# Algebraic ray trace analysis of Spatial Heterodyne Spectrometers

JOHN M. HARLANDER

*Space Systems Research Corporation, 2000 Duke Street, Suite 300, Alexandria, VA, 22314, USA and Department of Physics and Astronomy, 420 Fourth Ave South, St Cloud State University, St Cloud, MN 56301, USA*
*jmharlander@stcloudstate.edu*

**Abstract:** Algebraic ray traces of various configurations of Spatial Heterodyne Spectrometers (SHS) are developed to derive general, approximate, formulas for resolving power, fringe localization plane and admissible off-axis angle for each configuration. Michelson, all-reflective and field widened configurations are considered separately. The derived formulas for each configuration are tested against exact numerical ray traces using optical design software and in general found to be in good agreement.



## 1. Introduction

Spatial Heterodyne Spectroscopy (SHS) is a technique for interference spectroscopy that offers many advantages for high spectral resolution measurements of faint, diffuse sources. When compared to conventional spectrometers SHS instruments can be made more compact, more sensitive, have relaxed alignment and fabrication tolerances and reduced telemetry requirements. As a result, many SHS instruments have been developed or proposed for remote sensing applications from space or ground-based platforms [1-6]. A close relative of SHS is Doppler Asymmetric Spatial Heterodyne (DASH) has been developed to measure atmospheric winds using isolated emission lines [7-9]. Although not considered directly here, DASH instruments can be considered as SHS configurations with an offset aperture so the analysis presented here pertains to DASH as well. SHS is a multiplex technique where spectral information is obtained by a Fourier transform of an interferogram. As a result, multiplex noise must be considered when evaluating the performance of the technique, particularly on dense spectra. The fundamental characteristic of all SHS instruments is an interferometer that produces, a wavenumber-dependent spatially-heterodyned two-beam Fizeau fringe pattern that is recorded by a position-sensitive detector. The system throughput is characteristic of an interferometer as the change in fringe frequency is second order or higher with off axis angle in the interferometer. This results in the well-known throughput advantage of interferometers compared with slit spectrometers. This paper analyzes a variety of SHS configurations from a fundamental ray tracing perspective with the objective of obtaining expressions for the resolving power, fringe localization plane, and maximum allowable off-axis angle within the interferometer for each configuration. We begin with the basic Michelson-based configuration, continue to all-reflective, and finish with an analysis of field widened systems. For each configuration the approximate derived analytic expressions are checked against exact numerical ray tracing of specific examples of that configuration.

## 2. Michelson configuration: amplitude splitting beamsplitter

The basic Spatial Heterodyne Spectrometer consists of a Michelson interferometer with the return mirrors replaced by diffraction gratings and an imaging detector to record the static fringe pattern. Figure 1 shows the configuration with an input aperture, a collimating lens, the interferometer elements (beamsplitter and gratings) and the exit optics required to image the fringe localization plane onto the detector. The input aperture, located at the focal plane of lens $L_1$, determines the range of angles incident on the gratings. The solid angle subtended by the aperture from lens $L_1$ multiplied by the grating area determines the etendue of the system. The tilt angle $\theta_i$ for each grating is chosen so that rays incident on the gratings parallel to the optical axis of a selected wavenumber will retro-reflect. This wavenumber, $\sigma_o$, satisfies the Littrow condition for each grating; $m_{1,2}=2\sigma_o a_{1,2} \sin(\theta_{1,2})$ where $a_i$ is the groove spacing and $m_i$ is the order number of each grating. If the gratings are positioned along the optical axis equidistant from the beamsplitter zero path difference is along this axis. When identical gratings are positioned the same distance from the beamsplitter and used in equal and opposite order ($\theta_2 = -\theta_1$ in figure 1 and the formalism below), a two-sided interferogram is produced with the fringes localized on a plane perpendicular to the optical axis. The resolving power, solid angle field-of-view and passband of the symmetric configuration have been shown to be [10]:

$$R = \frac{\sigma}{\delta\sigma} = 4W\sigma \sin\theta \qquad (1)$$

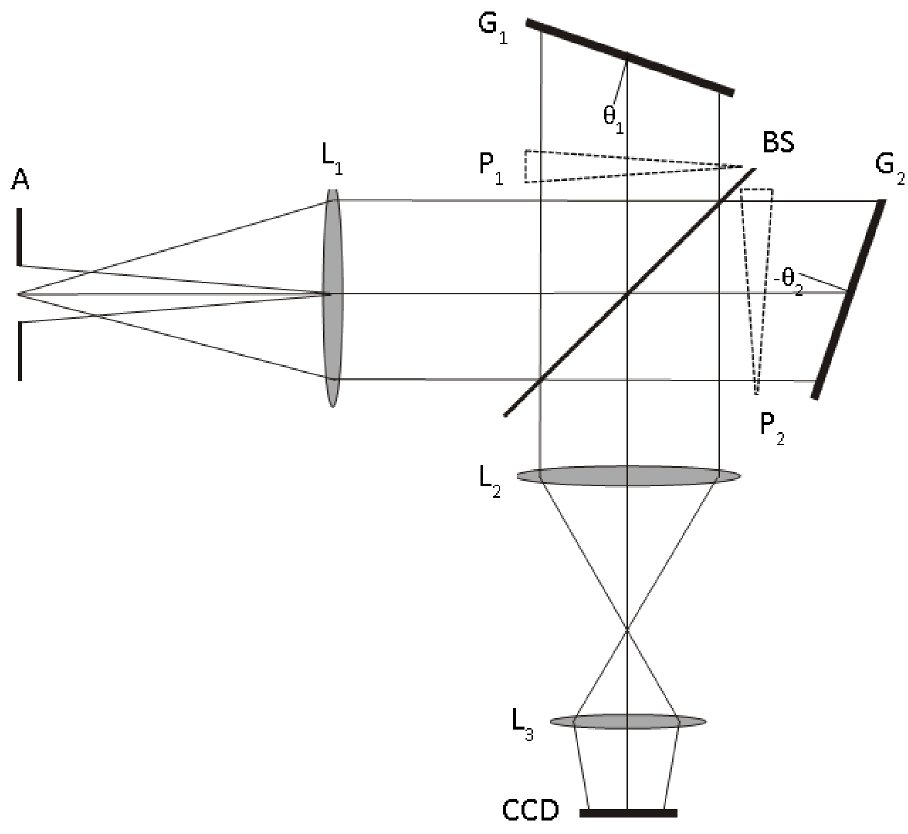


Fig 1. Basic SHS system. Diffraction gratings $G_1$ and $G_2$ terminate the arms of a Michelson interferometer. Aperture A at the focal plane of collimating lens $L_1$ determines the angles incident on the gratings. The fringes are localized on a plane near the gratings that is imaged by lenses $L_2$ and $L_3$ on a position-sensitive detector (CCD). Field widening prisms $P_1$ and $P_2$ (not present in the basic Michelson configuration) are introduced in section 4.

where W is the width of the gratings imaged onto the detector, $\sigma$ is the wavenumber ($\sigma \equiv 1/\lambda$), $\delta\sigma$ is the minimum resolvable wavenumber, and $\theta$ is the Littrow angle of the gratings. The maximum solid angle field of view $\Omega$ at the gratings is

$$\Omega = \frac{2\pi}{R} \quad (2)$$

and

$$\Delta\sigma_{max} = \frac{N\delta\sigma}{2} \quad (3)$$

where $\Delta\sigma_{max}$ is the passband and N is the number of detector pixels sampling the interferogram. The solid angle field of view is characteristic of Fabry-Perot and Michelson FTS interferometers operating at the same resolving power.

### *2.1 Ray tracing the basic configuration*

Figure 2 shows the gratings from Figure 1 unfolded as they appear from the exit of the interferometer along with the coordinates that will be used for algebraic ray tracing. For generality, gratings with two different groove densities (Littrow angles) are illustrated. Point O is the origin of the coordinate system (y is out of the plane of the figure) and the point where gratings $G_1$ and $G_2$ appear to coincide. Grating $G_1$ is inclined at angle $\theta_1$ to the x axis given by the Littrow condition, $\sin\theta_1 = m/(2a_1\sigma_o)$ where m is the order number, $a_1$ is the grating groove spacing and $\sigma_o$ is the Littrow wavenumber. The ray shown in bold is launched from point A at angles $\beta$ (shown in the figure) and $\phi$ (out of the figure plane measured perpendicular to the figure) perpendicular to plane P. This ray strikes grating $G_1$ at point B, diffracts from the grating at angles $\beta'$ and $\phi'$ and exits the interferometer through point C. For illustrative purposes a second ray is shown launched at the same angles from plane P but reflecting off grating $G_2$ (which in general has a different Littrow angle than $G_1$), exiting through point C. At the input to the interferometer these two rays are in a collimated beam (see Figure 1). They are therefore in phase on plane P perpendicular to the rays. The interference at point C may be calculated by determining the optical path (OP) for both rays between plane P and point C as a function of the output position x, y, and z and input angles $\beta$ and $\phi$. These variables are chosen as the independent variables as they are the same for the two interferometer arms. Output angles $\beta'$ and $\phi'$ are considered dependent variables because they are in general different for the two arms. The optical path difference (OPD) and therefore the phase difference at point C can then be calculated by a difference of the optical paths (OPD = $OP_2 - OP_1$). Once the OPD is calculated the interference pattern for monochromatic light is given by:

$$I(x, y, z) = I_0\{1 + \cos[2\pi\sigma OPD(x, y, z)]\} \quad (4)$$

Where $I_o$ is the incident intensity and $\sigma$ is the wavenumber of light which may or may not be equal to the Littrow wavenumber $\sigma_o$.

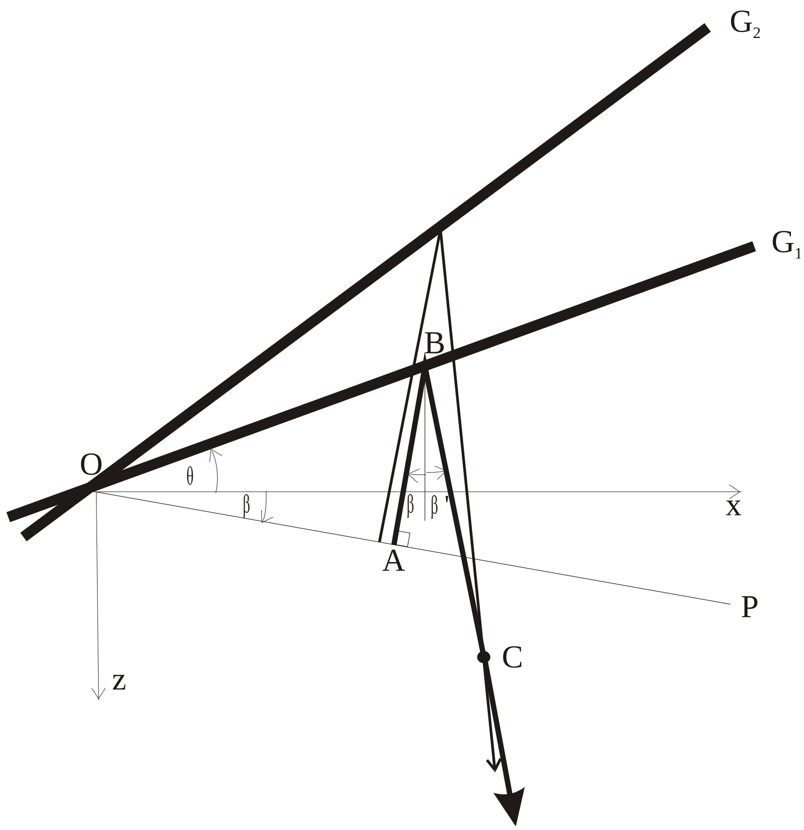


Fig 2. Basic Michelson configuration as viewed from the exit of the interferometer. The gratings $G_1$ and $G_2$ appear superimposed. For clarity, angles $\phi$ measured perpendicular to the x-z plane are not shown.

The optical path for ray ABC can be found by combining the geometric paths AB + BC plus the additional path introduced by the diffraction grating. The grating path is assumed zero at point O where the images of the gratings coincide. At point B the grating optical path is given by:

$$OP_g = -\frac{m(OB)}{a\sigma} = -2(OB)\frac{\sigma_0}{\sigma}\sin\theta = -2x_g\frac{\sigma_0}{\sigma}\tan\theta \tag{5}$$

Where OB is the distance along the grating from O to B in the plane of the figure, m is the order number, $\sigma$ is the wavenumber and $x_g$ is the x coordinate of the ray at point B. The paths BC and AB can be determined from the following equations:

$$BC = \frac{(z + x\tan\theta)\cos\theta}{\cos(\theta+\beta')\cos\phi} \tag{6}$$

$$x_g = x - BC\cos\phi'\sin\beta' \tag{7}$$

$$y_g = y - BC\sin\phi' \tag{8}$$

$$z_g = z - BC\cos\phi'\cos\beta' \tag{9}$$

$$AB = x_g\sin\beta\cos\phi + y_g\sin\phi - z_g\cos\beta\cos\phi \tag{10}$$

Where, x, y, z are the coordinates of the ray at point C and $x_g$, $y_g$, $z_g$ are the coordinates at point B.

The relationship between the primed and unprimed angles is given by the grating equation:

$$2\sigma_0 \sin\theta = \sigma \cos\phi\,(\sin(\theta+\beta) + \sin(\theta-\beta')) \tag{11}$$

$$\phi' = \phi \tag{12}$$

Using equations 11 and 12 to eliminate primed variables and expanding equations 5-10 to second order in β, ϕ, and $\Delta\sigma \equiv \sigma - \sigma_o$, the optical path OP = AB + BC + $OP_g$ is:

$$OP = z + x\beta + y\phi + 2x\frac{\Delta\sigma}{\sigma}\tan\theta - 2\beta\frac{\Delta\sigma}{\sigma}\tan\theta\,(z + x\tan\theta) - 2\left(\frac{\Delta\sigma}{\sigma}\right)^2 (z + x\tan\theta)\tan^2\theta - \beta^2\left(\frac{z}{2} + x\tan\theta\right) - \phi^2\left(\frac{z}{2} + x\tan\theta\right) \tag{13}$$

To determine the interference at point C produced by an interferometer with two gratings with Littrow angles $\theta_1$ and $\theta_2$ at wavenumber $\sigma_0$ the appropriate Littrow angle is substituted into equation 13 and the difference $OPD = OP(\theta_2) - OP(\theta_1)$ is taken:

$$OPD = 2x\frac{\Delta\sigma}{\sigma}(\tan\theta_2 - \tan\theta_1) - 2\beta\frac{\Delta\sigma}{\sigma}\{x[\tan^2\theta_2 - \tan^2\theta_1] + z[\tan\theta_2 - \tan\theta_1]\} - x(\beta^2+\phi^2)(\tan\theta_2 - \tan\theta_1) - 2\left(\frac{\Delta\sigma}{\sigma}\right)^2\{x[\tan^3\theta_2 - \tan^3\theta_1] + z[\tan^2\theta_2 - \tan^2\theta_1]\} \tag{14}$$

The first term in equation 14 is proportional to Δσ and is the term that generates the desired Fizeau fringes. If this were the only term in the OPD the fringe frequency for a given wavenumber ($f_x$ = (OPD)σ/ x) would be linear in Δσ and independent of input angle. The second and third terms in equation 14 shift the fringe frequency with off axis angle as they depend on β or ϕ. The second term is linear in β and can be thought of as a focus error. It can be made zero by setting $z_{Loc} = -x(\tan\theta_2 + \tan\theta_1)$ which identifies the fringe localization plane. This plane is where the fringes have the highest contrast and is the plane that should be imaged onto the detector by the optics following the interferometer, notionally $L_2$ and $L_3$ in Figure 1. Note that for general $\theta_1$ and $\theta_2$ the localization plane not perpendicular to the optical axis of the interferometer ($z_{Loc}$ is a function of x). Substituting the fringe localization plane condition for z into equation 14 gives:

$$OPD = 2x\frac{\Delta\sigma}{\sigma}(\tan\theta_2 - \tan\theta_1) - x(\beta^2+\phi^2)(\tan\theta_2 - \tan\theta_1) - 2x\left(\frac{\Delta\sigma}{\sigma}\right)^2 \tan\theta_2\tan\theta_1\,(\tan\theta_2 - \tan\theta_1) \tag{15}$$

The last term in equation 15 does not depend on input angle and is simply a quadratic shift in fringe frequency with wavenumber. At high spectral resolution over a narrow passband this term is negligible (Δσ/σ is small) and in any case its effect can be determined by wavelength calibration of the instrument. Ignoring this term, the resolving power of the instrument can be calculated by finding the wavenumber difference for which the number of fringes across the entire grating image changes by one for on-axis rays. Assuming a two-sided interferogram with the grating crossing point (point O in figure 2) in the center of the grating image, and setting Δσ=δσ= σ/R as the minimum resolvable wavenumber, the number of fringes between x=0 and x= $x_{max}$ for one-half additional fringe across one-half of the grating aperture is

$$\frac{1}{2} = 2x_{max}\frac{\sigma}{R}(\tan\theta_2 - \tan\theta_1) \quad \text{or} \tag{16}$$

$$R = 4x_{max}\sigma(\tan\theta_2 - \tan\theta_1) \tag{17}$$

The maximum allowable off-axis angles can be determined from the second term in equation 15. Allowing for a fringe shift with off-axis angle of one fringe over the grating image (consistent with a shift of no more than spectral resolution element) gives

$$\frac{1}{2} = \sigma x_{max}(\beta^2+\phi^2)_{max}(\tan\theta_2 - \tan\theta_1) \tag{18}$$

Solving for solid angle and combining with equation 17 gives

$$\Omega \cong \pi(\beta^2+\phi^2)_{max} = \frac{2\pi}{R} \tag{19}$$

Which confirms the result shown in equation 2.

As shown earlier, the localization plane with highest fringe contrast is given by $z_{Loc}$ =-x(tan$\theta_2$ + tan$\theta_1$) which for arbitrary $\theta_1$ and $\theta_2$ is not perpendicular to the optical axis as z is a function of x. The most common implementation of the Michelson configuration is when identical gratings are used in equal and opposite order. In this case $\theta_2=-\theta_1\equiv\theta$ and the localization plane is along the x axis at z =0 (no tilt with respect to the optical axis). The resolving power given in equation 17 then reduces to

$$R = 8x_{max}\sigma \tan\theta = 4W\sigma \sin\theta \tag{20}$$

where W = $2x_{max}/\cos\theta$ is the width of the grating. Using the grating equation at the Littrow angle ($m\lambda/a = 2\sin\theta$) the resolving power is simply equal to the total number of grooves (both gratings) imaged. This confirms the result that the SHS achieves the theoretical resolution of the diffraction gratings while equation 19 indicates a throughput characteristic of an interferometer.

### *2.2 Comparison with exact ray trace*

To verify the above approximate algebraic results for the Michelson configuration, exact numerical ray tracing using ZEMAX was performed for specific instrument parameters. The ZEMAX model assumes a plane wave incident into the interferometer with the two arms modeled as separate configurations. At the output interference of the exiting plane waves determines a fringe pattern on an arbitrary plane. The model is run multiple times varying the input angles of the plane waves to determine how the fringe frequency and phase changes with off-axis angle. The resolving power was determined by finding the wavenumber change ($\Delta\sigma$) that results in adding (or subtracting) ½ fringe at the edge of the aperture (cf. equations 16-19). The localization plane was determined by finding the distance z along the optical axis for which the point at the center of the aperture (x=0) has a path difference of zero for all input angles. Away from this plane the fringe patterns corresponding to different input angles are not in phase (cf. second term in equation 14) and the fringes summed over all angles are therefore out of focus. The instrument modeled for comparison has identical gratings in equal and opposite order corresponding to $\theta_2=-\theta_1$ in the above analytic expressions. The instrument parameters chosen for the model are shown at the top of Table 1. The output parameters of resolving power, fringe localization plane and the maximum angles $\beta_{max}$ and $\phi_{max}$ predicted by the above equations are compared to the values calculated using the ZEMAX model. The localization plane is measured from the origin of the coordinate system shown in Figure 2 and is perpendicular to the z axis. The agreement between the approximate analytic expressions and the exact ray trace indicates that expansions to second order in off-axis angles used in the analytic expressions are sufficient to describe the properties of the instrument.

**Table 1. Analytic vs. ray trace results for the Michelson configuration**

| Parameters | $\sigma_0$ (cm$^{-1}$) | 1/a (l/mm) | $\theta_2=-\theta_1$ (deg) | $x_{max}$ (mm) |
|---|---|---|---|---|
| Input | 18,181.8 | 1000 | 15.96 | 10 |
| Output | R=$\sigma/\delta\sigma$ | $z_{Loc}$ (mm) | $\beta_{max}$ (rad) | $\phi_{max}$ (rad) |
| Analytic | 41,600 | 0 | 0.00693 | 0.00693 |
| ZEMAX | 41,600 | 0 | 0.00693 | 0.00693 |

Figure 3 shows the change in OPD with off-axis angle, expressed as a change in number of fringes at x = $x_{max}$, predicted by the second term in equation 15 and indicated by the ZEMAX model for the instrument parameters in Table 1. The analytic expression prediction is the solid line while the ZEMAX model results are the asterisks. The dotted lines indicate the maximum off-axis angle for which the number of fringes changes by ½ at the edge of the grating (cf. with Table 1).

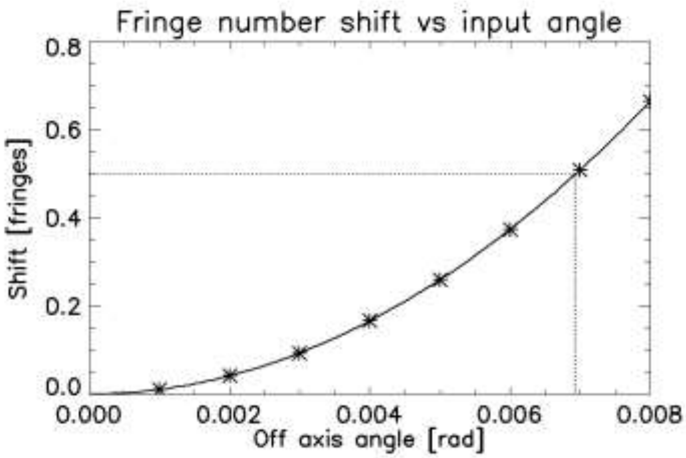


Fig 3. Shift in OPD with off-axis angle at $x=x_{max}$ predicted by equation 15 (solid line) and the ZEMAX model (asterisks). Note that only the dependence on β is shown in the plot as the ϕ dependence is identical. See text for additional details.

## 3. Instruments with grating beamsplitters

For spectral regions where transmitting optics are not available SHS interferometers can be designed in all-reflection configurations using diffraction gratings as beam splitters and combiners. Figure 4 shows three different configurations, each of which have the input and output beams normally incident on the gratings and use equal and opposite diffraction orders to split and recombine the beams. The configurations shown in figure 4A and B both use the same grating for beamsplitting and combining. The input and output beams for these configurations can be separated by using either a split aperture at the focal plane of the collimating lens or roof mirror(s) in the arm(s) to displace the beam into the plane of the page. From a ray tracing perspective configurations 4A and 4B are identical, however, configuration 4B has the advantage of being a common path system where light from both arms reflects from each interferometer element, resulting in greater stability. Configuration 4C employs separate gratings for beamsplitting and combining. It has the advantage that by proper choice of gratings it can produce fringes localized on a plane after the beam combining grating thus eliminating the need for the exit optics to reimage the fringes on the detector. (lenses $L_2$ and $L_3$ shown in Figure 1).

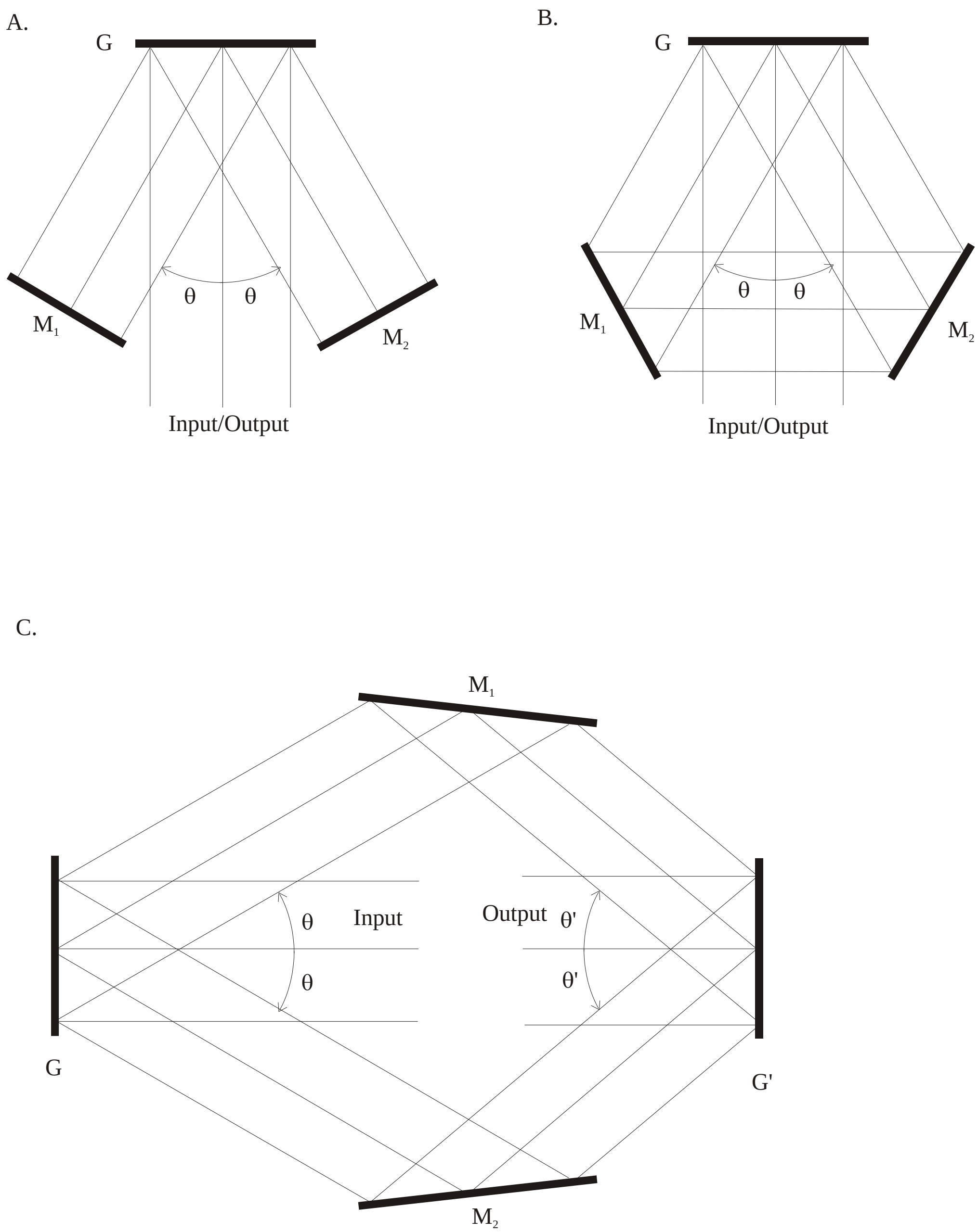


Fig. 4 Three all-reflection SHS configurations.

### *3.1 Analytic ray trace of the all-reflection configurations*

Figure 5 shows the coordinates used to trace a ray through one arm of the all-reflective configurations. The figure is drawn without the plane reflections and with the gratings oriented so that the optical path at the edge of the interferometer (maximum x) has the same sign for both gratings. This is the case for configuration 4A and 4B. For the configuration 4C the path introduced by the gratings have opposite sign which will be treated in the formalism by changing the sign of θ' in the derived equations. By appropriate choice of value and sign for θ and θ' all three of the configurations shown in Figure 4 will be analyzed in terms of independent variables x, y, and z (output coordinates) and β, ϕ (input angles), similar to the approach in section 2 for the Michelson configuration. The out-of-plane angle ϕ, not shown on Figure 5, is measured perpendicular to the x, z plane.

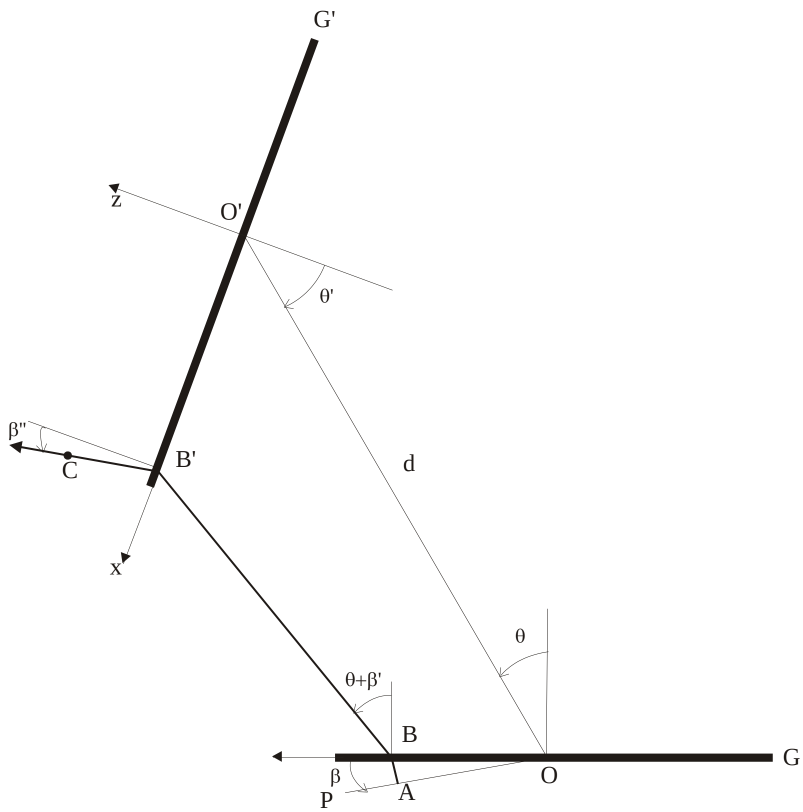


Fig. 5. Unfolded representation of the all-reflection configurations. The variables used in the formalism are indicated on the figure.

The optical path from point A on plane P at the input to point C at the output of the interferometer is given by:

$$OP = AB + OP_G + BB' + OP_{G'} + B'C \quad (21)$$

where $OP_G$ and $OP_{G'}$ are the optical paths introduced by gratings G and G' and the other terms are the line segments indicated on figure 5. Assuming O and O' are the centers of symmetry of the input and output gratings (see figure 4), the optical path at the gratings for each arm is the same for a ray traveling between O and O'. For points away from O and O', the optical paths introduced by the gratings are:

$$OP_G = OB\frac{\sigma_0}{\sigma}\sin\theta \text{ and} \tag{22}$$

$$OP_{G'} = OB'\frac{\sigma_0}{\sigma}\sin\theta' \tag{23}$$

where $\sigma_o$ is the alignment wavenumber (m/a = $\sigma_o$ sin θ for grating G with groove spacing a and m/a' = $\sigma_o$ sin θ' for grating G' with groove spacing a') and OB and OB' are distances measured in the plane of the figure along the gratings. The relationships between β, β', and β'' indicated in the figure are determined by the grating equations:

$$\sigma_0\sin\theta = \sigma\cos\phi\,(-\sin\beta + \sin(\theta+\beta')) \tag{24}$$
$$\sigma_0\sin\theta' = \sigma\cos\phi\,(\sin\beta'' + \sin(\theta'-\beta')) \tag{25}$$

$$\phi'' = \phi' = \phi \tag{26}$$

Following the analysis in section 2, the OPD can be obtained from the difference in optical paths between the rays starting on plane P and traversing the two interferometer arms to point C.

$$OPD = OP(-\theta,-\theta') - OP(\theta,\theta') \tag{27}$$

Using geometry to determine the length of line segments AB, BB' and B'C and expanding the resulting OPD to second order in β, ϕ, and Δσ gives:

$$OPD = 2x\frac{\Delta\sigma}{\sigma}\cos\theta'(\tan\theta'+\tan\theta) - \beta^2 x\frac{\cos\theta'}{\cos^2\theta}(\tan\theta'+\tan\theta) - \phi^2 x\cos\theta'(\tan\theta' + \tan\theta) - 2\frac{\beta}{\cos\theta}\frac{\Delta\sigma}{\sigma}[d\tan\theta + z\cos^2\theta'(\tan\theta'+\tan\theta)] - \left(\frac{\Delta\sigma}{\sigma}\right)^2 x\tan^2\theta\cos\theta'(\tan\theta' + \tan\theta) \tag{28}$$

The term linear in β represents the focus and can be set to zero by choosing

$$z_{Loc} = -\frac{d}{\cos^2\theta'}\frac{1}{(1+\tan\theta'/\tan\theta)} \tag{29}$$

Note that $z_{Loc}$ does not depend on x which means that the localization plane is perpendicular to the optical axis of the instrument. Substituting equation 29 for z in equation 28, ignoring the term quadratic in Δσ/σ and following the analysis in section 2, the resolving power and maximum off axis angles for the general two-grating system are given by:

$$R = 4x_{max}\sigma\cos\theta'(\tan\theta'+\tan\theta) = 2W'\sigma\cos\theta'(\tan\theta'+\tan\theta) \tag{30}$$

Where W' is the width of the beam at the fringe localization plane.

$$\beta_{max}^2 = 2\cos^2\theta/R \quad\text{and}\quad \phi_{max}^2 = 2/R \tag{31}$$

A comparison of equation 19 calculated for the Michelson configuration shows that the maximum ϕ is the unchanged while the maximum β angle is cos θ times smaller for the all-reflection configurations. This decrease is due to the anamorphic angular magnification in the plane of Figure 5 introduced by diffraction grating G used at normal incidence and results in an elliptical rather than circular pattern of angles at the input aperture. It follows that in practice an elliptical aperture input aperture (A in Figure 1) should be used for these configurations.

With the formalism above, the properties of the three configurations shown in Figure 4 can be determined. For configurations 4A and 4B θ’ = θ and the expression for the resolving power reduces to 4Wσ sinθ. Which is the same as for the Michelson configuration, however, here W is the width of fringe image at the fringe localization plane rather than the grating width.

For θ’ = θ equation 29 reduces to $z_{loc} = -d/(2\cos^2\theta)$. The negative sign indicates that the fringes are localized inside the interferometer. Imaging this virtual localization plane requires at least one focusing element between the interferometer and the detector to re-image the fringes onto the detector.

In configuration 4C θ’ and θ have opposite sign leading to a resolving power

$$R = 2W'\sigma\cos\theta'\,(\tan\theta' - \tan\theta) \qquad (32)$$

If | θ’ | > | θ | then tan θ’/tan θ < -1 and it is clear from equation 29 that the fringes are localized at positive values of z. This places the fringe localization plane after the beam combining grating which makes the fringes real and eliminates the need for reimaging the fringes with exit optics.

### *3.2 Verification of the analytic expressions for the all-reflection configurations*

Following the analysis in Section 2, a ZEMAX model for the all-reflection configurations have been constructed and used to verify the formulas derived for them. As indicated earlier configurations 4A and 4B are identical from a ray tracing perspective so they are treated together while configuration 4C, using two different gratings is modeled separately. Table 2 shows in inputs for the configurations and the resolving power, fringe localization plane and maximum allowable off-axis angle for the analytic expressions and ZEMAX model. The fringe localization plane distance $z_{Loc}$ is measured from the plane of the second grating. Negative distance indicates behind the grating (virtual fringes) while positive indicates after the grating (real fringes).

As for the Michelson configuration the comparison between the approximate analytic ray trace analysis and the exact ZEMAX model indicates the validity of the analytic method.

**Table 2. Comparison of the analytic method and ZEMAX ray trace.**

| Parameter | $\sigma_0$ (cm$^{-1}$) | 1/a (l/mm) | 1/a' (l/mm) | $\theta$ (deg) | $\theta'$ (deg) | $x_{max}$ (mm) | d (mm) |
|---|---|---|---|---|---|---|---|
| Conf 4A,4B | 20,000 | 1000 | 1000 | 30 | 30 | 10 | 80 |
| Conf 4C | 20,000 | 500 | 1000 | -14.48 | 30 | 10 | 133.81 |
| Output | | $R=\sigma/\delta\sigma$ | $z_{Loc}$ (mm) | $\beta_{max}$ (rad) | | $\phi_{max}$ (rad) | |
| Conf 4A,4B | Analytic | 80,000 | -53.33 | 0.00433 | | 0.00500 | |
| | ZEMAX | 80,000 | -53.33 | 0.00433 | | 0.00500 | |
| Conf 4C | Analytic | 21,100 | +144.34 | 0.00921 | | 0.00951 | |
| | ZEMAX | 21,100 | +144.34 | 0.00921 | | 0.00951 | |

Figure 6 shows a plot of the change in OPD (fringes) with off-axis angles at $x=x_{max}$. The left plot is for configurations 4A and 4B. The right plot is for configuration 4C. Solid black indicates the analytic OPD shift with $\beta$ while dashed black shows the analytic trend with $\phi$. Results of the ZEMAX model are shown with asterisks indicating the change with $\beta$ and squares the trend with $\phi$. The dotted lines indicate the maximum angles ($\beta_{max}$ and $\phi_{max}$ cf. table 2) for a fringe shift of ½ at the edge of the fringe image ($x=x_{max}$), consistent with a smearing of one spectral resolution element.

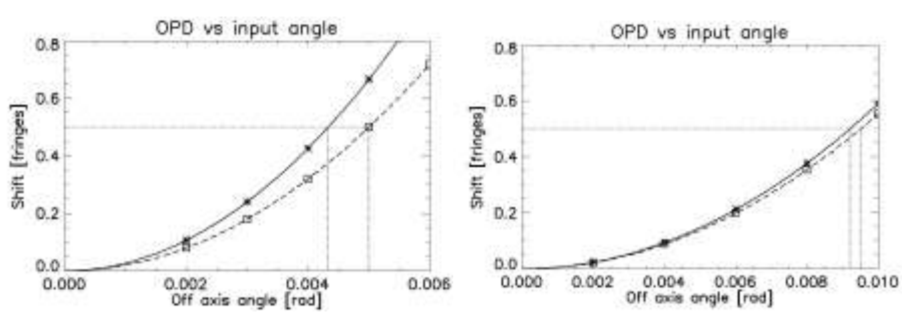


Fig 6. Plot of OPD change with off axis-angles for the all-reflection configurations.

### 4. Field widening with fixed prisms

The field of view limits for the Michelson configuration imposed by equation 19 can be exceeded by inserting fixed field widening prisms in each arm of the interferometer schematically shown as $P_1$ and $P_2$ in figure 1. The prism angle is chosen so that from a geometrical optics point of view the gratings appear coincident, much like a Michelson interferometer at zero path difference. The maximum field of view, limited by prism aberrations, can be much larger than a system without field widening prisms resulting in larger system etendue and instrument sensitivity. In the following sections we describe an analysis of the Michelson configuration with identical gratings used in equal but opposite order ($\theta_2 = -\theta_1$ in section 2) but with the addition of field widening prisms. The expansion of the OPD with off axis angle is more complex in these cases and has been accomplished with the help of the symbolic mathematics software Maple. The detailed design of field widened interferometers is likely best accomplished numerically using optical design software that performs an exact ray trace of each configuration. The goal of the algebraic analysis provided here is to understand the limiting prism aberrations and provide starting points for optimization using a numeric ray trace. As for the sections 2 and 3 above, each section concludes with a comparison of the approximate analytic equations with an exact numerical ray trace.

Section 4.1 considers the case where the prisms are used at the angle of minimum deviation. In this geometry an axial ray enters and leaves the prism at identical angles and the total ray deviation is minimum. At minimum deviation the prism apex angle and the angle of incidence of the optical axis on the prism are coupled so adding the prism introduces only one additional independent variable to the analysis. In section 4.2 we consider the general case where the prism is used at an arbitrary angle of incidence resulting in two additional independent variables, the prism apex angle and the angle of incidence the optical axis makes with the prisms. In both cases the interferometer arms are assumed to be symmetric in the sense that they use the same gratings in equal and opposite order and identically oriented prisms. Section 4.3 considers field widening when the refractive index of the prisms is equal to 2 or greater when by proper choice of prism angles astigmatism can be eliminated.

### *4.1.1 Field widening with prisms at minimum deviation*

Figure 7 shows the unfolded representation of one interferometer arm when field widening prisms are introduced into the arms of the interferometer. The optical axis along which both arms are assumed to have the same path is shown as the ray from point O on an input plane to point $O_6$ on an output plane. Distances $t_0$, $t_1$ and $t_2$ are measured along the axial ray between the points labeled $O_i$ indicated on the figure. The condition of minimum deviation can be enforced by setting $\sin\eta = n\sin(\alpha/2)$ where $\eta$ is the angle of incidence of the axial ray at the first surface of the prism, $\alpha$ is the prism apex angle, and n its index of refraction. At minimum deviation the axial ray also exits the prism at the angle $\eta$. For simplicity, the effect of prism dispersion will be deferred to the end of this section. An off-axis ray enters the interferometer from plane P at point A. The ray enters at angles $\beta$ and $\phi$ with respect to the axis where $\beta$ is measured in the plane of the figure and $\phi$ is the out of plane angle, measured perpendicular to the figure (not shown in the figure). The optical path, OP, between points A and C can be calculated by determining the sum of paths along the ray

$$OP = AB_1 + nB_1B_2 + B_2B_3 - 2\,O_3B_3\frac{\sigma_0}{\sigma}\sin\theta + B_3B_4 + nB_4B_5 + B_5C \tag{33}$$

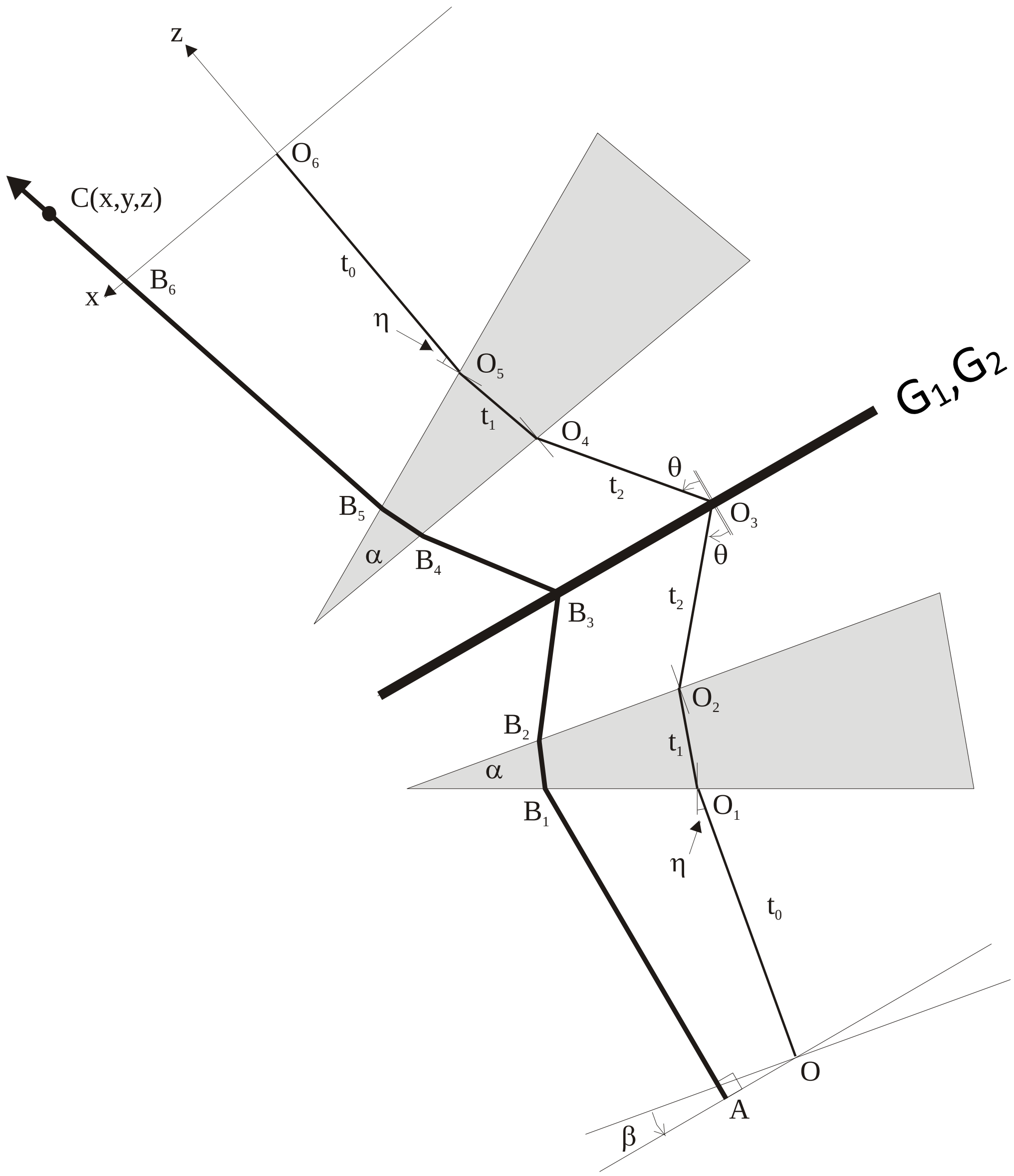


Fig. 7. Unfolded representation of the field widened SHS. Section 4.1 describes the case where the prisms are used at minimum deviation. In this case the axial ray makes an angle of $\eta = \sin^{-1}(n \sin(\alpha/2))$ with respect to the normal at points $O_1$, $O_2$, $O_4$, and $O_5$. Section 4.2 describes the general case where the angles on opposite sides of the prism are not necessarily equal.

where $O_3B_3$ is measured in the plane of the figure and n is the index of refraction of the prisms. The term containing $O_3B_3$ is the path introduced by the diffraction grating. The optical path difference (OPD) between the two arms at point C can determined by taking the difference between the optical paths for the two arms:

$$OPD = OP(x, y, z, \beta, \phi, \sigma, -\eta, -\alpha, -\theta, t_0, t_1, t_2) - OP(x, y, z, \beta, \phi, \sigma, \eta, \alpha, \theta, t_0, t_1, t_2) \quad (34)$$

where the negative signs in the first term reverse the orientation of the prism and grating of the second arm with respect to the first. Using geometry to evaluate the line segments and expanding the OPD to fourth order in β, ϕ and second order in $\Delta\sigma \equiv \sigma - \sigma_o$ yields after much algebra:

$$OPD \approx -4x\frac{\Delta\sigma}{\sigma}\tan\theta + 2x\beta^2\left[\tan\theta - 2\tan\eta\frac{n^2-1}{n^2\cos^2(\alpha/2)}\right] + 2x\phi^2\left[\tan\theta - 2\tan\eta\frac{n^2-1}{n^2}\right] + 4\,\beta\frac{\Delta\sigma}{\sigma}\tan\theta\left[t_0 + t_1\frac{\cos^2\eta}{n\cos^2(\alpha/2)} + t_2 + z\right] + 4x\left(\frac{\Delta\sigma}{\sigma}\right)^2(\tan\theta)^2\left[\tan\theta - 2\tan\eta\frac{n^2-1}{n^2\cos^2(\alpha/2)}\right] + x\alpha(\beta^2+\phi^2)^2\frac{n^2-1}{2n^3} \quad (35)$$

The first term in this equation shows that the linear change in fringe frequency with wavenumber is identical to the non-field widened case described in section 2, resulting in a resolving power that is equal to the configuration without field-widening prisms. The effect of the $\Delta\sigma^2$ term is independent of input angles and can be determined in the wavenumber calibration of the instrument. It will be ignored in the following analysis. The localization plane of the fringes can be determined by choosing the distance z so that the term linear in β is zero. The result from the fourth term in equation 35 is:

$$z_{Loc} = -\left[t_0 + t_1\frac{\cos^2\eta}{n\cos^2(\alpha/2)} + t_2\right] \approx -\left[t_0 + \frac{t_1}{n} + t_2\right] \quad (36)$$

where the approximation is good to first order in prism angles. Equation 36 shows that the fringes are localized on a plane perpendicular to the optical axis at the paraxial image of the point where the gratings cross. This result is similar to the non-field widened case with the addition of the $t_1/n$ term that represents the geometric image of the gratings through a plane glass slab of thickness $t_1$.

The limiting field of view is determined by the second ($\beta^2$), third ($\phi^2$) and last terms $[\beta^2 + \phi^2]^2$ in equation 35. The second and third terms differ by the term containing $\cos^2(\alpha/2)$ which results from prism astigmatism due to the non-symmetric angles of incidence at the prisms. The last term that is quartic in angle is from spherical aberration due to the different thickness of the two prisms along all rays except the optical axis. Ignoring the last term for the moment, by appropriate choice of prism angle α, (or its equivalent η) either, but not both, of the angular terms can be made zero. Setting the $\beta^2$ term equal to zero will result in a field of view that is large in the plane of figure x (then limited by spherical aberration). The field of view in the ϕ direction (perpendicular to the figure) will then be smaller; determined by the resulting nonzero coefficient of the $\phi^2$ term. A field of view that is large in the ϕ direction and small in the β direction can be obtained by choosing the prism angle so the coefficient of the $\phi^2$ term is zero. A compromise between these two extremes is to choose the prism angle so the coefficients have equal magnitude and opposite signs. The limiting angles in the two directions is then the same, while a plot of path difference on the β - ϕ plane has the saddle shape characteristic of astigmatism.

Setting the angular coefficients equal in magnitude and of opposite sign results in the transcendental equation:

$$\tan\theta = \tan\eta \frac{n^2-1}{n^2}\frac{2n^2-\sin^2\eta}{n^2-\sin^2\eta} \tag{37}$$

where the condition for minimum deviation $\sin\eta = n\sin(\alpha/2)$ has been used to eliminate the angle $\alpha$. Once equation 37 is solved for $\eta$, the angle $\alpha$ can then be determined from the condition for minimum deviation.

By substituting equation 37 into equation 35 the field of view at the gratings limited by astigmatism can be determined. Since the OPD trend with angle is quadratic in each direction and of opposite sign a plot of the OPD in the $\beta$ - $\phi$ plane has the saddle shape characteristic of astigmatism. It follows that the OPD change with angle is zero along the diagonals where $\phi = \pm\beta$. The aperture defining the angles into the interferometer (A in figure 1) can then be made square rather than elliptical as in the non-field widened cases. Following the analysis in section 2 it can be shown that for a square field the maximum solid angle for the minimum deviation prism arrangement limited by astigmatism is given by:

$$\beta_{max}^2 = \phi_{max}^2 = \frac{2}{R}\left\{\frac{1+\cos^2(\alpha/2)}{1-\cos^2(\alpha/2)}\right\} \tag{38}$$

Where R is the resolving power.

#### *4.1.2 Effect of prism dispersion*

Because of prism dispersion the resolving power for the field widened system is slightly higher than given above. An analysis of prism dispersion indicates that to first order in index change, $\delta n$, the optical path difference changes by $-4x\alpha(\delta n)$ where $\delta n$ is the change in index of refraction from the Littrow wavenumber. Adding this term to the first term in equation 35 the expression for the resolving power corrected for prism dispersion can be written as:

$$R = R_0 + 8\alpha x_{max}\sigma^2\frac{\delta n}{\delta\sigma} = R_0 - 8\alpha x_{max}\frac{\delta n}{\delta\lambda} \tag{39}$$

where $R_0$ is the resolving power without prism dispersion.

#### *4.1.3 Comparison of analytic expressions and exact ray trace.*

The two rows of table 3 show the input parameters used to compare the analytic and ZEMAX models for the configuration with the prisms at minimum deviation. The prism angle of incidence $\eta$ was calculated using equation 37 after which the prism angle $\alpha$ was calculated using the condition for minimum deviation. The bottom three rows show the output parameters of resolving power, location of the fringe localization plane, and the maximum angles predicted by the theory and the ray trace. Note that the fringe localization plane $z_{Loc}$ is calculated from the same input plane for both cases and that prism dispersion has been included in the calculation of resolving power R.

**Table 3. Comparison of analytic expression with ray trace**

| Parameter | $\sigma_0$ ($cm^{-1}$) | 1/a (l/mm) | $\theta_2$=-$\theta_1$ (deg) | prism | n | $\alpha$ (deg) | $\eta$ (deg) | $x_{max}$ (mm) |
|---|---|---|---|---|---|---|---|---|
| Input | 18,181.8 | 1000 | 15.96 | N-Bk7 | 1.51852 | 18.331 | 13.998 | 10 |
| Output | R=$\sigma/\delta\sigma$ | | $z_{Loc}$ (mm) | | $\beta_{max}$ (rad) | | $\phi_{max}$ (rad) | |
| Analytic | 42940 | | -66.3 | | 0.061 | | -0.061 | |
| ZEMAX | 42980 | | -66.1 | | 0.061 | | -0.062 | |

Figure 8 is a plot of the OPD shift, measured in fringes, vs input angle for the configuration with prisms at minimum deviation. The sold black line is the prediction of the approximate analytic model in the $\beta$ direction while the dashed black line is for the $\phi$ direction when the prism angles are chosen to balance the astigmatism between the two directions as in equation 37. The asterisks and squares are the corresponding values from the ZEMAX ray trace. The dotted lines show the angles at which the shift in fringes is ±0.5 from the ZEMAX model. Note as mentioned earlier, the aperture defining the angles into the interferometer (A in figure 1) can be made square rather than elliptical as in the non-field widened cases. The throughput gain associated with this configuration over a non-field widened interferometer at this resolving power is then (2*0.062)*(2*0.061)/(2$\pi$/42980) ≈ 100 where equation 19 has been used for the throughput without field widening.

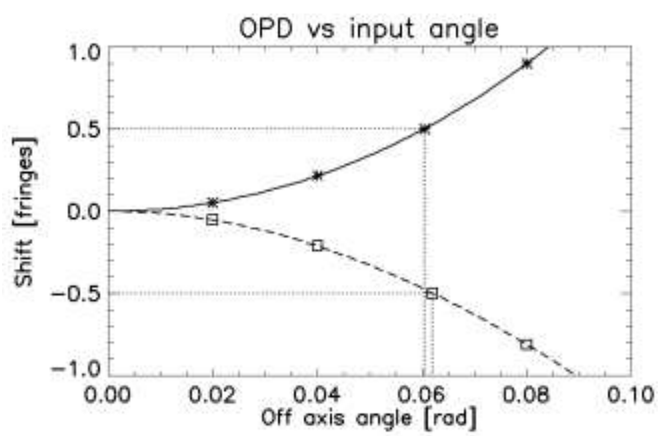


Fig. 8. Plot of the OPD vs input angle for the field widened configuration with the prisms at minimum deviation.

### *4.2 Field widening for arbitrary prism incident angle.*

By allowing an angle of incidence at the prism other than minimum deviation the field of view of the interferometer can be made larger than predicted by equation 38. The analysis follows that given in section 4.1 except the condition $\sin\eta = n \sin(\alpha/2)$ is no longer used with the result that $\eta$ and $\alpha$ are independent variables. If the prism index of refraction is larger than 2, it will be shown that the terms quadratic in off-axis angle can be set to zero (zero astigmatism). This means that the lowest order remaining terms are higher than second order in off-axis angle. To simplify the results but retain the important features the OPD expression (equation 34) will be expanded in variables $\eta$ and $\alpha$ in addition to $\beta$, $\phi$, and $\Delta\sigma$ to fifth order.

### *4.2.1 Term-by-term analysis of OPD*

Expanding equation 34 to fifth order in the variables $\beta$, $\phi$, $\Delta\sigma$, $\eta$, and $\alpha$ and grouping them in terms of $\Delta\sigma$, $\beta$, and $\phi$ results in a series of six terms analogous to those in equation 35. For clarity, each of the six terms will be considered separately. The first term, linear in $\Delta\sigma$ only, determines the resolving power and is

$$\frac{\Delta\sigma}{\sigma} 4x\tan\theta\,\{\,1-\frac{\alpha^2}{2}(n^2-1)+\frac{\eta\alpha}{n}(n^2-1)-\alpha^4\left(\frac{n^4}{8}+\frac{n^2}{12}-\frac{5}{24}\right)+\eta\alpha^3\left(\frac{n^3}{2}-\frac{n}{3}-\frac{5}{6n}\right)+ \eta^2\alpha^2\left(\frac{1-n^4}{n^2}\right)+\eta^3\alpha\left(\frac{5n}{6}-\frac{1}{3n}-\frac{1}{2n^3}\right)\} \quad (40)$$

Although this expression appears complicated, its first term is equal the first term in equation 35, furthermore it can be shown that the if the prism angles are not included in the expansion the exact expression, including all terms is given by

$$\frac{\Delta\sigma}{\sigma} 4x\tan\theta\left[\frac{\cos\eta_1\cos\eta_3}{\cos\eta\cos\eta_2}\right] \quad (41)$$

where $\eta_1$ is the angle of refraction at $O_1$, $\eta_2$ is the angle of incidence at $O_2$, and $\eta_3$ is the angle of refraction at $O_2$ in figure 7. The term in square brackets can be interpreted as the magnification of the aperture in the x dimension introduced by the prism (which is 1 at minimum deviation). The resolving power is changed slightly from earlier versions due to the magnification introduced by the prism resulting in a beam width at the grating that is slightly different than the beam width exiting the interferometer.

The next term of interest in the expansion is the $(\beta\Delta\sigma/\sigma)$ term that serves to localize the fringes (cf. fourth term in equation 35). This term is equal to:

$$\beta\frac{\Delta\sigma}{\sigma}2\tan\theta\,\{z[-2+(n^2-1)(\alpha^2-2\eta\alpha/n)]+t_0[-2+(n^2-1)(\alpha^2-2\eta\alpha/n)]+ \frac{t_1}{n}\left[-2+(n^2-1)\left(\alpha^2-\frac{2\eta\alpha}{n}+2\eta^2/n^2\right)\right]+t_2[[-2+(n^2-1)(\alpha^2-2\eta\alpha/n)]]\} \quad (42)$$

Setting this term to zero and solving for z serves to locate the fringe localization plane. Expanding the result to first order in prism variables gives:

$$z_{Loc} = -\left[t_0+\frac{t_1}{n}+t_2\right] \quad (43)$$

which is the same as the righthand side of equation 36. Since equation 42 is independent of x the fringes are localized on a plane perpendicular to the gratings which is to first order in prism angles equal to the geometrical image of the gratings.

We will now turn our attention to the higher order terms in off axis angle. The terms of immediate interest are the quadratic and quartic terms in β and ϕ (cf. the second, third and last term in equation 35). As the lowest terms in the expansion, they determine to the maximum field of view. For the arbitrary prism angle of incidence these terms are given by:

$$\beta^2 x\{-2\tan\theta + 2\alpha\frac{n^2-1}{n} + 2\alpha\eta\tan\theta\frac{n^2-1}{n} - \alpha^2\tan\theta\,(n^2-1) + \alpha\eta^2(n^2-1)\frac{5n^2+3}{n^3} - 2\alpha^2\eta(n^2-1)\frac{3n^2+1}{n^2} + 2\alpha^3(n^2-1)\frac{3n^2+1}{3n}\} \quad (44)$$

$$\phi^2 x\{-2\tan\theta + 2\alpha\frac{n^2-1}{n} - 2\alpha\eta\tan\theta\frac{n^2-1}{n} + \alpha^2\tan\theta\,(n^2-1) + \alpha\eta^2(n^2-1)\frac{n^2+1}{n^3} - 2\alpha^2\eta\frac{n^2-1}{n^2} + 2\alpha^3\frac{n^2-1}{3n}\} \quad (45)$$

$$[\beta^2+\phi^2]^2 x\alpha\frac{n^2-1}{2n^3} \quad (46)$$

Where equations 44 and 45 are different due to prism astigmatism. Following the discussion in section 4.1 by ignoring the spherical aberration term (equation 46), we set the expressions in curly brackets from equations 44 and 45 equal in magnitude but opposite in sign to obtain a symmetric field of view. The result is, after some algebra:

$$\tan\theta = \alpha\frac{n^2-1}{n}\left[1 + (3n^2+2)\frac{n^2\alpha^2 - 3n\alpha\eta + 3\eta^2}{6n^2}\right] \quad (47)$$

Using equation 47 to eliminate tanθ in either equations 44 or 45 gives the resulting angular quadratic term as:

$$\beta^2 \text{ or } \phi^2 \rightarrow \pm\alpha x\frac{n^2-1}{n^3}[n^2\alpha^2 - n\alpha\eta(n^2+2) + \eta^2(2n^2+1)] \quad (48)$$

Equations 47 and 48 are a general set which allow the calculation of both the prism angle α and angle of incidence at the prism η. By minimizing the magnitude of equation 48 as a function of η, the maximum solid angle is obtained. Equation 47 can then be used to determine the optimum prism angle. Next, we treat two separate cases depending on whether the index of the prism is greater or less than 2.

### *4.2.2 Analysis of prism astigmatism for n < 2.*

In spectral regions where the index of refraction of the prism material is less than 2, the FOV limited by astigmatism is obtained by minimizing equation 48 as a function of η which results in the condition:

$$\eta = \frac{n\alpha}{2}\frac{n^2+2}{2n^2+1} \text{ (n<2) minimum astigmatism} \quad (49)$$

Substituting this into equation 48 gives the maximum off-axis angle

$$\beta_{max}^2 = \phi_{max}^2 = 16\frac{2n^2+1}{R\alpha^2 n^2(4-n^2)} \tag{50}$$

Where R is the resolving power and the first term in equation 47 and equation 17 for $\theta_1 = -\theta_2$ have been used. This expression leads to a gain in solid angle compared to an SHS without field widening (or a single channel Fabry-Perot or Michelson) of:

$$G = 8\frac{2n^2+1}{\alpha^2 n^2(4-n^2)} \tag{51}$$

For n = 1.5 equation 51 gives a value of gain that is ≈ 1.4 times larger than the minimum deviation configuration discussed in section 4.1. Due to the difference term in the denominator, equation 51 predicts that the closer the index of refraction is to 2, the larger the gain compared with a non-field widened system.

Substituting equation 49 into equation 47 gives an expression for the optimum prism angle in terms of the index and grating angle. The equation is:

$$\tan\theta = \alpha\frac{n^2-1}{n}\left[1+\alpha^2(3n^2+2)\frac{7n^4-2n^2+4}{24(2n^2+1)^2}\right] \tag{52}$$

In designing an instrument of resolving power R and grating aperture W, the equation $R = 4W\sigma\sin\theta$ could be used to determine the grating angle θ. Equation 52 could then be used to determine the prism angle α and equation 49 could be used to determine η, the angle of incidence of the axial ray on the prism.

### *4.2.3 Comparison with exact ray trace*

As was done for the previous configurations, the formulas derived in section 4.2.2 were compared with and exact ray trace. The results are shown in Table 4 and figure 9. The same grating aperture (10 mm), groove density (1000 l/mm) and wavenumber (18,181.8 $cm^{-1}$ = 550 nm) as were used in section 4.1.2 were modeled for ease of comparison with the case with the prisms at minimum deviation. Here the focus is on the change in OPD with input angle for the two cases. The prism angle of incidence η was obtained from equation 49 and the prism angle α from equation 52. For this case there is a greater difference between the maximum allowable off-axis angles predicted by the analytic method and the exact ray trace, likely due to the small angle approximations for the prism angles made in the Taylor expansions. Taking the ZEMAX values as limiting the field of view the throughput gain associated with this case relative to a configuration without field widening is (2*0.064)*(2*0.096)/(2π/42270) ≈ 165, which is somewhat larger than with minimum deviation prisms.

**Table 4. Comparison of analytic expression with ray trace for arbitrary prism angle of incidence.**

| Parameter | $\sigma_0$ (cm$^{-1}$) | 1/a (l/mm) | $\theta_2$=-$\theta_1$ (deg) | prism | n | $\alpha$ (deg) | $\eta$ (deg) | $x_{max}$ (mm) |
|---|---|---|---|---|---|---|---|---|
| Input | 18,181.8 | 1000 | 15.96 | N-Bk7 | 1.518 52 | 18.256 | 10.635 | 10 |
| Output | R=$\sigma$/$\delta\sigma$ | | $z_{Loc}$ (mm) | | $\beta_{max}$ (rad) | | $\phi_{max}$ (rad) | |
| Analytic | 42940 | | -66.4 | | 0.074 | | -0.074 | |
| ZEMAX | 42270 | | -66.6 | | 0.064 | | -0.096 | |

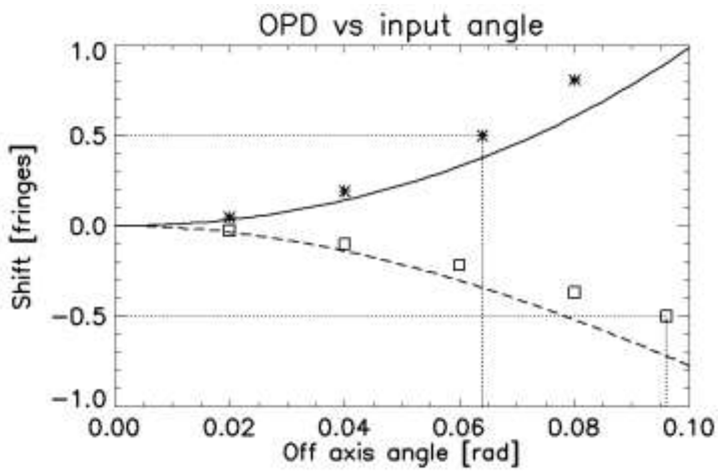


Fig. 9. Plot of OPD vs. input angle for arbitrary prism angle of incidence. Solid black the analytic prediction in the β direction while dashed black is in the ϕ direction. The asterisks and squares are the limits provided by the ZEMAX model.

*4.3.1 Evaluation of prism angles for index of refraction greater than 2.*

In spectral regions where prism materials are available that have an index of refraction of 2 or greater, the factor in brackets in equation 48 can be set to zero by the condition:

$$\eta = \frac{n\alpha}{2}\left[n^2 + 2 \pm \frac{n\sqrt{n^2-4}}{2n^2+1}\right] \qquad (53)$$

This relationship along with equation 47 then provide a set of equations for both α and η for which the quadratic term in off axis angles is zero. The leading non-zero term is then from spherical aberration (equation 46) and leads to a predicted maximum off axis angle given by

$$\beta_{max}^4 = \phi_{max}^4 = \frac{8n^2}{R} \qquad (54)$$

where R is the resolving power.

### *4.3.2 Comparison of analytic method and exact ray trace for high index prisms*

Table 5 and Figure 10 compare the approximate analytic prediction with the exact ZEMAX ray trace for an interferometer with high index ZnS prisms. The combination of equations 53 and 47 were used to determine the prism angles. Figure 8 shows the lines (analytic) and points (ZEMAX) with the same, rather than opposite curvature which indicates that prism astigmatism has been eliminated, however the analytic method predicts a larger field of view. Even so the ZEMAX ray trace indicates a gain in through put of (π*0.134*0.135)/(2π/46750) ≈ 420 over the Michelson configuration without field widening prisms. For this calculation an elliptical limit on the input angles was used since the OPD change with input angle is not an astigmatic saddle but has the same sign in both directions.

**Table 5. Comparison of analytic expression with ray trace for high index prisms.**

| Parameter | $\sigma_0$ ($cm^{-1}$) | 1/a (l/mm) | $\theta_2$=-$\theta_1$ (deg) | prism | n | α (deg) | η (deg) | $x_{max}$ (mm) |
|---|---|---|---|---|---|---|---|---|
| Input | 18,181.8 | 1000 | 15.96 | ZnS | 2.38616 | 8.054 | 3.559 | 10 |
| Output | R=σ/δσ | | $z_{Loc}$ (mm) | | $\beta_{max}$ (rad) | | $\phi_{max}$ (rad) | |
| Analytic | 47790 | | -64.2 | | 0.182 | | 0.182 | |
| ZEMAX | 46570 | | -64.8 | | 0.134 | | 0.135 | |

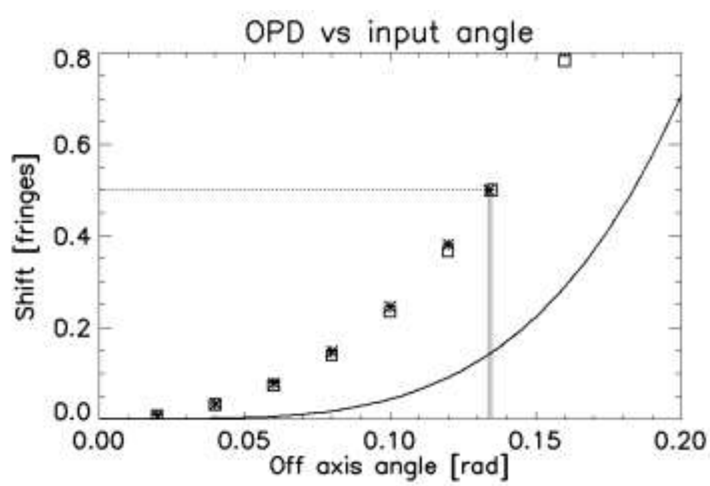


Fig. 10. Plot of analytic prediction (black line) with ZEMAX ray trace (symbols) for prisms with index of refraction greater than 2. Prism astigmatism has been eliminated. Although the analytic method predicts a larger field of view the field of view of the ray trace results in a larger throughput than prisms with index of refraction less than 2.

## 5. Summary

Multiple spatial heterodyne spectrometer configurations have been considered with an emphasis on deriving approximate expressions for instrument parameters, based on the expansion of the OPD for small angles. Formulas were derived for the resolving power, fringe localization plane, and admissible off-axis rays when viewing diffuse sources for each configuration. These were then compared with exact ray traces of example interferometers for each configuration.

For the basic Michelson and all-reflection configurations that do not have field widening prisms discussed in sections 2 and 3, the approximate formulas were found to accurately represent the exact results so that further modeling is likely not necessary when designing these instruments. Because they involve much larger angles, the approximate formulas derived for the configurations with field widening prisms in section 4 were not as accurate when compared to the numeric ray tracing results. The formulas derived in section 4 should therefore be viewed as starting points for further optimization using ray tracing software such as ZEMAX or by numerical calculation of the OPD. Note also that there are often considerations other than maximizing the throughput when designing field widened interferometers. For example, for stability the SHIMMER [2] and MIGHTI [9] interferometers were both monolithic designs using fixed spacers to hold the prisms and gratings to the beam splitter. To simplify the alignment of the monolith, the grating surface was made parallel to the prism surface closest to it which made the spacer between the prism and grating plane parallel rather than wedged resulting in a small decrease in the maximum field of view in favor of easier alignment. Furthermore, the MIGHTI interferometer utilized multiple orders of the gratings to simultaneously observe three separate spectral bands. Due to different indices of refraction for the three different bands, the prism angles could not be optimized simultaneously for all bands so a compromise was struck that slightly reduced the maximum angles at the interferometer.

**Funding.** A portion of this work was supported by the NASA Explorers Program through contract NNG12FA42I.

**Data Availability.** No data were generated or analyzed in the presented research.

**Acknowledgments.** The author acknowledges Fred Roesler for his contributions to the early development of the SHS concept and Christoph Englert for many fruitful conversations. A portion of this work was supported by NASA's Explorers Program.

**Disclosures.** The author declares no conflicts of interest.

## References

1. J. M. Harlander, R.J. Reynolds, and F.L. Roesler, "Spatial heterodyne spectroscopy for the exploration of diffuse interstellar emission lines at far-ultraviolet wavelengths," Astrophys. J., **396**, 730–740, (1992).
2. J.M. Harlander, F.L. Roesler, J.G. Cardon, C.R. Englert, and R.R. Conway, "Shimmer: a spatial heterodyne spectrometer for remote sensing of Earth' middle atmosphere," Appl. Opt., **41**, 1343–1352, (2002).
3. J. B. Corliss, W. M. Harris, E. J. Mierkiewicz, and F. L. Roesler, "Development and field tests of a narrowband all-reflective spatial heterodyne spectrometer," Appl. Opt. **54**, 8835-8843 (2015)
4. Sona Hosseini, "Characterization of cyclical spatial heterodyne spectrometers for astrophysical and planetary studies," Appl. Opt. **58**, 2311-2319 (2019)
5. J.A. Langille, B. Solheim, A. Bourassa, S. Degenstein, S. Brown, and G.G. Shepherd, "Measurement of water vapour using an imaging field-widened spatial heterodyne spectrometer," Appl. Opt., **56**, 4297–4308, (2017).

6. M. Kaufmann, F. Olschewski, K. Mantel, B. Solheim, G.G. Shepherd, M. Deiml, J. Liu, R. Song, Q. Chen, et al., “A highly miniaturized satellite payload based on a spatial heterodyne spectrometer for atmospheric temperature measurements in the mesosphere and lower thermosphere”, Atmospheric Measurement Techniques **11**(7), 3861-3870, (2018)
7. C.R. Englert, D.D. Babcock, and J.M. Harlander, “Doppler asymmetric spatial heterodyne spectroscopy (DASH): concept and experimental demonstration,” Appl. Optics, **46**, 7297–7307, (2007).
8. B. Solheim, S. Brown, C. Sioris & G. Shepherd,“SWIFT-DASH: Spatial Heterodyne Spectroscopy Approach to Stratospheric Wind and Ozone Measurement,” Atmosphere-Ocean, **53:1**, 50-57,(2015), DOI: 10.1080/07055900.2013.855160
9. J.M. Harlander, C.R. Englert, C.M. Brown, K.D. Marr, I.J. Miller, V. Zastera, B.W. Bach, and S.B. Mende,. “Michelson interferometer for global high-resolution thermospheric imaging (MIGHTI): monolithic interferometer design and test”, Space Science Reviews, **212**(1-2), 601–613. (2017). DOI: 10.1007/s11214-017-0374-4
10. J.M. Harlander, “Spatial Heterodyne Spectroscopy: Interferometric performance without scanning”, Ph.D. dissertation (University of Wisconsin-Madison, 1991).