of the grating, and

$$\text{rect}(x) = 1(|x| < 0.5), 0.5(|x| = 0.5), 0(|x| > 0.5). \quad (27)$$

Because $g(x)$ is a periodic function with period u , the Fourier conjugate (This Fourier transform can be physically interpreted as a plane-wave expansion in three-dimensional space.) of $g(x)\text{rect}(x/D)$ is

$$\mathcal{F}[g(x)\text{rect}(x/D)](n\kappa_x) = \left(S(n\kappa_x) \sum_{m=-\infty}^{\infty} \delta(n\kappa_x - m/u) \right) * \frac{\sin(D\pi n\kappa_x)}{\pi n\kappa_x}, \quad (28)$$

where m , n , and $S(n\kappa_x)$ are the diffraction order, index of refraction, and diffraction efficiency, respectively, and where $\kappa_x = k_x/(2\pi)$ and k_x is the wavenumber in the x -direction. Here, we assume an ideal diffraction efficiency; that is

$$S(n\kappa_x) \sum_{m=-\infty}^{\infty} \delta(n\kappa_x - m/u) = \delta(n\kappa_x - m_{\text{design}}/u), \quad (29)$$

where m_{design} means the diffraction order intended to be observed. Equation (28) then becomes

$$\mathcal{F}[g(x)\text{rect}(x/D)] = \frac{\sin(D\pi(n\kappa_x - m_{\text{design}}/u))}{\pi(n\kappa_x - m_{\text{design}}/u)}. \quad (30)$$

The inverse Fourier transform of Eq. (30) is as follows:

$$g(x)\text{rect}(x/D) = e^{2\pi i m_{\text{design}} x/u} \text{rect}(x/D). \quad (31)$$

On the other hand, the grating equation for a reflective grating is given by

$$u(\sin \alpha + \sin \beta) = m_{\text{design}} \lambda / n. \quad (32)$$

Using $n\kappa_x$, Eq. (32) can be rewritten as

$$n\kappa_{x0} + n\kappa_x = m_{\text{design}}/u, \quad (33)$$

where κ_{x0} is $\sin(\alpha)/\lambda$. By substituting Eq. (33) into Eq. (31), we obtain

$$g(x)\text{rect}(x/D) = e^{2\pi i n(\kappa_{x0} x + \kappa_x x)} \text{rect}(x/D). \quad (34)$$

Equation (34) shows that, when we assume an ideal diffraction efficiency [Eq. (29)], we can model a diffraction grating as a “mirror” that reflects plane waves from the direction of κ_{x0} to the direction of κ_x . In the simulations, we define the direction of κ_{x0} as the peak direction of the spectrum of the plane wave incident on the grating.

7 Appendix B: Evaluation of $dl/d\beta$

The quantity $dl/d\beta$ can be evaluated using matrix ray-tracing based on Gaussian optics.¹⁹ We explain the method here. Hereafter, we denote the distance from the center of the grating surface to the grating window by d_1 , the distance from the grating window to the thin lens by d_2 and the distance from the thin lens to the detector by d_3 .

The ray height, h , and the angular direction of the ray, θ , in radians can be written in the vector form, which is expressed as

$$\begin{pmatrix} h \\ \theta \end{pmatrix}. \quad (35)$$

Thus, the vector

$$\begin{pmatrix} 0 \\ d\beta \end{pmatrix} \quad (36)$$

means the chief ray (on the grating surface) with the tiny first-order-direction angle $d\beta$.

The transfer matrix from the grating surface to the detector can then be calculated with the equation that is expressed as

$$\begin{pmatrix} 1 & d_3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{d_3} & 1 \end{pmatrix} \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix} \begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & nd_3 \\ -\frac{1}{d_3} & n\left(1 - \frac{d_2}{d_3}\right) - \frac{d_1}{d_3} \end{pmatrix}. \quad (37)$$

Thus, the chief ray with the tiny direction angle on the grating surface propagates onto the detector, which is given by the equation

$$\begin{pmatrix} 0 & nd_3 \\ -\frac{1}{d_3} & n\left(1 - \frac{d_2}{d_3}\right) - \frac{d_1}{d_3} \end{pmatrix} \begin{pmatrix} 0 \\ d\beta \end{pmatrix} = \begin{pmatrix} nd_3 d\beta \\ \left(n\left(1 - \frac{d_2}{d_3}\right) - \frac{d_1}{d_3}\right) d\beta \end{pmatrix}. \quad (38)$$

Since the height of the ray on the detector can be interpreted as dl , we obtain the expression for $dl/d\beta$, which is given by

$$dl/d\beta = nd_3. \quad (39)$$

Acknowledgments

This research is part of conceptual design activity for the infrared astronomical space mission SPICA, which was a candidate for the ESA Cosmic Vision M5 and a JAXA strategic L-class mission. The SMI consortium is an international team and the authors of this paper include members of this consortium who were in charge of the development of the SPICA Mid-Infrared Instrument; this consortium is led by Japanese universities and ISAS/JAXA.

References

1. S. Notsu et al., “Candidate water vapor lines to locate the H₂O snowline through high-dispersion spectroscopic observations. I. The case of a T Tauri star,” *ApJ* **827**, 113 (2016).
2. S. Notsu et al., “Candidate water vapor lines to locate the H₂O Snowline through high-dispersion spectroscopic observations. II. The case of a Herbig Ae star,” *ApJ* **836**, 118 (2017).
3. N. Fujishiro et al., “Free-form reflective optics for mid-infrared camera and spectrometer on board SPICA,” *Proc. SPIE* **10564**, 105640G (2017).
4. H. Kaneda et al., “SPICA mid-infrared instrument (SMI): conceptual design and feasibility studies,” *Proc. SPIE* **10698**, 106980C (2018).
5. H. P. Stahl, “Survey of cost models for space telescopes,” *Opt. Eng.* **49**, 053005 (2010).
6. G. R. Wiedemann, H. H. Dave, and D. E. Jennings, “Immersion grating and etched gratings for infrared astronomy,” *Proc. SPIE* **1946**, 622–628 (1993).
7. J. P. Marsh, D. J. Mar, and D. T. Jaffe, “Production and evaluation of silicon immersion gratings for infrared astronomy,” *Appl. Opt.* **46**, 3400–3416 (2007).
8. M. Gully-Santiago et al., “Near-infrared metrology of high-performance silicon immersion gratings,” *Proc. SPIE* **8450**, 84502S (2012).
9. Y. Sarugaku et al., “Development of CdZnTe immersion grating for spaceborne application,” *Proc. SPIE* **8442**, 844257 (2012).
10. Y. Ikeda et al., “Machined immersion grating with theoretically predicted diffraction efficiency,” *Appl. Opt.* **54**, 5193 (2015).
11. J. Rayner et al., “iSHELL: a 1-5 micron R = 80,000 immersion grating spectrograph for the NASA infrared telescope facility,” *PASP* **134**, 015002 (2022).
12. Wyrowski Photonics GmbH, “Version History,” <https://www.wyrowski-photonics.com/service/version-history.html> (2021).
13. D. J. Schroeder, *Astronomical Optics*, Academic Press, San Diego (2000).
14. M. Born and E. Wolf, *Principles of Optics*, Cambridge University Press, Cambridge (1999).
15. T. Wada et al., “SPICA mid-infrared instrument (SMI): the latest design and specifications,” *Proc. SPIE* **11443**, 114436G (2020).
16. J. W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, San Francisco (2005).
17. K. Matsushima, H. Schimmel, and F. Wyrowski, “Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves,” *J. Opt. Soc. Am. A* **20**, 1755–1762 (2003).
18. J. G. Robertson, “Quantifying resolving power in astronomical spectra,” *Publ. Astron. Soc. Aust.* **30**, e048 (2013).
19. E. Hecht, *Optics*, Pearson Education, Addison-Wesley (2002).

Biographies of the authors are not available.