Using birefringent elements and imaging Michelsons for the calibration of high-precision planet-finding spectrographs

J. Schou

Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany
e-mail: schou@mps.mpg.de

March 21, 2022

ABSTRACT

Context. One of the main methods used for finding extrasolar planets is the radial velocity technique, in which the Doppler shift of a star due to an orbiting planet is measured. These measurements are typically performed using cross-dispersed echelle spectrographs. Unfortunately, such spectrographs are large and expensive, and their accurate calibration continues to be challenging.

Methods. A commonly used way to introduce a calibration signal is to insert an iodine cell in the beam. Disadvantages of this include that the lines are narrow, do not cover the entire spectrum, and light is absorbed. Here I show that inserting a birefringent element or an imaging Michelson, combined with Wollaston prisms, eliminates these three shortcomings while maintaining most of the benefits of the iodine approach.

Results. The proposed designs can be made very compact, thereby providing a convenient way of calibrating a spectrograph. Similar to the iodine cell approach, the calibration signal travels with the stellar signal, thereby reducing the sensitivity to spectrograph stability. The imposed signal covers the entire visible range, and any temperature drifts will be consistent and describable by a single number. Based on experience with similar devices that were used in a different configuration by the Helioseismic and Magnetic Imager, it is shown that the calibration device can be made stable at the 0.1 m/s level over a significant wavelength range on short to medium timescales.

Conclusions. While the design is promising, many details still need to be worked out. In particular, a number of laboratory measurements are required in order to finalize a design and estimate actual performance, and it would be desirable to make a proof of concept.

Key words. Instrumentation: spectrographs - Techniques: radial velocities - Stars: oscillations

1. Introduction

The design of spectrographs for highly precise and accurate Doppler measurements for planet finding and asteroseismology is driven by several different scientific requirements, the simultaneous satisfaction of which has resulted in very large and expensive instruments. Driving requirements include the efficient use of the photons collected by the telescope, a desire for extreme stability, and a need to match available detectors. These requirements have led to almost all such spectrographs being some variation of a cross-dispersed echelle design (Eversberg & Vollmann 2015). Indeed, 51 Pegasi, the first planet discovered by the radial velocity method, was originally detected using such spectrographs: ELODIE (Baranne et al. 1996), the Advanced Fiber Optic Echelle (AFOE) (Brown et al. 1994), and the Hamilton Echelle Spectrometer (Vogt 1987).

A persistent problem with such spectrographs is that their inherent stability is not adequate for detecting Earth-sized planets in the habitable zone around Sun-like stars, which typically cause a stellar reflex motion of \(\approx 0.1\) m/s. Fischer et al. (2016) provide a review of the challenges involved in this and discuss various methods currently in use. These methods generally involve introducing a calibration signal into the spectrograph. One approach is to introduce absorption lines, typically using an iodine cell (Brown 1990; Butler et al. 1996). Another approach is to use emission lines, which are typically generated using a thorium-argon (ThAr) lamp (Brown 1990; Lovis & Pepe 2007). Alternatively, a line spectrum may be generated using a laser comb (Phillips et al. 2012; Wilken et al. 2012; Metcalf et al. 2019). Finally, a reference spectrum may be generated using a Fabry-Pérot etalon (Wildi et al. 2010; Reiners et al. 2014; Bauer et al. 2015; Schwab et al. 2015; Jennings et al. 2020). In the case of iodine cells and ThAr lamps, the wavelengths of the spectral lines are inherently stable. Laser combs and Fabry-Pérots are referenced to an external reference.

The calibration signal can either be superimposed onto the stellar light or introduced in a separate beam, effectively creating additional orders on the detector. The former approach has the disadvantage of removing signal photons or adding photons without signal, in both cases degrading the data. The latter approach has the disadvantage that the stellar and calibration signals do not travel together and thus do not have the same sensitivity to instrumental drifts. Moreover, a larger CCD is required than for the former approach due to the doubled orders. Additional problems include that while the wavelengths are very stable, the line strengths are not, leading to problems if the lines are unresolved; that it is difficult to cover the entire wavelength range; and that some of the lines from ThAr lamps, for instance, are much brighter than others, leading to saturation and bleeding into different orders.

Desirable features of a calibration signal thus include not degrading the signal-to-noise ratio (S/N) by adding or subtracting...
photons, covering the entire spectrum, that the stellar light and calibration signal follow each other through the relevant parts of the instrument, and that it is itself inherently stable or can be easily calibrated. None of the designs mentioned above have all of these features.

For solar observations, the requirements are quite different. Here there is a need to image an extended object with good spatial resolution in a variety of variables, including Doppler shift. On the other hand, photons are abundant, and as a consequence, it is possible to use a narrow wavelength range and to accept substantial photon losses. For making Doppler images, this has typically led to the use of so-called “filtergraphs”. In these, images are taken through narrow-band tunable filters, which are realized by various combinations of dielectric filters, birefringent (Lyot) elements, Michelson interferometers, and Fabry-Pérot elements. Examples include the ground-based Global Network for Solar Oscillations Group (GONG; Harvey et al. 1996), the Michelson Doppler Imager (MDI; Scherrer et al. 1995) on board the Solar and Heliospheric Observatory (Domingo et al. 1995), the Helioseismic and Magnetic Imager (HMI; Schou et al. 2012) on board the Solar Dynamics Observatory (Pesnell et al. 2012), and the Polarimetric and Helioseismic Imager (Solanki et al. 2020) on board Solar Orbiter (García-Merariordiga et al. 2021), each of which uses a different combination of filters. For making spatially resolved Doppler images of giant planets, Mach-Zehnder interferometers, which are somewhat similar to the Michelsons mentioned above, have been used (Schmider et al. 2007; Gonçalves et al. 2019).

In Sect. 2 I briefly discuss the operation of birefringent filters and wide-field imaging Michelsons. In Sect. 3 I show how such a device can be incorporated into a spectrograph and how this may be used to make a compact design and provide a good calibration. In Sect. 4 a generalization of the design is explored, and it is shown how this can be used to improve the performance and/or decrease the size of a spectrograph.

Sect. 5 provides a summary. The present paper is only intended to show a concept, thus no detailed layout of a spectrograph is presented. Nonetheless, a number of practical details are discussed in Appendix A.

It should be noted that some of the ideas discussed here have been discussed in connection with the Externally Dispersed Interferometer (EDI; Erskine et al. 2014; Edelstein et al. 2012; Erskine et al. 2011) and the Dispersed Fixed-Delay Interferometer (DFDI; van Eyken et al. 2010), which incorporate a Michelson interferometer. For a detailed theoretical understanding of these, see van Eyken et al. (2010), who also provide an extensive list of references.

2. Basics of imaging filters

In the solar case, imaging filters are used, that is, filters through which an extended object can be imaged in a narrow passband. Many different types of imaging filters have been used, including dielectric interference filters, Fabry-Pérot filters, birefringent filters, and Michelson interferometers. In the following subsections, I briefly discuss the latter two, as they are particularly relevant for the present purpose. For details of these types of filters, their use in HMI, and references to more detailed papers, see Couvidat et al. (2012).

2.1. Birefringent filters

A simple birefringent filter element consists of a piece of birefringent material followed by an exit polarizer. The input light must be polarized, typically either circularly or linearly at 45° to the fast and slow axis of the birefringent element. Various birefringent materials can be used, with calcite being the most common in Lyot filters for solar use, often combined with KDP (KH₂PO₄) or ADP (NH₄H₂PO₄) for temperature compensation. Recently, LiNbO₃ has also been studied by Tomczyk et al. (2016), who also list other candidate materials. Of these, α-BaB₂O₄, TeO₂, TiO₂ and YVO₄ appear to be good candidates, but have not been used in this context. However, before these materials are used, their optical quality and the temperature and wavelength dependence of their birefringence will need to be determined.

Locally (i.e., ignoring secular variations), the transmission as a function of wavelength is given by

\[ T(\lambda) = \frac{1 + B \cos(2\pi(\lambda - \lambda_0)/FSR)}{2}, \]

where \( \lambda \) is the wavelength, B is the contrast, FSR is the so-called free spectral range (determined by the length and birefringence of the element), and \( \lambda_0 \) is the position of a transmission peak. If the entrance or exit polarization is changed by 90° (or between left circular polarization and right circular polarization), \( \lambda_0 \) is changed by FSR/2, that is, the peaks and troughs of Eq. (1) are exchanged. Another factor of 0.5 needs to multiplied on the transmission for the first element in a series in order to take the loss of light in the input polarizer into account. In reality, B<100% and the FSR will change slowly with wavelength. However, as discussed in Secs. 3.1 and 3.2.1, these issues are of little consequence.

An important property of a birefringent element is that all the pieces can be greased (oiled) together, resulting in low reflective losses and a lack of sensitivity to air pressure. One may also choose to temperature-compensate an element by using two different materials (in the case of HMI, calcite+ADP or calcite+KDP depending on the element, as described in Sec. 3.2.2). For a simple element, the central wavelength depends significantly on both the angle to normal and the azimuthal direction relative to the optical axis. For solar imaging applications, the elements are thus typically wide-fielded by using two crossed half-elements with a half-wave plate in between, which results in a lower dependence on the incidence angle and removes the azimuthal dependence.

A Lyot filter is constructed by placing several birefringent elements, generally with a factor of two in the FSR, in series, using the fact that the exit polarizer of one element can act as the entrance polarizer of the next element and tuning each to the design wavelength. When this design is used with ideal polarizers, the transmission is 0.5 at the design wavelength, regardless of the number of elements.

While a Lyot filter is typically optimized to work at a single wavelength, a single element will work over a very large wavelength range, only limited by the transmission of the material used, which in the case of calcite ranges from 4000Å to about 2μm. If all the elements in a Lyot filter are made tunable, resulting in what is known as a universal birefringent filter, the entire transmission range can, in principle, be covered. To accomplish the tuning, a quarter-wave plate and a rotatable half-wave plate are added to each element between one of the polarizers and the birefringent crystal. This allows one to tune \( \lambda_0 \) of the element and obtain a transmission peak at any desired wavelength.

HMI uses a Lyot with five elements (E1 through E5) with nominal FSRRs of 690 mÅ, 1380 mÅ, 2758 mÅ, 5516 mÅ, and 11032 mÅ. Of these, E1 is tunable. The average contrasts are all \( \geq 96\% \).
An alternative type of imaging filters is polarizing Michelsons (Title & Ramsey 1980; Title 2010), such as those used by GONG, MDI, and HMI. In these, the input light is divided into two polarizations in a polarizing beamsplitter, reflected in the two arms and recombined in the polarizing beamsplitter. With the appropriate addition of various waveplates, the light exits the fourth face of the beamsplitter.

While very different in design, the properties are quite similar to those of the birefringent filters described in the previous section. In particular, the transmission profile is also given by Eq. (1), they can be tuned by rotating a half-wave plate and be temperature-compensated. Main differences include that the light is reflected by 90°, a possibility of an improved angular dependence of the central wavelength (fourth order instead of second order), that they can often be made smaller, and that they do not depend on materials that are difficult to obtain. Downsides include that at least for the design used for MDI and HMI, one of the legs is air- or vacuum-spaced, meaning that they are highly sensitive to air pressure (if not operated in vacuum) and temperature. Temperature compensation using solid legs with two different types of glass, for example, has not been used and would require careful attention to mechanical stresses.

HMI contains two Michelsons, the so-called Wide Band (WB) Michelson with an FSR of 344 mÅ, and the Narrow Band (NB) Michelson with an FSR of 172 mÅ. Both of these are tunable. As for the Lyot element, this is done using a rotating half-wave plate. In a stellar context, imaging and temperature-compensated, but nonpolyrizing, Michelsons have also been investigated in the context of the EDI and DFDI designs (Mahadevan et al. 2008; Wan & Ge 2010; Wang et al. 2012).

3. One possible simple design

The idea of an EDI is basically to insert a classical Michelson into the collimated beam of a standard spectrograph, in order to increase the effective resolution of the spectrograph. A downside of such an arrangement is that half of the light is lost.

Clearly, it would desirable to make use of all of the light. The solution proposed here is to split the input light into two complementary polarization states and to replace the classical Michelson with one of the imaging filters described in the previous section. A sketch of the proposed design is shown in Fig. 1. There are multiple options for how to incorporate the device into a spectrograph, one of which is shown in Fig. 2. In this setup, the device is placed in a short section of collimated light immediately after the fiber. After this, the light can either be focused back to a point or continue as it would have without the device. The latter option is shown here. Given that they are likely to be the preferred option and given that they make for simpler drawings, I have chosen to use a birefringent element for the illustrations rather than a Michelson interferometer.

The splitting of the light into horizontal and vertical polarization can be achieved using a so-called Wollaston prism. Such a prism consists of two oppositely oriented prisms of a birefringent material with crossed optical axes cemented together. When light is passed through such a prism, the horizontally polarized part of the light will be refracted by the same angle and the vertically polarized part by the same angle. By choosing the angle of the prisms appropriately, the divergence angle can be set to the desired value.

As discussed in the previous section, the transmission of the two beams at a fixed output polarization will be oppositely modulated. One beam will be modulated by the transmission in Eq. (1) with one value of \( \lambda_0 \) and the other beam with FSR/2 added to \( \lambda_0 \) (i.e., with peaks and troughs exchanged).

Clearly, using a polarizer on the output will result in the loss of half the light. To overcome this problem, the output polarizer can also be replaced by a Wollaston prism. Again, this will cause each of the beams inside the element to be split into two beams with opposite polarization and opposite modulation. In total, there will be four output beams, two with the one modulation pattern and two with the opposite pattern, as illustrated in Fig. 1. Examples of the resulting spectra are shown in Fig. 3. Basically, each of these spectra is similar to a spectrum taken through an iodine cell, with the irregular iodine spectrum replaced by a regular sinusoidal modulation. The spectra are even more similar to those obtained by inserting a Fabry–Pérot into the beam (Wildi et al. 2010; Bauer et al. 2015). However, unlike in those approaches, no light is lost. The sum of the four spectra contains all the input light.

In order for the four beams to remain separate, the deflection angle of the second Wollaston prism (\( \alpha_2 \)) must be chosen to be different from that of the first prism (\( \alpha_1 \)), most simply by setting \( \alpha_2 \approx 2 \alpha_1 \) or \( \alpha_2 \approx \alpha_1/2 \). The smaller of the two angles must be chosen such that the spectra are spaced by at least their widths, that is, by at least a fiber width. Since the distance from the fiber exit to the device would be very roughly similar to the length of the device, it follows that the offset of the beams at the exit will very roughly be one fiber diameter. Regarding the choice of \( \alpha_2 \approx 2 \alpha_1 \) versus \( \alpha_2 \approx \alpha_1/2 \), this will depend on a more detailed analysis. The latter has the advantage of minimizing the angles in the device, thereby minimizing any angular effects and thermal gradients. The latter has the advantage of placing the complementary spectra (A,C and B,D) next to each other.

One way to look at the setup in Fig. 2 is to note that if one were to look into the device, one would see four images of the fiber, placed above each other, with each image representing one of the output beams. This is similar to the well-known double images seen through calcite crystals.

This design has several advantages over a standard cross-dispersed echelle design, some of which have been pointed out in connection with the EDI and DFDI designs: The modulation provides a convenient and consistent calibration signal across the spectrum; the modulation signal can be made extremely stable and is easy to calibrate; and it increases the efficiency of the spectrograph in the sense that a lower noise level on the Doppler shift can be obtained with a spectrograph of the same resolution or alternatively that the same noise performance can be achieved with a lower spectrograph resolution. Below I address these three items and discuss some advantages and disadvantages.

3.1. Noise performance

With a desired Doppler velocity noise level much below 1 m/s, photon noise will dominate readout noise. At higher noise levels (lower signal), the readout noise will become significant, with the exact numbers depending on observing strategy, wavelength coverage, and so on. For the present discussion, I assume that photon noise is dominant, but given a more detailed design, the accuracy of this should be revisited. With this assumption, it is straightforward to estimate the noise performance, given a high-resolution spectrum and a model of the instrument. To do so, a model of the spectra is made (see Fig. 3), including the finite spectral resolution. This model is linearized with respect to the model parameters (in particular, the Doppler velocity), the noise on the incoming spectrum is modeled as photon noise, and
Fig. 1. Sketch of the simple design. Unpolarized (green) light enters on the left. This is split into vertically (red) and horizontal (blue) polarized light by the first Wollaston prism (W1). Inside the birefringent element, the polarization of the light depends on wavelength and position. After the birefringent element, the light is split by the second Wollaston prism (W2), with twice the offset angle of the first. The transmissions of the four beams as a function of wavelength are shown on the right. The top (A) and bottom (D) beams are identically modulated. Similarly, the middle two beams (B and C) are identically modulated. For illustrative purposes, the relative distances are not to scale, the angles are vastly exaggerated and are not two-to-one, and only the center ray of each collimated beam is shown. In reality, the collimated beam would nearly fill the element, the spacing of the rays would be on the order of the fiber diameter, and \( \alpha_2 \approx 2\alpha_1 \) or \( \alpha_2 \approx \alpha_1 / 2 \). For clarity, the Wollaston prisms are spaced away from the element. In reality, they would likely be in contact. Also, they are shown with zero thickness.

Fig. 2. One possible way to insert the device into a spectrograph. Here it is inserted in a short section of collimated beam close to the fiber. Alternatively, the light can be refocused after the device.

The noise on the parameters is determined from a least-squares fit using standard statistical methods. For the present paper, a high-resolution Fourier transform spectrometer (FTS) spectrum of the Sun observed by Reiners et al. (2016), with a resolution of \( \approx 10^6 \), is used. The wavelength range 6000 Å to 6250 Å is used here for illustrative purposes, as it is mostly free of terrestrial blends and contains the spectral line used by HMI. In all cases, the performance is shown relative to an ideal spectrogram (infinite resolution) receiving the same number of photons per unit wavelength at the focal plane.

Figure 4 shows the performance as a function of FSR, calculated using a least-squares fit, as described above. The performance of this finite-resolution spectrograph is, of course, worse than that of an ideal infinite resolution spectrograph. At high resolution, such as shown here, the performance with an added Lyot element is only very slightly better (dip around 0.2Å) than the performance with an unmodified spectrograph.

As all the light makes it to one of the four spectra regardless of the contrast, the original spectrum can be recovered by summing them. As a result, the noise performance is never worse than an unmodified spectrograph, independent of the contrast.

While the addition of the element only improves the performance slightly at a typical resolution of \( R=100\,000 \), the improvement is substantial at lower resolutions, as shown in Fig. 5. At a resolution of \( 10^4 \), which is achievable with a simple spectrograph, the improvement is about a factor of 2.5. Indeed, this is one of the main advantages of the EDI.

3.2. Use as a calibration signal

The main use of the modulation is as a calibration signal. Advantages here include that the entire spectrum is modulated in a uniform way, that no light is lost (making use of all photons), that no light is added (not adding photons and thus photon noise), and
An advantage of filter elements, such as those discussed here, is that all wavelengths follow essentially the same path through the elements and that the absolute position of the modulation signal is thus described by a single number. To see why this is the case, one needs to consider how much each of the two beam paths inside the element change with wavelength, what the temperature gradient is across this distance, and how the fringe positions change with temperature.

If we consider the setup in Fig. 2, the distance between the fiber and the element will very roughly be the same as the length of the element. As we need to offset the two images of the fiber by about one fiber width, it thus follows that the two beams will be spaced by roughly a fiber width at the output of the device, or something like 100μm. If the Wollaston prisms are made of quartz, the offset will vary by $\approx \pm 2\%$ between 4000 Å and 8000 Å. In other words, the different wavelengths in each beam will travel within $\approx \pm 2\mu m$ of each other or less than a part in a few thousand of the width of the collimated beam (on the order of 10 mm). With a difference of even a fraction of a degree across the device, the difference will thus be well under 0.1mK, which, as discussed in Sec. 3.2.2, is adequate. Temperature gradients along the rays do not enter the computation. Similarly, the difference between the two beam paths is not important, as they can be calibrated separately and averaged.

The other question is if the positions of the calibration fringes can be linearized over the relevant temperature range (some mK). To estimate this, one may use the results in Fig. 20 of Couvidat et al. (2012). At a distance of $\Delta T$, $(dT/dT)/\Delta T \approx 1.4 \times 10^{-5} K^{-2} \Delta T$. In other words, the quadratic term is tiny compared to the acceptable linear drifts, as discussed in Sec. 3.2.2.

Given these estimates, it should be a very good approximation to assume that the drift can be written as a function of wavelength times a single function of temperature. This is unlike a regular spectrograph, where thermal expansion effects can vary across the spectrum, and it is unlike ThAr lamps, where the relative intensities depend on more than one factor. The birefringent elements have the further advantage that all the modulation occurs while the light travels through a solid material, in other words, they are insensitive to air pressure, again unlike an in-air spectrograph or an air-spaced Michelson.

Two important properties to consider are the effects of the photon noise on the calibration and the effect of the stability. These are discussed below.

### 3.2.1. Photon noise

The photon noise on the calibration is illustrated in Fig. 4. The calibration noise is substantially less than that on the Doppler signal for a wide range of FSRs, including those of the HMI Michelsons and the narrowest Lyot elements. This is as expected, as the Doppler signal is only sensitive to the photons in the lines, while the calibration signal uses all wavelengths.

Unlike for the stellar signal, the noise in the calibration signal will decrease roughly in inverse proportion to a decreased contrast. However, for HMI, the contrasts are in the range of 96% to 100%, which combined with the large margin means that this should not represent a problem. One could potentially use this large margin to simplify the design by replacing the Wollaston prisms by partial polarizers, which would result in a single spectrum instead of four. With weak polarizers, most of the light will be transmitted and the calibration signal will be weak, but strong enough for the calibration. This is more or less equivalent to in-
serting a weak Fabry-Pérot and using the fringes for calibration, which may be physically simpler.

A way to overcome flatfield errors in the calibration signal is to add a rotating half-wave plate between one of the Wollaston prisms and the birefringent element. If this is rotated by 45°, the spectra are effectively exchanged, meaning that the flat-field difference between the spectral pairs can be canceled, at least to lowest order. While this might at first appear to introduce an unacceptable risk of additional noise due to the angular repeatability of the rotation, this is actually a small effect, as encoders can be made extremely accurate. For HMI, the rms of the repeatability of the angular position of the rotating waveplates is roughly $10^{-6}$ of a revolution. For the E1 element, this corresponds to an apparent Doppler shift of $c \times 4 \times 10^{-6} \times 0.69 \, \text{Å}/6173 \, \text{Å}$, or about 0.13 m/s, where $c$ is the speed of light and the factor 4 comes from the fact that the polarization repeats four times when a half-wave plate is rotated once. The HMI encoders were not designed for extreme accuracy, and a much better performance is available. However, it is not clear that this added complexity is necessary, and it should be pointed out that light losses and reflections from the additional waveplates need to be considered.

### 3.2.2. Temperature stability

As mentioned earlier, the temperature dependence of Lyot elements can be dramatically reduced by combining different materials. At 6173 Å, calcite has a temperature dependence of $-17$ (m/s)/mK, KDP $-80$ (m/s)/mK, and ADP $-350$ (m/s)/mK. Thus the temperature dependence can be canceled by combining a piece of calcite with a thinner piece of KDP or ADP rotated by 90°. Generally, ADP is preferable given the larger temperature dependence, which means that a thinner element can be used.

The design specification for the HMI Lyot elements was $10$ mÅ/K (about 0.5 (m/s)/mK). As the temperature and wavelength dependence of the birefringence of ADP and KDP were not known accurately enough, it was decided to make test elements and measure their temperature dependence. As suspected, this resulted in thickness ratios slightly different from those determined from published values.

The actual temperature stability was measured both standalone and inside the instrument, giving somewhat inconsistent numbers (but still within the specification). This is likely due to differences in the temperature zeropoints of the measurements. Indeed, the slope appears to go to zero about 1-2 degrees above the nominal operating temperature in Fig. 20 of Couvidat et al. (2012). It may be noted that as the linear term per construction dominates, especially since the measurements are made over a range of multiple K versus the mK range relevant for the actual operation. When the conservative 10 mÅ/K specification and an oven stability of around 1 mK (similar to what is seen with the HMI oven) is used, one obtains a 0.5 m/s expected short-term stability. Regarding this number, it is important to note that it is possible to control the temperature to better than 1 mK, especially as it is a small item, and that any temperature changes will be very slow.

It is important to keep in mind that the compensation is only for bulk temperature changes (or linear temperature gradients by splitting the elements), not differential ones. It is thus essential that temperature gradients across the elements are minimized. For HMI, this was done by surrounding the element with a conductive tube and limiting the view of other components with different temperatures.

![Fig. 6. Wavelength dependence of the temperature dependence. The solid black line shows the temperature dependence of calcite. Dashed and dash-dotted black lines show ADP and KDP, scaled to having the same value as calcite at 6173 Å. The top red lines show the results for calcite+ADP and calcite+KDP, optimized for 6173 Å. The dotted lines show the HMI specification. The calcite numbers are from Eq. (10) of Title et al. (1976). The ADP and KDP numbers are derived from the 201 K and 300 K numbers in Table 1 of Ghosh & Bhar (1982) and temperature expansion coefficients of $4.2 \times 10^{-6}$ and $44 \times 10^{-6}$ for ADP and KDP, respectively (Couvidat et al. 2012).](image-url)

Given that the temperature drift of ADP is much larger than that of calcite, it may be shown that in order to compensate to $\epsilon$ times the calcite drift, the thickness ratio has to be accurate to $\approx \epsilon$. In other words, the ADP thickness does not have to be extremely accurate.

### 3.2.3. Wavelength dependence of the sensitivity

The tests described above were performed at 6173 Å, which leaves the question of how well the temperature compensation works at other wavelengths. As already mentioned, the temperature and wavelength dependence of the birefringence of ADP and KDP are poorly known. One set of numbers were published by Ghosh & Bhar (1982); the results obtained using these numbers are shown in Fig. 6. This shows a substantial reduction of the temperature dependence at all wavelengths. However, it is far from perfect. By using a combination of all three materials it is possible to obtain an almost perfect cancelation, but unfortunately this also leads to much longer elements. However, given the uncertainties in the published values, including those in Ghosh & Bhar (1982), in particular, over which wavelength range they are reliable, more accurate measurements of the materials are needed before a final design is made.

Depending on the results of the more detailed measurements of the various candidate materials and the desired wavelength range and stability, the required temperature stability can be determined. If this stability is not achievable, it may be necessary to calibrate more frequently or even continuously.

### 3.2.4. Long-term stability

To illustrate the long-term stability of such elements, Fig. 7 shows the drift of the three tunable elements in HMI. The drifts are given by the peak wavelength $\lambda_0$ in Eq. (1) and converted to the equivalent velocity. The physical origins of the drifts are un-
known, but might be due to effects such as radiation damage to the glass in the Michelsons or settling of the glue used to assemble them. For the Lyot element, in particular, the apparent drift is likely spurious and due to the drift in other preceding filter elements, such as the front window or the dielectric blocking filter before the Lyot. The annual variations and the jump at the power cycling and end of each observing night. In this way, the light loss can be doubled in one direction, which will have massive impacts on the downstream spectrograph.

In summary, it is not clear that a practical arrangement can be found to reduce the number of spectra to two. However, given the significant potential benefits, this possibility should clearly be investigated in more detail.

Imaging Michelsons can in principle be constructed in large sizes, but they are tricky to manufacture, require a different optical layout due to the bending of the light, and most likely will have to be kept in vacuum to avoid pressure effects. Finally, new calibration algorithms will have to be developed. Experiences from using Fabry-Pérot systems may be applicable for this; see, for example, Bauer et al. (2015).

4. A possible extended design

A possible way to extend the design is to place a series of elements in series, similar to what is done in a Lyot filter. The most straightforward arrangement has factors of 2.0 between the FSRs, like in a Lyot filter. Wollaston prisms would be placed before the first element, between them, and after the last one, with calibration can be turned on and off without disturbing the absolute calibration will need to be determined. Another possibility would be to add the absolute calibration signal in an unused part of the spectrum, such as in the ultraviolet, infrared, or parts of the spectrum heavily contaminated by telluric lines.

3.3. Disadvantages

While the proposed design has several advantages, it also has some disadvantages. One obvious disadvantage is that the complexity of the instrument is increased. The hope is that this is more than compensated for by the ability to calibrate the entire spectrum consistently and possibly reduce the requirements on the inherent stability of the instrument. It may also be offset somewhat by a reduced resolution and thus size of the spectrograph.

Four separate spectra are now generated, compared to one for the iodine approach and two for the ThAr and laser comb approaches, requiring a larger CCD area. To some extent, this is compensated for by the possible reduction in the spectrograph resolution, but this improvement is nowhere near a factor of four or even two. In addition to the increased CCD area, the additional spectra will also increase the impact of readout noise, which may or may not represent a significant issue.

It should be straightforward to reduce the factor to three by overlapping the spectra B and C in Fig. 1. Reducing the number of spectra to two is substantially more difficult. One solution would be to introduce an additional amount of birefringence in one of the two beams exiting the first Wollaston prism such that they have approximately the same polarization entering the birefringent element. Using the angular dependence of birefringence and using one straight beam and another offset by a large angle or tilting the two beams by a large angle can both achieve this. However, to overlap the resulting pairs of spectra, these two beams must be parallel through the device, and it is unclear how this can be achieved simultaneously.

In the present design, the two beams overlap almost completely, and using the spatial separation is therefore likely impossible. Separating the two beams completely should be straightforward using a combination of polarizing beamsplitters and mirrors. Unfortunately, it is unclear how these two beams can be made to overlap afterward. If this is not done, the beam size will be doubled in one direction, which will have massive impacts on the downstream spectrograph.

In summary, this can be achieved simultaneously.

The simplest approach to obtaining the time-dependent component is likely to add the calibration device to a spectrograph with an iodine cell, possibly modified to have weaker absorption lines than normal. The iodine spectrum then provides the absolute calibration, and the calibration device extends this to the part of the spectrum that is not covered by the iodine lines. A similar approach may, of course, also be used with ThAr light in separate spectra, but this negates the advantage of having a common optical path for the stellar light and the calibration signal. It also requires $2 \times 4 = 8$ subspectra per order.

Since the drifts are expected to be very slow, it is tempting to consider only calibrating infrequently, such as at the beginning and end of each observing night. In this way, the light loss from an iodine cell, for instance, can be avoided, but whether the
the deviation angle doubled between each. The result would be $2N+1$ spectra, with $N$ being the number of elements. With $N = 2$ elements, one would see four pairs of spectra (as for the $N = 1$ case previously considered, there will be pairs of identical spectra), multiplied by

$$(1 + \cos(x))(1 + \cos(2 + x))$$

where $x = 2\pi(\lambda - \lambda_0)/FSR_0$. FSR0 is the FSR of the first element, and the contrasts have been assumed to be unity.

Fig. 8 shows the performance of such a system with various numbers of elements. The performance indeed improves as the number of elements is increased. Unfortunately, the performance stops increasing as more elements are added, ending with an efficiency significantly below unity. In other words, it is not possible to use this idea to make a high spectrograph resolution.

The FSR required to achieve the best performance decreases with both increasing resolution and increasing number of elements. However, with many elements, the dependence of the performance on the FSR is quite weak as long as the FSR resolves the line. With seven elements, the ideal FSR is roughly 0.06 Å at $R=10^4$ and 0.02 Å at $R=10^3$. Yet, as illustrated in Fig. 8, the performance is almost as good at the FSR of HMI, which is 0.172 Å. This also illustrates that a buildable set of elements achieves a good performance. Essentially, the only change required to the HMI filters would be to replace the polarizers between the elements by Wollaston prisms.

5. Conclusion

Birefringent elements and Michelson interferometers represent a promising way to calibrate cross-dispersed echelle spectrographs. Based on experience from solar instrumentation, it has been shown that the stability of such filters is adequate for calibration at 0.1 m/s level over a substantial wavelength range. While the approach shows great promise, an actual spectograph has not been constructed, and there are still a number of outstanding issues and undoubtedly many details that need to be worked out. One of the first issues to address will be to obtain accurate estimates of the temperature and wavelength dependence of the birefringence for the various candidate materials. Similarly, the uniformity of the optical properties, including the birefringence, need to be measured. Simply adding this device to an existing spectrograph is not likely to be possible, as there is not enough space between the orders to accommodate four spectra. Having said that, it may be possible to make a proof of concept using an existing spectrograph. If one of the Wollaston prisms is replaced by a simple polarizer, only two spectra will be generated, which can likely be accommodated in a spectrograph using a parallel calibration spectrum.

Making a proof of concept will hopefully identify any problems that have not been identified so far. Having done this and with the detailed measurements in hand it should be possible to decide on various tradeoffs and make a more detailed design.

Acknowledgements. The author would like to thank Matthias Ammler-von Eiff, Tim Brown, Patrick Gaulme, Jack Harvey, Rick Hessman, Greg Kopf, Martin Bo Nielsen, Andreas Querrenbach, Angar Reiners, Dick Sheehan, and the anonymous referee for useful suggestions, discussions and help with various datasets.

References
Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
Bauer, E. F., Zechmeister, M., & Reiners, A. 2015, A&A, 581, A117
Brown, T. M. 1990, in Astronomical Society of the Pacific Conference Series, Vol. 8, CCDs in astronomy, ed. G. H. Jacoby, 335–344
Brown, T. M., Noyes, R. W., Nisenson, P. K., Korzennik, S. G., & Horner, S. 1994, PASP, 106, 1285
Butler, R. P., Marcy, G. W., Williams, E., et al. 1996, PASP, 108, 500
Coudévid, S., Schou, J., Sheehan, R. A., et al. 2012, Sol. Phys., 275, 285
Domingo, V., Fleck, B., & Poland, A. I. 1995, Sol. Phys., 162, 1
Edelstein, J., Erskine, D., Sirk, M., Vanderburg, A., & Wishnow, E. H. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 4
Erskine, D. J., Edelstein, J., Mairhead, P., et al. 2011, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8146, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 0
Erskine, D. J., Edelstein, J., Sirk, M., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 17
Eversberg, T. & Vollmann, K. 2015, Spectroscopic Instrumentation: Fundamentals and Guidelines for Astronomers
Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
García-Marrirodriaga, C., Pascos, E., Strandmoe, S., et al. 2021, A&A, 646, A121
Ghosh, G. & Bhar, G. 1982, IEEE Journal of Quantum Electronics, 18, 143
Gonzáles, J., Schmidt, F. X., Gaulme, P., et al. 2019, Icarus, 319, 795
Halverson, S., Roy, A., Mahadevan, S., & Schwab, C. 2015, ApJ, 814, L22
Harvey, J. W., Hill, F., Hubbard, R. P., et al. 1996, Science, 272, 1284
Hoeksema, J. T., Baldwin, C. S., Bush, R. I., Schou, J., & Scherrer, P. H. 2018, Sol. Phys., 293, 45
Jennings, J., Terrien, R., Fredrick, C., et al. 2020, OSA Continuum, 3, 1177
Loupis, C. & Pepe, F. 2007, A&A, 468, 1115
Mahadevan, S., Ge, J., Fleming, S. W., et al. 2008, PASP, 120, 1001
Metcalf, A. J., Anderson, T., Bender, C. F., et al. 2019, Optica, 6, 233
Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, Sol. Phys., 275, 3
Phillips, D. F., Glenday, A. G., Li, C.-H., et al. 2012, Optics Express, 20, 13711
Reiners, A., Banyal, R. K., & Ulbrich, R. G. 2014, A&A, 569, A77
Reiners, A., Mroitzek, N., Lemke, U., Hinrichs, J., & Reinsch, K. 2016, A&A, 587, A65
Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, Sol. Phys., 162, 129
Schmieder, F. X., Gay, J., Gaulme, P., et al. 2007, A&A, 474, 1073
Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, Sol. Phys., 275, 229
Schwab, C., Sturmer, N., Gurevich, Y. V., et al. 2015, PASP, 127, 880
Solanki, S. K., del Toro Iniesta, J. C., Woch, J., et al. 2020, A&A, 642, A11
Title, A. M. 2010, ISSI Scientific Reports Series, 9, 327
Title, A. M., Pope, T. P., Ramsey, H. E., & Schoolman, S. A. 1976, Development of birefringent filters for spaceflight, Unknown
Title, A. M. & Ramsey, H. E. 1980, Appl. Opt., 19, 2046
Tomczyk, S., Mathew, S. K., & Gallagher, D. 2016, Journal of Geophysical Research (Space Physics), 121, 6184
van Eyken, J. C., Ge, J., & Mahadevan, S. 2010, ApJS, 189, 156
Vogt, S. S. 1987, PASP, 99, 1214
Wan, X. & Ge, J. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7734, Optical and Infrared Interferometry II, ed. W. C. Danchi, F. Delplancke, & J. K. Rajagopal, 77343N
Wang, J., Ge, J., Wan, X., Lee, B., & De Lee, N. 2012, PASP, 124, 598
Wildi, F., Pepe, F., Chazelas, B., Lo Curto, G., & Lovis, C. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, ed. I. S. McLean, S. K. Ramsay, & H. Takami, 77354X
Wilken, T., Curto, G. L., Probst, R. A., et al. 2012, Nature, 485, 611
Appendix A: Practicalities

The use of birefringent elements and imaging Michelsons is not common. Therefore I discuss a number of details in this appendix. The details are relevant to the assembly of such elements and to how they may be incorporated into a spectrograph.

Appendix A.1: Placement and size

One might think that the element can be placed almost anywhere in the optical path, but unfortunately, both types of elements have an angular dependence of $\lambda_0$. For simple elements such as those described here, the dependence is quadratic in angle and thus focal ratio.

A calcite birefringent element will change by about $\pm 3 \, \AA$ across an f/10 beam, which will wash out the modulation at the FSRs considered here. As discussed in Couvidat et al. (2012), this can be improved substantially by splitting the element into two, with a half-wave plate in between, resulting in a $0.15 \, \AA$ shift from the center to the edge of an f/10 beam. However, as it is difficult to make waveplates perfectly achromatic, the reduction will be imperfect at some wavelengths.

For an imaging Michelson, the uncompensated-for shift is similar to the birefringent elements. However, in this case, the angular dependence can be reduced to fourth order by a suitable choice of refractive indices and leg lengths, thereby making the effect negligible, even for fast beams. A downside of this arrangement is that the arms must be made of different materials, leading to other problems, and the problems with making achromatic waveplates.

One solution to the angular dependence problem is to place the element in a collimated beam. Unfortunately, the beam sizes of the collimated beams in large spectrographs are generally quite large compared to the size of the elements used so far in the solar case, and they likely exceed the available crystal sizes. Having said that, a Michelson could potentially be made large enough.

A better solution to this problem is probably to make a short section of collimated light shortly after the fiber. This arrangement has the advantage that the beam can be made very small, which will also facilitate the temperature control.

The fast beams generally used both at the fiber output and near the camera mean that it is almost certainly impossible to use birefringent elements there. An imaging Michelson could potentially be made large enough and have a low enough angular dependence to be placed elsewhere in the beam, such as near the fiber or immediately before the detector.

For a wide-fielded birefringent element, the angular dependence of the wavelength is

$$\delta v \approx \frac{c}{4n_e^2 n_o} \frac{n_o - n_e}{n_e} \theta^2,$$

where $\theta$ is the angle away from normal, $n_o$ is the refractive index of the ordinary ray, and $n_e$ is that of the extraordinary ray. From this it follows that the element should be mounted with an angular stability of around 0.00018 radians or 0.6 arcmin to achieve a stability of 0.1 m/s. For a perfectly wide-fielded Michelson, the leading-order term is fourth order, and so the required stability will be substantially reduced, likely limited by the accuracy of the widefielding.

A side effect of the proposed design is that the light falling on the Echelle will be linearly polarized. Whether this is an advantage or a disadvantage is to be determined. If circular is preferred, this can be achieved (at least approximately) by adding a quarter-wave plate, again subject to the availability of achromatic waveplates. As pointed out by Halverson et al. (2015), a varying polarization can lead to significant calibration problems, and so the well-defined and extremely stable polarization in the present design may represent a significant advantage.

A possible problem with birefringent elements in a noncollimated beam is that they may introduce optical aberrations (astigmatism) and that the different polarizations (the four spectra) may have different imaging properties.

Appendix A.2: Birefringent elements

An important question to answer is what materials to use. Clearly, some amount of birefringence is needed. However, it need not be very high. The HMI E1 element, which has been used for illustration here, consists of 30.62 mm calcite and 7.056 mm ADP for a total length of less than 40 mm, which should be easy to accommodate in a spectrograph, even if mounted in a highly stable oven. Given this, other materials with somewhat lower birefringence may also be considered.

As already mentioned, the temperature and wavelength dependence of the birefringence is poorly known for many of the relevant materials. Given this, one of the first tasks will be to identify reliable measurements or to make laboratory measurements.

If large elements are needed, a serious consideration is materials availability. The birefringent elements used for solar instrumentation have been about 35 mm in diameter, and it is unclear if calcite crystals substantially larger than that can be obtained. Apart from availability, the fact that calcite, ADP, and KDP are soft and moderately hygroscopic is another reason to consider alternative materials.

For temperature compensation, and if needed, wide-fielding, multiple pieces need to be used. To avoid losses and reflections, it is desirable to contact these pieces. This must be done with extreme care. Not only are materials like calcite optically anisotropic, their thermal expansion is also highly anisotropic and different between the materials. If two elements are glued together and subjected to a modest temperature change, the stresses can cause the crystals to break. Indeed, the original MDI Lyot filter was completely destroyed by a 10 to 15 C temperature change. To avoid this, the elements can be greased (oiied) together. The redesigned MDI filter, which was assembled in this manner, was by accident subjected to more than 100 C of temperature change without breaking.

Appendix A.3: Imaging Michelsons

In the case of Michelsons, there are no issues with material availability for any conceivable size. However, achieving good temperature compensation does require a careful selection of materials.

A practical problem with fabricating Michelsons is the reliance on polarizing coatings and waveplates, both of which can have a significant wavelength dependence. At some wavelengths, the coating may not split the light 50/50 into the two linear polarizations and the waveplates will deviate from quarter-wave retardance. One result of this is that the contrast will decrease. This is not necessarily a significant problem, as it will mostly affect the calibration noise and as there is a large margin. A more significant problem is that some of the light may be lost by being reflected back out the entrance face of the Michelson, resulting in increased noise. The extent to which carefully de-
signed coatings and wide-band waveplates can reduce this problem will need careful modeling.