High-precision calibration of spectrographs

T. Wilken,1* C. Lovis, 2 A. Manescau, 3 T. Steinmetz, 1,4 L. Pasquini, 3 G. Lo Curto, 3 T. W. Hänisch, 1 R. Holzwarth 1, 4 and Th. Udem 1

1 Max-Planck-Institut für Quantenoptik, Hans-Kopfermann–Strasse 1, 85748 Garching, Germany
2 Observatoire de l’Université de Genève, 51 ch. des Maillettes, 1290 Versoix, Switzerland
3 European Southern Observatory, Karl-Schwarzschild-Strasse 3, 85748 Garching, Germany
4 Menlo Systems GmbH, Am Klopferspitz 19, 82152 Martinsried, Germany

Accepted 2010 March 9. Received 2010 February 22; in original form 2009 November 9

ABSTRACT
We present the first stringent tests of a novel calibration system based on a laser frequency comb (LFC) for radial velocity measurements. The tests were obtained with the high-resolution, optical spectrograph, High Accuracy Radial velocity Planet Searcher. By using only one echelle order, we obtain a calibration repeatability of 15 cm s\(^{-1}\) for exposures that are several hours apart. This is comparable with a simultaneous calibration using a Th–Ar lamp that makes use of all 72 echelle orders. In both cases, the residuals are compatible with the computed photon noise. Averaging all LFC exposures, recorded over a few hours, we could obtain a calibration curve with residuals of 2.4 m s\(^{-1}\). Thanks to the adjustable and optimally chosen line density of the LFC, we resolve a periodicity of 512 pixels in the calibration curve that is due to the manufacturing process of the CCD mask. Previous Th–Ar calibration was unable to resolve these systematic deviations, resulting in a deviation of up to 70 m s\(^{-1}\) from the true calibration curve. In future, we hope to be able to make use of all echelle orders in order to obtain a calibration repeatability below 1 cm s\(^{-1}\) and absolute calibration within a few m s\(^{-1}\).

Key words: instrumentation: detectors – instrumentation: spectrographs – techniques: radial velocities – techniques: spectroscopic.

1 INTRODUCTION
Recent years have seen a growing interest in precision spectroscopy in astrophysics (e.g. Santos et al. 2008). Among the many applications considered, some are quite demanding in terms of spectrograph performance, such as the discovering of Earth-like extrasolar planets or the direct detection of cosmic acceleration. The majority of the extrasolar planets discovered so far have been detected by radial velocity (RV) techniques, that is, detecting the planets by measuring the small periodic Doppler shift modulation induced by their orbital motion in the host stellar spectrum. Most of those planets are quite massive and of short period, although a vast population of Neptune-mass objects seem to be present around a considerable fraction of stars. There is a strong interest in finding planets in habitable zones and Earth twins. The RV variations induced by the Earth in the solar spectrum are however very small: only 9 cm s\(^{-1}\) with a 1-yr period.

An extremely high precision is also required for the direct measurement of the cosmic acceleration. Such a temporal evolution of the Hubble constant may be detected by observing the Doppler shift of Ly\(\alpha\) forest absorption lines towards quasi-stellar objects (QSOs) over a reasonable period (20–30 yr). The cosmic acceleration of suitable objects (QSOs at redshifts below \(\sim 5\)) is expected to be of the order of about 1 cm s\(^{-1}\) yr\(^{-1}\) (Liske et al. 2008), corresponding to a frequency shift of 20 kHz at a wavelength of 500 nm (600 THz).

Another critical application that will benefit from improved calibration is the search for a possible variability of the physical constants (e.g. Webb et al. 1999; Reinhold et al. 2006). The basic measurement consists of determining the difference in wavelength when comparing different transitions in the intergalactic medium in the absorption spectra towards QSOs.

In order to measure these tiny variations in Doppler shift, the spectrograph must have the best possible repeatability, i.e. the wavelength calibration must be sufficiently stable during the observation time. Absolute calibration is not required at the same level as repeatability, provided that systematic shifts are constant during this time. Comparing observations over very long periods of time will make it very difficult to maintain all parameters of the spectrometer constant at a few cm s\(^{-1}\). While up to now largely ignored, absolute calibration of a spectrograph may hence become a key ingredient when comparing measurements from different epochs and telescopes.

*E-mail: tobias.wilken@mpq.mpg.de
So far Th–Ar lamps and iodine absorption cells have been used to generate a large set of calibration lines. Irregularities in line spacings and intensities, blending and ageing limit their repeatability and the lines’ absolute frequencies are only known to the $10^{-7}$ level [see Murphy et al. (2007) for a thorough analysis of the different calibrators, discussing the weak points of each]. Today even the best spectrographs are limited in stability, due to their calibration light source, to several 10 cm s$^{-1}$ (e.g. Lovis & Pepe 2007). Hence, when designing new spectrographs, a better calibration source is needed. A very promising approach is based upon the use of spectrally filtered laser frequency combs (LFCs), which are the best candidates to become the ideal tool for improved calibration (e.g. Murphy et al. 2007; Braje et al. 2008; Li et al. 2008; Steinmetz et al. 2008).

An LFC is capable of generating thousands of equidistant, yet tuneable, calibration lines with equal intensity. Both repeatability and absolute frequency uncertainty depend only on the atomic clock to which each line is referenced, reaching easily below the required $10^{-10}$ level. Their potential for astronomy has been shown in the lab (Braje et al. 2008; Li et al. 2008) and, as a proof of principle, also in the infrared (IR) at a telescope (Steinmetz et al. 2008).

The main challenge is to generate an LFC with a sufficiently large mode spacing that can be resolved by a typical astronomical spectrograph. Even if for precision measurements the highest spectral resolution is sought, feasibility and efficiency considerations limit de facto the resolving power of astronomical (non solar) spectrographs to $R \sim 100,000$, or about 5 GHz. Fabry–Perot cavities (FPCs) can serve as mode filters to increase the LFC’s fundamental mode spacing which for fibre lasers, like the ones we use, is typically below 1 GHz. Fibre amplifiers can replenish the power lost in the filtering process to enable non-linear processes such as frequency doubling and spectral broadening necessary to match the LFC’s spectrum to the spectrograph’s optical bandwidth.

In this work, we demonstrate that a precision spectrograph such as the High Accuracy Radial velocity Planet Searcher (HARPS, e.g. Mayor et al. 2003; Lovis & Pepe 2007) can be calibrated with significantly lower uncertainty using an LFC. Although we use only one of HARPS’s 72 echelle orders, the repeatability is already on a par with a Th–Ar calibration using all orders. The absolute wavelength calibration surpasses the Th–Ar calibration by a factor of 20. Spectrally broadening the LFC to cover all orders will further improve the repeatability and enable calibrations on the cm s$^{-1}$ scale.

Section 2 of this work is devoted to the description of the test set-up, while in Section 3 the analysis and the results are described.

### 2 THE SET-UP

#### 2.1 The HARPS spectrograph

The calibrations were performed with the HARPS spectrograph at European Southern Observatory’s 3.6-m telescope at La Silla, Chile. The spectrograph consists of a white pupil design, grism crossdispersed R4 echelle, housed in a temperature-controlled vacuum vessel. With a 1 arcsec aperture on the sky it has a resolving power of $R = 115,000$, corresponding to 5.2 GHz at 515 nm. It has been especially designed for superior stability (e.g. Mayor et al. 2003). HARPS covers the 380–690 nm range simultaneously, distributing the light over 72 dispersed echelle orders on to a mosaic of two 2K × 4K CCD detectors. A second fibre is used to record simultaneously a Th–Ar reference spectrum to monitor the spectrograph’s residual instabilities. Thanks to its excellent performance – namely a repeatability below 1 m s$^{-1}$ (3 × 10$^{-9}$) over several years (e.g. Lovis et al. 2006) – a large number of Neptunian exoplanets and super-Earths have been discovered with this instrument (Bouchy et al. 2009, fig. 4).

#### 2.2 The laser comb system

The LFC is based on a mode-locked Yb-doped fibre laser with a repetition rate of $f_r = 250$ MHz and a central wavelength of 1030 nm. The optical frequencies of the LFC modes are given by $m f_r + f_0$, where $m$ is a large integer (of the order $10^6$) and $f_0 < f_r$ is a radio frequency comb offset (e.g. Udem, Holzwarth & Hänsch 2002). Two FPCs with free spectral ranges of $m_1 f_r$ and $m_2 m_1 f_r$ are used in series to suppress the unwanted LFC modes (e.g. Murphy et al. 2007) (see Fig. 1). The modes transmitted by the FPCs are then given by $f_n = n m_1 m_2 f_r + f_0 = n f_r f_0 + f_0$, with the larger mode spacing $f_r = m_1 m_2 f_r$ and the integer filter ratios $m_1$ and $m_2$. The comb offset after filtering is $f_r f_0$ and might differ from $f_0$ by a multiple of $f_r$. Just like $f_r$, the common LFC offset frequency $f_0$ is controlled by a Rb atomic clock with an accuracy of $5 \times 10^{-11}$ (1.5 cm s$^{-1}$). The mode number $n$ which resolves the remaining 250 MHz ambiguity was determined by comparison with the Th–Ar calibration curve of the CCD.

Using more than one filter cavity has the advantage of achieving a higher suppression of the unwanted modes (Steinmetz et al. 2009). The cavities were tilt-locked to the reflection signal of a continuous wave laser (Shaddock, Gray & McClelland 1999), which

![Figure 1. Experimental set-up. A Yb-fibre-based laser frequency comb was filtered with two Fabry-Perot cavities (FPC 1 + 2), which results in a mode spacing of $f_r = 18$ GHz. In the following second harmonic generation stage the centre wavelength frequency was doubled to 515 nm. Traversing a fibre shaker, the comb light was fed together with light from a conventional Th–Ar calibration lamp via the telescope’s calibration unit (TCU) through the TFP to the HARPS spectrograph. F1, F2 and F3 refer to the fibres between the different units, having a core diameter of 1 mm, 300 μm and 70 μm, respectively. At the telescope’s focal plane (TFP), apertures can be set, such that light coming from any of the F2 fibres or the telescope can be coupled to any of the F3 fibres. Fibre amplifiers situated before, between and directly after the FPCs are not illustrated. They compensate for the power losses due to the rejected modes of the filter stages.](image-url)
was locked to the LFC and whose frequency was monitored with a wavemeter. This ensures that the same comb modes were continuously transmitted by the FPCs. With \( m_1 = 8 \) and \( m_2 = 9 \), we provide an LFC with a mode spacing of 18 GHz for calibration, which is 3.4 times the resolution of HARPS and slightly above the theoretical optimum (Murphy et al. 2007). Calibration at 12 GHz was tested, but we opted for a slightly larger mode spacing, which minimizes the overlap between the wings of adjacent modes, to facilitate the analysis. The cavity mirrors have a reflectivity of 99.2 per cent, resulting in a finesse of \( \sim 400 \). Both cavities are plano-concave, where the radii of curvature of the concave mirrors are chosen to minimize the transmission of higher order spatial modes as described by Steinmetz et al. (2009). The suppression of the unwanted modes is calculated to be better than 50 dB, which was confirmed by radio frequency spectra.

Three fibre amplifiers (before, between and after the filter FPCs) provide enough power for frequency doubling in a periodically poled potassium–titanyl–phosphate crystal. The calibration light generated is centred at around 514.9 nm with a full width at half-maximum of 2.1 nm. The spectrum covers one echelle order with 362 modes, and its width is presently limited by the crystal’s phase matching bandwidth. In future, this will be improved by broadening the spectrum in a highly non-linear fibre.

### 2.3 The fibre interface

We used a multimode fibre (F1) to couple the LFC light to the HARPS calibration unit which also contained the Th–Ar lamp. The output facet of F1 was positioned at the focal plane of a lens system that illuminates one out of two multimode fibres (F2). The lens system was designed for the Th–Ar lamps and has a demagnification of 10:1, projecting the LFC light on to a 100 μm spot on the input facet of the F2 fibres. Each fibre then projects a spot of 750 μm diameter on the telescope’s focal plane (TFP). The projected light is collected by one out of two fibres (F3) that bring the light to the spectrograph. At this point, light from each of the F2 fibres or starlight from the telescope can be coupled to any of the F3 fibres. In the path to the spectrograph, a passive optical scrambler exchanges the F3 fibre near-field and far-field. At the spectrograph’s entrance, the F3 fibres are displaced along the direction of the grating rules such that their images are \( \approx 15 \) pixels apart in the cross-dispersed direction on the CCD.

### 3 Analysis and results

The LFC calibration curve is determined by fitting individual comb modes with Gaussians, which serve as a good approximation of the instrument’s point spread function (PSF) (see Fig. 2). The linewidth of the LFC modes is of the order of 1 MHz which could be reduced if required. It can be neglected in comparison with the PSF’s width of about 5 GHz. We determine the drift of the spectrograph through the mean shift of the observed line positions between different acquisitions, weighted with the uncertainty of each line position (derived from the fit). Alternatively, the cross-correlation between different exposures can be used to determine the global shift of the spectrum on the detector. Various algorithms to determine the drift of the spectrograph performed equally well and the results agreed within the statistical uncertainty. The calibration with Th–Ar is determined using the method described by Bouchy, Pepe & Queloz (2001). This method cannot be used to measure the shift of the comb calibration as only a global flux variation can be tolerated.

![Figure 2. Laser frequency comb calibration spectrum. A part of the extracted comb spectrum](image)

A single, 80 s exposure was integrated to arrive at the given photoelectron count rates. The inset shows one full echelle order (out of 72) with the LFC’s spectral envelope. The range of the zoomed region is highlighted.

The calibration repeatability between subsequent 80 s exposures is determined relative to an arbitrarily selected reference acquisition. We do this with the LFC as well as with the Th–Ar spectra, and Fig. 3 depicts the results for a series of 50 acquisitions. The top panel shows that, using only one echelle order and the LFC, we can determine a Doppler shift variation of an object with a statistical uncertainty of 18 cm s\(^{-1}\) within 80 s. The middle panel shows the same data for the Th–Ar calibration using all 72 available echelle orders. Calculating the difference between the LFC and Th–Ar data (shown in the bottom panel) results in a slightly reduced SD of 15 cm s\(^{-1}\) as compared to the LFC data. This is due to a partial correlation of the data, caused by the drift of the instrument. After eliminating the instrument’s drift, the LFC calibration is at the calculated photon noise limit for a single 80 s exposure. Averaging over all 50 acquisitions, and thereby neglecting the small instrumental drift, a statistical uncertainty of 2.1 cm s\(^{-1}\) is obtained from the full LFC data set. While this can always be further reduced by taking more data, the repeatability will eventually be limited by systematic uncertainties and the ability to average for sufficiently long time periods given the instrument’s finite intrinsic stability.
within the statistical uncertainty.

Table 1 shows that intentional shifts of the LFC’s offset frequency 

<table>
<thead>
<tr>
<th>( \Delta f_0 ) (MHz)</th>
<th>Measured shift (MHz)</th>
<th>( \sigma ) (MHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.4</td>
<td>-0.48</td>
<td>0.24</td>
</tr>
<tr>
<td>-0.2</td>
<td>-0.33</td>
<td>0.32</td>
</tr>
<tr>
<td>0.4</td>
<td>0.22</td>
<td>0.29</td>
</tr>
<tr>
<td>2.0</td>
<td>2.09</td>
<td>0.39</td>
</tr>
<tr>
<td>4.0</td>
<td>4.19</td>
<td>0.17</td>
</tr>
<tr>
<td>8.0</td>
<td>7.93</td>
<td>0.21</td>
</tr>
<tr>
<td>16.0</td>
<td>16.08</td>
<td>0.26</td>
</tr>
<tr>
<td>120.0</td>
<td>119.97</td>
<td>0.22</td>
</tr>
</tbody>
</table>

3.2 Scanning the LFC

A very important advantage of the LFC over Th–Ar lamps is its tunability – via the mode spacing and offset frequency – such that CCD pixels can be scanned over individually. This feature provides the means to calibrate interpixel variations in size and sensitivity. Table 1 shows that intentional shifts of the LFC’s offset frequency ranging from 120 MHz down to 200 kHz (60 m s\(^{-1}\) to 10 cm s\(^{-1}\) in Doppler shift variation) can be traced reliably by the spectrograph which is again limited by the photon noise.

3.3 Impact of fibre coupling

To provide reproducibility for the ambitious long-term projects in astronomy, calibration with an absolute frequency scale seems to be the best solution. For this purpose, we have identified the largest systematic calibration uncertainty which is due to an incomplete wavefront matching between the star light and the calibration light. Lacking a workable scheme for adaptive optics that compensates atmospheric turbulences in the visible, multimode fibres must be used. This leads to two effects that can induce systematic calibration shifts. First, any inhomogeneous illumination of the fibre input facet will be partly preserved through the fibre and result in an inhomogeneous output pattern (e.g. Baudrand & Walker 2001). Secondly, for each frequency component spatial fibre modes can interfere, producing inhomogeneous output patterns (e.g. Grupp 2003) and spectral modulations that vary with fibre bending, temperature etc. Consequently, the barycentre of the beam entering the spectrograph may shift. The spectrograph’s imaging system cannot perfectly compensate for this effect, and the detected line positions are prone to a systematic shift.

While we cannot differentiate the two effects, we could reduce both by moving and bending the fibre during the exposures with a fibre shaker. The device actively scrambles the spatial fibre modes on a time-scale much shorter than the exposure time. Tests with different mode field diameters for fibre F1 showed best performance when using the largest core available (1 mm). With the fibre shaker enabled, the systematic shifts due to the multimode fibres are below the photon noise limit, while disabling the shaker the calibration repeatability of a single 80 s exposure increases to 30 cm s\(^{-1}\). Our results show that the design of future high-precision instruments should incorporate optimized fibre mode scrambling or, more ideally, single-mode fibres to mode match the calibration light with the incoming star light. The use of single-mode fibres coupled to an adaptive optics system, which has thus far been demonstrated for IR wavelengths, could avoid these problems without significantly sacrificing light collection efficiency.

3.4 Absolute frequency calibration

For HARPS, a calibration uncertainty of 15 cm s\(^{-1}\) corresponds to \(2 \times 10^{-4}\) pixels so that the pixel to frequency map (calibration curve) has to be known extremely well to enable an absolute calibration. Note, however, that this is required on average over the full spectrum only. Due to the sparse line distribution of Th–Ar lamps, CCD pixel inhomogeneities have so far been lost in the large-scale interpolations. Meaningful calibration curves are limited to about five parameters (fourth-order polynomial per echelle order; Lovis & Pepe 2007). Using the same procedure with the LFC calibrator, we could determine these CCD inhomogeneities for the first time in a calibration curve. The residuals of this fit plotted against the line position show discontinuities every 512 pixels (Fig. 4). These derive from the manufacturing process of the CCD which is written with a 512 pixel mask, leading to variations in size or sensitivity at the edges of the mask.

Including the discontinuities in the fitted calibration curve and increasing the polynomial order leads to a 2.4 m s\(^{-1}\) SD of the residuals. This value is significantly above the photon noise and measures the unmodelled pixel inhomogeneities and sensitivity variations. Comparing the Th–Ar and LFC calibration curves (blue line in Fig. 4), one finds large local deviations of up to 60 m s\(^{-1}\). Those can now be modelled using the LFC which significantly improves the absolute calibration.

An upper limit for the absolute frequency uncertainty can be given by measuring the Th–Ar line positions with the LFC calibration curve and comparing the results with the published values (Palmer & Engleman 1983). While the latter are only known to an uncertainty of about 50 m s\(^{-1}\), our much more accurate measurements are consistent within this limit. Assuming that the fibre coupling has an effect of probably less than 1 m s\(^{-1}\) (which can be concluded from the repeatability without mode scrambler), we are confident that the LFC’s calibration curve improves the absolute
Figure 4. Detector inhomogeneities unveiled by residuals of the calibration curve. Polynomials are fitted to the pixel versus frequency distribution and the residuals are plotted. Fitting functions to the LFC data are a global fourth-order polynomial (red dots) and eight piecewise eighth-order polynomials (green crosses) that cover 512 pixels each (72 parameters). Variations in pixel size, shape, position etc. which have their origin in the manufacturing process lead to the pattern with this period. The solid blue line shows the difference between the usual Th–Ar calibration curve (global fourth-order polynomial) and the LFC calibration curve (including the discontinuities). Due to the low line density of a Th–Ar lamp its calibration curve deviates strongly from the LFC calibration curve and discontinuities could previously not be detected.

4 CONCLUSIONS AND PERSPECTIVES

In summary, we have set up a frequency comb system with a mode separation and spectral bandwidth adequate for calibrating one echelle order of the HARPS spectrograph. We could demonstrate a repeatability compatible with the computed photon noise and comparable with a Th–Ar calibration but using only one echelle order instead of all 72. By using the LFC, we could measure the contribution of CCD pixel irregularities to the calibration curve for the first time. As a consequence, the spectrograph’s absolute calibration was improved by more than one order of magnitude. After spectrally broadening the LFC in a highly non-linear fibre, we aim to cover the whole visible range in near future. This system will further reduce the photon noise by more than an order of magnitude which is expected to enable Doppler shift measurements at the cm s$^{-1}$ level during one acquisition (80 s).

ACKNOWLEDGMENTS

We thank the La Silla Observatory and in particular Mr. Aldo Pizarro for the great support. TWH gratefully acknowledges support by the Max-Planck Foundation.

REFERENCES

Li C. H. et al., 2008, Nat, 452, 610
Lovis C. et al., 2006, Nat, 441, 305
Mayor M. et al., 2003, Messenger, 114, 20
Steinmetz T. et al., 2008, Sci, 321, 1335

This paper has been typeset from a TeX/LaTeX file prepared by the author.