

FLORIAN FRANZISKUS BAUER

**RADIAL VELOCITIES IN LOW-MASS STARS:  
IMPROVING THE WAVELENGTH SOLUTION OF ASTRO-  
NOMICAL SPECTROGRAPHS AND UNDERSTANDING  
STELLAR NOISE**

DISSERTATION





**Radial velocities in low-mass stars:**  
improving the wavelength solution of astronomical  
spectrographs and understanding stellar noise

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

“Doctor rerum naturalium”

der Georg-August-Universität Göttingen

im Promotionsprogramm PROPHYS

der Georg-August University School of Science (GAUSS)

vorgelegt von

**Florian Franziskus Bauer**

aus Pischelsdorf, Österreich

Göttingen, 2016

Betreuungsausschuss

---

Prof. Dr. Ansgar Reiners  
Institut für Astrophysik, Georg-August-Universität, Göttingen, Deutschland

Dr. Sandra Jeffers  
Institut für Astrophysik, Georg-August-Universität, Göttingen, Deutschland

Mitglieder der Prüfungskommission

---

Referent: Prof. Dr. Ansgar Reiners  
Institut für Astrophysik, Georg-August-Universität, Göttingen, Deutschland

Korreferent: Prof. Dr. Artie Hatzes  
Thüringer Landessternwarte Tautenburg, Physikalisch-Astronomische Fakultät,  
Friedrich-Schiller-Universität, Jena, Deutschland

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Stefan Dreizler  
Institut für Astrophysik, Georg-August-Universität, Göttingen, Deutschland

Prof. Dr. Wolfram Kollatschny  
Institut für Astrophysik, Georg-August-Universität, Göttingen, Deutschland

Prof. Dr. Claus Ropers  
IV. Physikalisches Institut, Georg-August-Universität, Göttingen, Deutschland

Prof. Dr. Andreas Tilgner  
Institut für Geophysik, Georg-August-Universität, Göttingen, Deutschland

Tag der mündlichen Prüfung: 09. Dezember, 2016

# Abstract

The radial velocity method has been used to discover hundreds of extrasolar planets in the past two decades. The continuous improvements in radial velocity precision have revealed several rocky planets in the habitable zones of low mass stars. However, a true Earth analogue has not yet been found because the precision needed for such a discovery is about one order of magnitude below what current state-of-the-art spectrographs can achieve. In order to find the radial velocity signal of an Earth-like planet in the habitable zone of a Sun-like star, several aspects of the radial velocity method have to be improved. This thesis focuses on two sectors: improving the wavelength solution of radial velocity spectrographs and understanding astrophysical noise sources, originating from spots on the stellar surface, that hinder the detection of small, Earth-sized exoplanets.

The radial velocity precision achieved by spectrographs is ultimately linked to their calibration. Currently hollow cathode lamps are used for wavelength calibration and state-of-the-art spectrographs have demonstrated precisions of 1 m/s by using these calibrators. To improve the wavelength solution of the next generation of instruments, Fabry-Pérot interferometers are already tested with current spectrographs. The device provides many lines for calibration but the absolute wavelength of Fabry-Pérot peaks are poorly constrained. Thus, these devices were only used for nightly drift checks up to now. This thesis presents a method to calibrate the Fabry-Pérot interferometers with absolute standards to use the dense grid of lines for the wavelength solution of echelle spectrographs. The HARPS and CARMENES instruments are a test ground for Fabry-Pérot interferometers. This work demonstrates that the use of Fabry-Pérot interferometers can substantially improve the wavelength calibration of these instruments. Hence, Fabry-Pérot interferometers are suitable to provide high precision calibration for the hunt of true Earth analogs with future spectrographs.

When the radial velocity precision of next generation instruments approaches the 10 cm/s level, astrophysical noise sources will become the limiting factor for planet detection. Hence, there is a need to understand and model this activity related radial velocity jitter. This thesis also presents simulations of radial velocity signatures originating from active regions like dark spots and bright plages on the stellar surface of solar type stars. Strong magnetic fields heat or cool active regions and hinder convection within them. To include convection effects in the simulation of radial velocity jitter, results from magneto-hydrodynamic models are incorporated. Observations of spot temperatures for stars other than the Sun are used in the models presented in this thesis to extend the parameter range for radial velocity jitter models to F and K type stars. With the results obtained in this thesis, the overall picture of activity related radial velocity jitter in the Sun and other stars improves so that in the future, these effects can be corrected when hunting for small, rocky exoplanets.



# Contents

<b>Abstract</b>	<b>i</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 High-precision wavelength calibration of echelle spectrographs</b>	<b>7</b>
2.1 Calibrators	7
2.1.1 Hollow cathode lamps	7
2.1.2 Laser frequency combs	8
2.1.3 Fabry-Pérot interferometers	11
2.2 Line spread functions and high-precision line positions	13
2.2.1 Overview	13
2.2.2 The concept of the effective line spread function (eLSF)	14
2.2.3 Reconstructing the eLSF	16
2.2.4 How accurate can we get? Comparing the performance of eLSF models	17
2.3 Using Fabry-Pérot etalons for wavelength calibration	23
2.3.1 Overview	23
2.3.2 Calibrating the Fabry-Pérot interferometer	23
Ideal FPI	23
Real FPI	24
2.3.3 Wavelength solution model	27
Mapping wavelengths to detector coordinates	28
Unequal pixel sizes	29
2.3.4 Calibrating HARPS using its FPI	30
Detector characterization and wavelength solution	30
Comparison to DRS	33
Comparison to LFC	33
Systematic high-frequency noise in the HARPS etalon	35
2.3.5 Influence of distortions in the wavelength solution on precise RV measurements	35
2.4 Calibrating the CARMENES spectrographs	37
2.4.1 The CARMENES survey	37
2.4.2 Instrument overview	38
2.4.3 CARMENES calibration resources	41
2.4.4 Line catalogs	42
2.4.5 Optimizing the hollow cathode lamp outputs	43
Lamp exposure times	43
Fiber position	46
2.4.6 Line catalog cross-referencing	47
2.4.7 The CARMENES Fabry-Pérot etalons	48
The CARMENES Fabry-Pérot etalons in the optics laboratory	48
Comparing the FPIs in CARMENES and the FTS	51

2.4.8	Systematics in the wavelength solution of CARMENES . . . . .	54
2.5	Summary . . . . .	58
2.6	Future outlook for further improvements . . . . .	60
<b>3</b>	<b>Simulating active stars</b>	<b>63</b>
3.1	Introduction . . . . .	63
3.2	Effect of active regions on stellar RVs . . . . .	65
3.2.1	Temperature . . . . .	65
3.2.2	Convection . . . . .	66
3.3	Methods . . . . .	69
3.3.1	Model grid . . . . .	69
3.3.2	Line profiles . . . . .	70
3.3.3	Disk integrated line profiles and radial velocities . . . . .	71
Step 1:	immaculate star . . . . .	72
Step 2:	spectra of active regions . . . . .	72
Step 3:	full disk integrated spectrum . . . . .	72
3.4	Results . . . . .	73
3.4.1	Influence of the magnetic field strength on RV curves . . . . .	73
RV curves	derived from MHD line profiles . . . . .	73
Convection	at different magnetic field strengths . . . . .	75
RV amplitudes	across different stellar types . . . . .	77
3.4.2	Comparison to the literature . . . . .	78
3.4.3	Influence of limb-angle-dependent convective blueshift and line profiles on RV curves . . . . .	79
The apparent	active region velocity across the disk . . . . .	79
Using Voigt	profiles with different convective blueshift parameters . . . . .	81
Changes	induced by convective blueshift models . . . . .	81
Changes	induced by line shape . . . . .	84
3.4.4	Influence of spot temperatures in F, G, and K stars . . . . .	87
Spot	temperatures . . . . .	87
The	symmetry parameter . . . . .	87
3.5	Summary and conclusion . . . . .	90
<b>4</b>	<b>Discussion and conclusion</b>	<b>93</b>
	<b>Acknowledgements</b>	<b>97</b>
	<b>Curriculum vitae</b>	<b>99</b>

# List of Figures

1.1	Fomalhaut b . . . . .	2
1.2	Tansits with Kepler . . . . .	3
1.3	Radial velocity method . . . . .	4
2.1	Image of a ThNe lamp . . . . .	8
2.2	Spectrum of a ThAr lamp . . . . .	9
2.3	Spectrum of the HARPS laser frequency comb . . . . .	10
2.4	Sketch of the laser frequency comb in Goettingen . . . . .	11
2.5	CARMENES Fabry-Pérot interferometers . . . . .	12
2.6	Line spread functions . . . . .	15
2.7	Line spread function reconstruction method . . . . .	18
2.8	Accuracy of measured line positions . . . . .	20
2.9	Phase error in position fitting . . . . .	21
2.10	Position error as a function of line FWHM . . . . .	22
2.11	Spectrum of a Fabry-Pérot interferometer . . . . .	23
2.12	Smooth cavity width . . . . .	26
2.13	Interference number of Fabry-Pérot lines . . . . .	27
2.14	HARPS wavelength solution residuals . . . . .	31
2.15	HARPS Fabry-Pérot cavity width . . . . .	32
2.16	HARPS wavelength solution differences . . . . .	34
2.17	RV error with a distorted wavelength solution . . . . .	36
2.18	CARMENES target sample . . . . .	38
2.19	HARPS and CARMENES wavelength coverage . . . . .	40
2.20	CARMENES VIS UAr spectrum . . . . .	43
2.21	CCD overspill in CARMENES VIS . . . . .	44
2.22	ThNe optimum exposure time . . . . .	45
2.23	UAr and UNe optimum exposure time . . . . .	45
2.24	Fiber position . . . . .	47
2.25	Cavity width of CARMENES VIS etalon measured with FTS . . . . .	49
2.26	Cavity width of CARMENES NIR etalon measured with FTS . . . . .	50
2.27	Cavity width of CARMENES VIS etalon measured in CARMENES . . . . .	52
2.28	Cavity width of CARMENES NIR etalon measured in CARMENES . . . . .	53
2.29	CARMENES VIS wavelength solution residuals . . . . .	55
2.30	CARMENES NIR wavelength solution residuals . . . . .	56
3.1	Schematic of the flux effect of spots and plages . . . . .	67
3.2	Schematic of the convective blueshift effect of spots and plages . . . . .	68
3.3	Influence of magnetic field strength on spot RVs . . . . .	74
3.4	Influence of magnetic field strength on plage RVs . . . . .	74
3.5	Active region velocity field as a function of $B$ . . . . .	76
3.6	RV model comparison . . . . .	79
3.7	Active region velocity field as a function of $\mu$ . . . . .	80
3.8	Comparison of spot RVs derived from MHD- and Voigt line profiles . . . . .	82

3.9 Comparison of plage RVs derived from MHD- and Voigt line profiles . .	83
3.10 FeI line profiles at different $\mu$ . . . . .	85
3.11 FeI line perturbations at different $\mu$ . . . . .	86
3.12 Spot temperature measurements . . . . .	88
3.13 Symmetry parameter . . . . .	89



# List of Tables

2.1	Pixel sizes of the HARPS CCD . . . . .	33
2.2	Table of CARMENES calibration sources. . . . .	41
3.1	Stellar parameters of simulated stars . . . . .	70
3.2	Active region velocity field as a function of $B$ . . . . .	75
3.3	RV difference using different convective blueshift models . . . . .	84



# List of Abbreviations

<b>HCL</b>	<b>H</b> ollow <b>C</b> athode <b>L</b> amp
<b>ThAr</b>	<b>T</b> horium <b>A</b> rgon
<b>ThNe</b>	<b>T</b> horium <b>N</b> eon
<b>UAr</b>	<b>U</b> ranium <b>A</b> rgon
<b>UNe</b>	<b>U</b> ranium <b>N</b> eon
<b>LFC</b>	<b>L</b> aser <b>F</b> requency <b>C</b> omb
<b>FPI</b>	<b>F</b> abry- <b>P</b> érot <b>I</b> nterferometer
<b>FTS</b>	<b>F</b> ourier <b>T</b> ransform <b>S</b> pectrograph
<b>rms</b>	root <b>m</b> ean <b>s</b> quare
<b>RV</b>	<b>R</b> adial <b>V</b> elocity
<b>HARPS</b>	<b>H</b> igh <b>A</b> ccuracy <b>R</b> adial velocity <b>P</b> lanet <b>S</b> earcher
<b>CARMENES</b>	<b>C</b> alar <b>A</b> lto high- <b>R</b> esolution search for <b>M</b> dwarfs with <b>E</b> xo-earths with <b>N</b> ear-infrared and optical <b>E</b> chelle <b>S</b> pectrographs
<b>CRIRES</b>	<b>C</b> Ryogenic high-resolution <b>I</b> nfra <b>R</b> ed <b>E</b> chelle <b>S</b> pectrograph
<b>SPIRou</b>	un <b>S</b> pectro <b>P</b> olarimètre <b>I</b> nfra- <b>R</b> ouge
<b>HPF</b>	<b>H</b> abitable zone <b>P</b> lanet <b>F</b> inder
<b>IRD</b>	<b>I</b> nfra <b>R</b> ed <b>D</b> oppler instrument for the Subaru Telescope
<b>ESPRESSO</b>	The Echelle <b>S</b> Pectrograph for <b>R</b> ocky <b>E</b> xoplanets and <b>S</b> table <b>S</b> pectroscopic <b>O</b> bservations
<b>CODEX</b>	<b>C</b> Osmic <b>D</b> ynamics and <b>E</b> Xo-earth experiment
<b>MHD</b>	<b>M</b> agneto <b>H</b> ydro <b>D</b> ynamic



# Chapter 1

## Introduction

Humanity has been speculating about the presence of other worlds and other forms of life different from what we know on Earth since the ancient times (Anaximander, 610 BC). In the 16<sup>th</sup> century Giordano Bruno proposed that stars are distant Suns that could host their own planets and possibly even life. However, it took hundreds of years until these philosophical ideas were followed by the first discovery claim of an extrasolar giant planet (Mayor & Queloz, 1995). That exoplanets were discovered only recently is related to the weak signals planets produce; modern high precision instruments are required for detection. Today more than 3400 exoplanets are known and new ones are discovered at a rate of one every few days<sup>1</sup>. Nevertheless, the majority of these planets are believed to be hostile to life because they are either too big, too hot, or always face their host star with the same hemisphere so that there are extreme temperature differences between day- and night-side. Discovering a planet similar to Earth where life could have possibly emerged is currently one of the biggest scientific goals of humanity because it could provide us the answer to the question: ‘are we alone?’ This fascination with other worlds has led to several methods to discover smaller and smaller planets around other stars, but all of these methods fall just short of finding a true Earth analogue.

One method to discover exoplanets is direct imaging. Planets emit most of their radiation in the infrared and also reflect a fraction of the host starlight. Therefore planets are not completely dark objects and can be detected by imaging. An example of a planet discovered with direct imaging is Fomalhaut b (Kalas et al., 2008) shown in Fig. 1.1. However, imaging an exoplanet is difficult because planets are extremely faint compared to their host stars. For example, Earth is about  $10^7$  times fainter than the Sun if observed in the infrared and the situation becomes even worse in the optical where the contrast diminishes to  $10^{10} : 1$  (Seager & Deming, 2010). Although modern instruments like SPHERE at VLT are designed to achieve contrast ratios of up to  $10^6$  in the near-infrared (Beuzit et al., 2008), indirect exoplanet detection methods (transits or radial velocity searches) still lead discovery statistics.

Opposed to direct imaging, the transit method benefits from the large star-planet contrast. As the planet passes between the observer and the star, a fraction of the stellar light is blocked, resulting in dimming of the star (Henry et al., 2000; Charbonneau et al., 2000). In Fig. 1.2, a schematic image of transit observations can be seen. Transit searches have made a large number of exoplanet detections mostly thanks to space born missions like CoRoT (Baglin & COROT Team, 1998) or Kepler (Basri et al., 2005). Space missions are expensive but it is necessary to avoid Earth’s atmosphere in transit searches because the atmosphere is the limiting noise source for high precision brightness measurements. The Kepler mission has discovered several Earth sized planets in the habitable zone of small host stars (Torres et al., 2015), but a true Earth analogue

---

<sup>1</sup><http://exoplanet.eu>

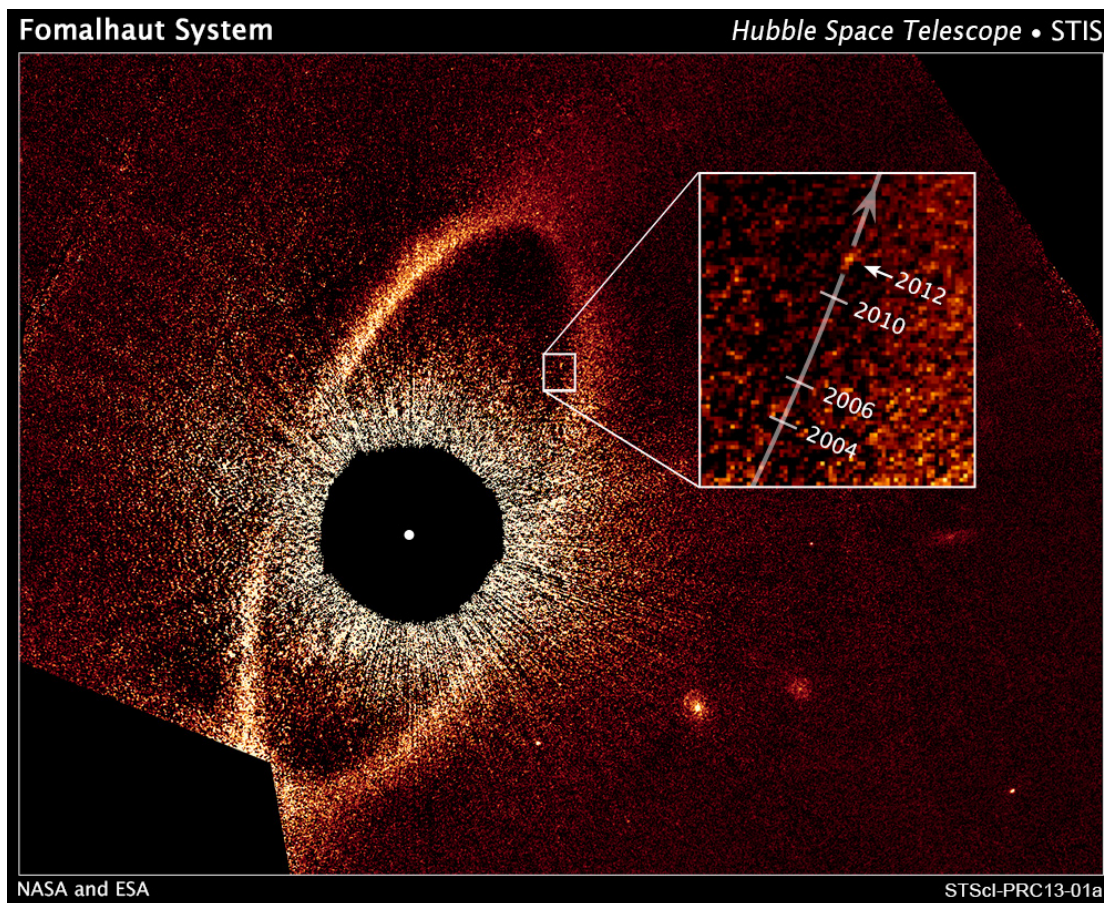


FIGURE 1.1: Picture of Fomalhaut b taken with the Hubble space telescope. Light from the central star is blocked by a coronagraph mask (central black area). A debris disk is visible as bright ring. The planet itself circles the star at the inner edge of the debris disk and is visible as dot, slightly brighter than the background. The planet was detected in multiple observations and the zoom-in image shows the planet's location on its orbit in 2004, 2006, 2010 and 2012. Image Credit: <http://www.nasa.gov>

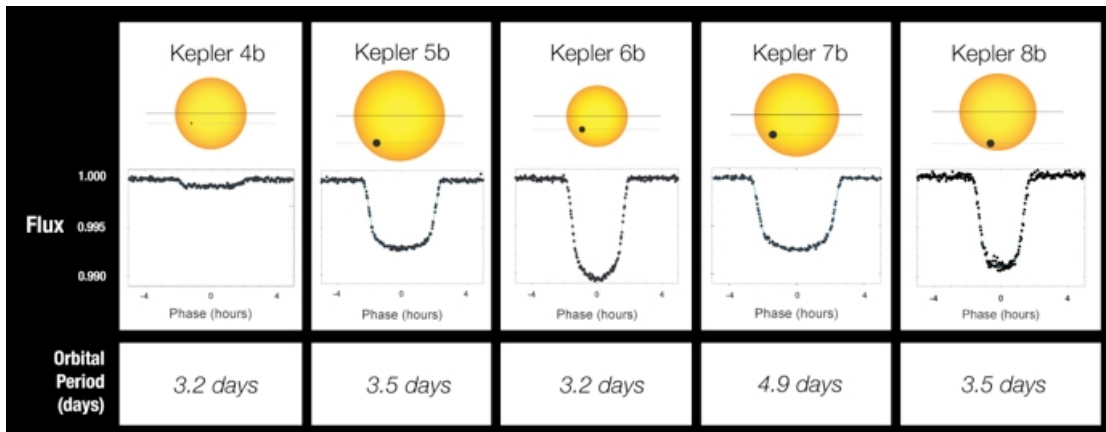


FIGURE 1.2: Schematic of transit observations with the Kepler space telescope. Several examples of Kepler discoveries are shown. The host stars and their planets differ in size so that the occultation of the star results in transit signals with different amplitudes in all cases. Image Credit: <http://kepler.nasa.gov>.

was just beyond the capabilities of the instrument. However, a next generation transit spacecraft, PLATO (Rauer et al., 2014), will be launched with the goal to detect Earth-sized planets in the habitable zones of solar type stars. Transits give access to the planet size, but no dynamical information can be extracted from the lightcurves. Hence, orbital parameters and planet properties like mass can only be determined by complementary techniques such as radial velocity observations.

The radial velocity (RV) technique was the first one to successfully detect a planet around a main sequence star (Latham et al., 1989). Stars orbited by planetary companions show a reflex motion around the systems' common center of mass. The stars' orbital velocity depends on the mass ratio between the host and companion and their distance. For example, Jupiter causes the Sun to move with a velocity of about 12 m/s while Earth imposes only 9 cm/s. As shown in Fig. 1.3, light emitted by the host star becomes Doppler shifted according to the stars' reflex motion for an observer who is outside the system. This Doppler shift can be made visible if the starlight is dispersed as a stellar spectrum. Stellar spectra have spectral lines caused by atoms absorbing light at specific wavelengths. These spectral lines act as markers in the RV method. Periodic red- or blueshifts of spectral lines indicate a motion of the star caused by the presence of an unseen, gravitationally interacting companion (Perryman, 2011).

Because Doppler shifts are measured with spectral lines, RV observations require spectrographs to disperse the light and give access to the spectral features. All RV instruments are generally designed as cross-dispersed echelle spectrographs (Baranne et al., 1996). This design allows for a high resolution (typically  $R > 50000$ ) paired with a large wavelength coverage (typically a few 1000 Å). High resolution is needed because the radial velocity information is in the gradient of the spectral features (Bouchy et al., 2001). Low resolution smears out the spectral features resulting in a loss of RV content. A larger wavelength coverage increases the number of spectral lines observed and therefore more RV information is collected by the observations. Current state-of-the-art RV spectrographs measure Doppler shifts with a precision down to about 1 m/s and keep this stability over many years (e.g. Dumusque et al., 2012). Currently, modern instruments use pixel detector technology like CCDs to record the stellar spectra. A Doppler shift of 1 m/s translates to a shift of the spectrum typically on the order of

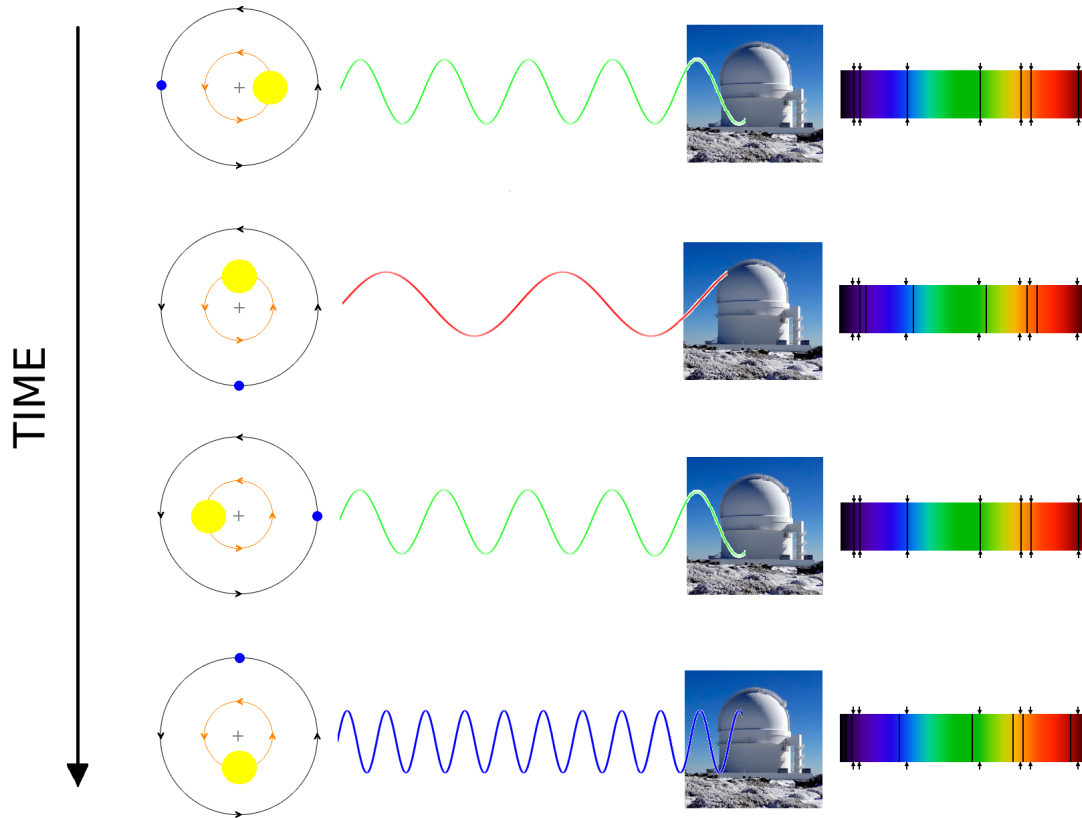


FIGURE 1.3: A schematic depiction of the radial velocity method. Host star and planet are indicated as yellow and blue circles respectively. The orange and black circles show the stars and planets orbit around the systems center of mass schematically. Each panel (from top to bottom) corresponds to a snapshot in time. When the host star moves perpendicular to the observers line of sight (snapshot 1 and 3), the observed spectrum is at rest and the radial velocity signal is 0. When the host star moves away from the observer (snapshot 2), the observed spectrum becomes red-shifted and a positive velocity is measured. The effect is opposite when the host star moves towards the observer (snapshot 4), which results in a blueshift of the spectrum and a negative velocity is measured.



1/1000 of a pixel. To achieve this precision, two different RV measurement strategies have been developed and successfully used for planet detections.

The first strategy is the iodine method (Marcy & Butler, 1992). A gas cell filled with iodine ( $I_2$ ) is inserted into the telescope beam before it enters the spectrograph. Hence, iodine absorption lines are superimposed on the stellar spectrum. The idea of this calibration method is to use the iodine lines as stable reference. If the stellar lines shift relative to the iodine lines, the cause is astrophysical and inherent to the star. If both the iodine and the stellar lines shift in the same way, the reason for the apparent Doppler shift is instrumental. Temperature changes in the spectrograph room change the optical alignment of components and can therefore result in apparent Doppler shifts. Pressure changes affect the refractive index of air, which then also results in apparent Doppler shifts. The iodine lines track all changes in the spectrograph and thus instrumental drifts can be calibrated. This makes spectrographs using the iodine method relatively simple and cheap. On the downside, about 20 - 30 % of light are lost because of the cell. Further, most of iodines' spectral information is contained in the region between 5000 and 6200 Å limiting the spectral range usable for RV measurements to a relatively narrow band (Perryman, 2011).

A second method to measure high precision RVs is to build ultra-stable spectrographs (Mayor et al., 2003). In this design all optical elements are placed inside a vacuum tank that is temperature and pressure stabilized. Therefore spectrograph drifts are minimized. To keep the spectrograph enclosed, light from the telescope is fed through an optical fiber to the instrument. Instead of using a gas cell as a calibration reference, light from a calibration unit is fed through a second fiber to the spectrograph. Hence, these stabilized RV instruments always record two spectra at the same time: the stellar spectrum and a calibration reference. Both fibers are very close to each other so that the light from the calibration fiber takes a similar path through the instrument as the stellar light contained in the science fiber. Any small changes of temperature or pressure inside the vacuum tank that result in small drifts of the instrument can be measured with the spectrum of the calibration reference. With stable spectrographs there are no light losses in gas absorption cells and no contamination of the stellar spectrum by superimposed calibration lines. However, the gain in efficiency comes with a high price tag because the optical design and stabilization are expensive. Additionally, stabilized instruments are very complex and temperature/pressure controlled environments are crucial for success.

To discover a true Earth analogue with an RV amplitude of 9 cm/s, new developments in ultra-stable spectrographs aim for a precision below the m/s level. In the optical, ESPRESSO at VLT is aiming for 20 cm/s (D'Odorico & CODEX/ESPRESSO Team, 2007) and CODEX at ELT is aiming for 2 cm/s (Pasquini et al., 2008). At the same time, the first infrared RV instruments are on the way. Among them are CARMENES (Quirrenbach et al., 2011), SPIROU (Artigau et al., 2011), IRD (Kotani et al., 2014), and HPF (Mahadevan et al., 2010). All these new projects face the same challenges to achieve their extreme precisions. On the one hand, building an instrument thermally stable and pressure controlled to a level enabling precisions below the m/s level is a difficult engineering task. On the other hand, calibrating the spectrograph to a cm/s level requires new and better standards. Reducing all instrumental systematics will unveil various astrophysical signals that hinder planet detections. These astrophysical noise sources need to be understood in detailed in order to mitigate them and enable planet detection (Fischer et al., 2016). This thesis is dedicated to methods that improve the calibration of modern radial velocity instruments and increase our understanding of astrophysical jitter that awaits us below the m/s level.

Chapter 2 describes the wavelength calibration of astronomical echelle spectrographs. We first introduce different wavelength standards and then describe how line positions of different sources can be determined with high accuracy. A technique how to use Fabry-Pérot interferometers (FPI) for improvements in wavelength calibration is introduced and applied to HARPS and CARMENES. We discuss our findings about detector imperfections and distortions found in the wavelength solution of the two high precision RV instruments.

With the next generation of extreme precision instruments, radial velocity jitter induced by stellar activity is a concern for planet detection. In Chapter 3 we describe how activity related surface features of stars produce RV signals and present simulations to estimate activity related RV curves. Results from magneto-hydrodynamic simulations are used to predict the influence of dark spots and bright plages on RV measurements of stars with spectral types F to K.

## Chapter 2

# High-precision wavelength calibration of echelle spectrographs

### 2.1 Calibrators

Accurate wavelength calibration is one of the most important requirements in obtaining high-precision RV measurements because measurements of the Doppler shift of in spectra requires a wavelength scale for the observations. Nowadays, state-of-the-art echelle spectrographs use pixel detectors like silicon CCDs for the optical or mercury cadmium telluride arrays (HgCdTe) for the near-infrared. Spectra are therefore recorded in the pixel domain of these detectors and the observations come without a direct wavelength scale. Standards of known wavelength are therefore needed to translate the pixel scale into the wavelength domain.

#### 2.1.1 Hollow cathode lamps

Hollow cathode lamps (HCLs) provide such standards. Since HCLs were invented more than a century ago in 1916 by Paschen, their design has constantly improved. Today HCLs consist of a cylindrical metal cathode (widely spread are thorium (Th) or uranium (U)) enclosed in a glass body which is filled with a buffer gas (common are argon (Ar) or neon (Ne)) (Huke et al., 2016; Crosswhite et al., 1955). Figure. 2.1 shows an image of a ThNe HCL in the optics laboratory in Göttingen.

As explained in Huke et al. (2016) if a voltage is applied to the lamp the negative potential of the cathode accelerates electrons away from it. Collisions between the accelerated electrons and the buffer gas result in ionized gas atoms which in turn accelerate towards the cathode. When ions hit the cathode, metal atoms are sputtered out and excited, resulting in the emission of narrow atomic emission lines when the electrons relax back into a lower energy state. As an example, Fig. 2.2 shows the spectrum of the HARPS ThAr lamp. To use the metal emission lines from a HCL as wavelength standards for astronomical echelle spectrographs, their line positions must be known with high accuracy and precision. Therefore spectra of the lamps are taken with high-resolution (typically  $R > 100000$ ) Fourier transform spectrographs (FTS) to produce standard line lists. (Palmer & Engleman, 1983; Lovis & Pepe, 2007; Kerber et al., 2008, and Redman et al., 2011, 2014). In some of these catalogs the measured line positions are combined with Ritz wavelength calculations. Typical uncertainties for the wavelength catalogs are a few 10 m/s (Redman et al., 2011, 2014). Although the low price tag and years of successful usage have made HCLs the most widely used standards in astronomical spectroscopy, they have a number of drawbacks.

- The number of lines is limited to the number of energy transitions within the cathode material. Even though the number of lines can be very large (about 10000

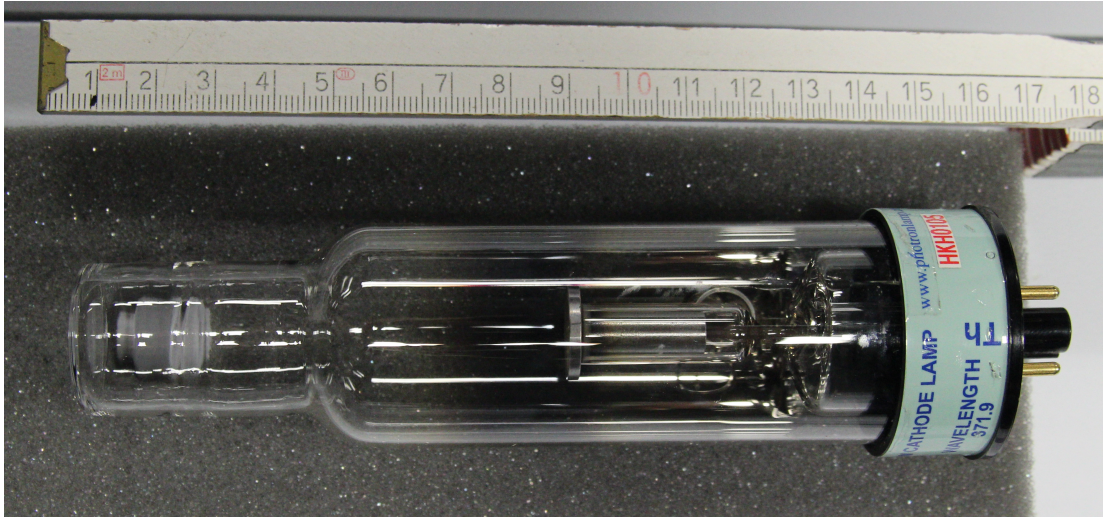


FIGURE 2.1: Image of a ThNe lamp in the optics laboratory in Göttingen

for thorium between 250 nm and 5500 nm; [Redman et al., 2011](#)) only 4000 – 8000 remain usable for optical high-resolution spectrographs ([Lovis & Pepe, 2007](#)). The rest of the lines are either too faint or are part of blends (groups of lines too close to be resolved by the spectrograph).

- Spacing of the lines is an additional problem. Regions with only weak or even without lines have to be interpolated by the wavelength solution of the spectrographs which gives uncertain results in poorly populated areas ([Bauer et al., 2015](#)).
- The buffer gas is also excited by free electrons and collisions with ions once the lamp is turned on. Hence not only metal lines but also strong gas emission lines are emitted by HCLs. Gas emission features are several orders of magnitude brighter than cathode lines and contaminate the calibration spectrum. Modern detectors are unable to handle such large dynamic ranges and therefore regions with strong gas lines are saturated and not usable for calibration, as seen in [Fig. 2.2](#). There are ways to decrease the dynamic range such as increasing the operational current of the lamp but this significantly reduces the lifetime. In practice, tradeoffs for the exposure time have to be carefully chosen for HCLs in order to keep the number of saturated gas lines low while maximizing the number of cathode lines available for calibration.

Spectrographs using HCLs as calibrators have demonstrated RV precisions on the order of 1 m/s (e.g. [Dumusque et al., 2012](#)). However, for the next generation of instruments aiming for cm/s precision ([D’Odorico & CODEX/ESPRESSO Team, 2007](#)) these lamps are insufficient and new standards are needed.

### 2.1.2 Laser frequency combs

Laser frequency combs (LFCs) are currently under development to overcome the problems of HCLs. An LFC is a device that produces equidistantly spaced spectral lines. Compared to HCLs (see [Fig. 2.2](#)) an LFC offers a higher number of calibration lines at a constant density. The dynamic range of LFCs is also high like in HCLs but several methods have been developed to flatten the spectra and achieve almost constant



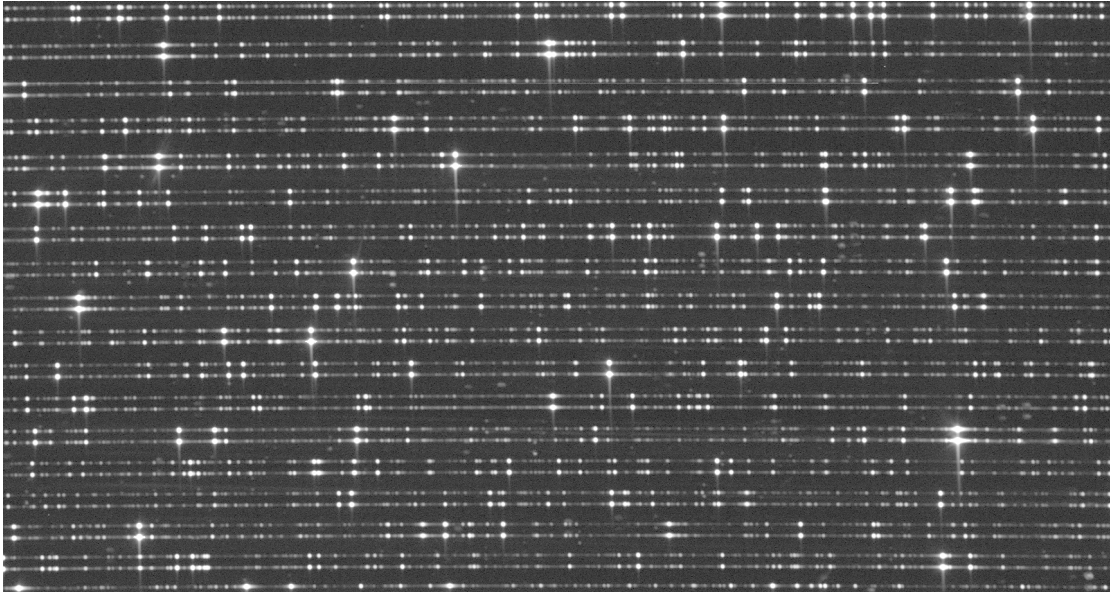


FIGURE 2.2: Raw echelle spectrum of the ThAr calibration lamp taken by HARPS.

amplitudes for all lines (Probst, 2015). Hence there are no saturation problems and no empty regions without lines. As an example the HARPS LFC spectrum is depicted in Fig. 2.3.

Another advantage of LFCs is that they can be adapted to the spectrographs they are designed for. The frequency of the lines produced by an LFC are only determined by two parameters :

$$f_n = f_0 + n * f_r, \quad (2.1)$$

where  $f_0$  is the offset frequency,  $f_r$  is the repetition rate and  $n$  is the mode number. The line spacing of a frequency comb is determined by the repetition rate  $f_r$  of the laser. In theory, the calibration source can therefore be built according to the instrument resolution by choosing a laser with the right repetition rate that produces well separated lines. In practice, however, lasers with repetition rates beyond the GHz regime are not yet available so that line separations required for astronomical spectrographs are not achievable without line filtering techniques.

As explained in Wilken (2010) and Probst, 2015, two center pieces are needed to produce the typical LFC pattern of sharp, equidistantly spaced, unresolved lines: a high power pulsed laser and a nonlinear fiber. The laser pulses with a frequency  $f_r$  in the MHz to GHz regime and emits light pulses with a length of a few fs. This results in trains of light a few cm long and with a spectral bandwidth of a few tens of nm. The temporal signal of the laser is fed into a crystal fiber with a very high dispersion. This non linear fiber stretches the wavelength range of the pulses and performs optical Fourier transformation to produce the frequency comb pattern.

For high-precision wavelength calibration, the comb pattern must be stable in frequency to below 1 m/s (Wilken et al., 2012). If the pattern moves constantly during one exposure, lines will smear out and asymmetric line shapes might be the result. Hence, both the offset frequency  $f_0$  and the repetition rate  $f_r$  need to be actively stabilized, which requires two active feedback loops (Wilken, 2010). The repetition rate  $f_r$  is detected by measuring a fraction of the laser light with a photo diode. The laser repetition rate  $f_r$  is then adjusted to a reference frequency. In practice, the adjustment of the laser

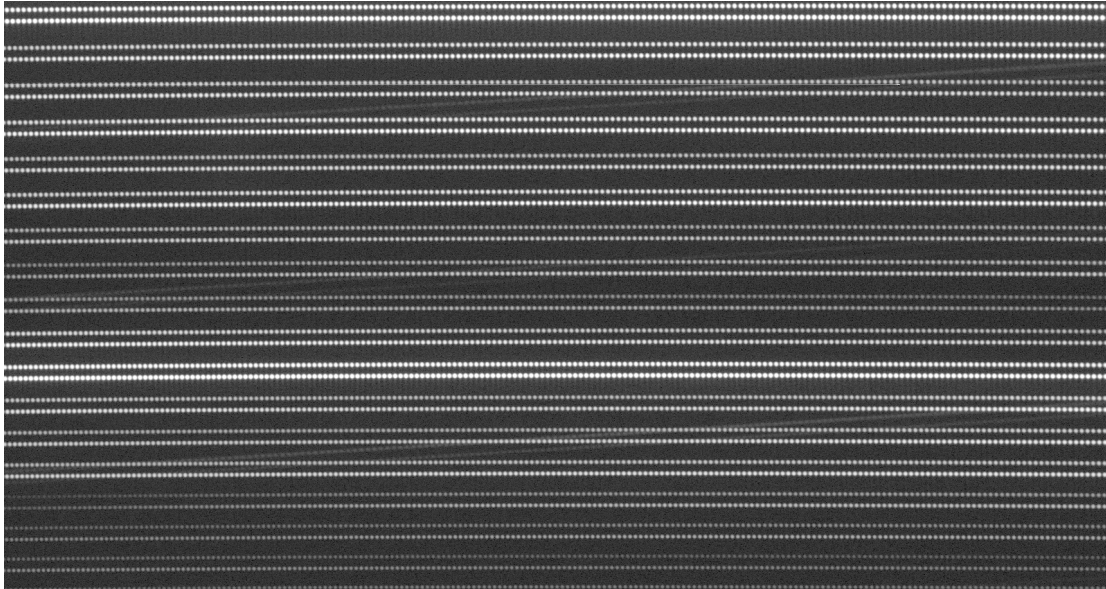


FIGURE 2.3: Raw echelle spectrum of the HARPS LFC.

repetition rate is done by adjusting the cavity width of the laser with a piezo crystal.

Stabilizing the offset frequency  $f_0$  is more elaborate and requires the comb to span at least one optical octave to apply self referencing. The laser light passes a beam splitter and one part of the beam is frequency doubled by a non-linear crystal. When combined again both beams form a beat signal with a frequency equal to  $f_0$  which is again detected by a photo diode. Within the time the laser pulse takes to do one round trip in the laser cavity  $T_r$ , the phase between the carrier wave and the pulse envelope shifts by  $\Delta\phi$ . This defines the offset frequency  $f_0$  by the relation:

$$f_0 = \frac{\Delta\phi}{2\pi T_r}, \quad (2.2)$$

To adjust and stabilize  $f_0$  in practice means controlling the group and phase velocity of the laser pulse. This is done by controlling the dispersion within the laser cavity either by shifting wedged glass in or out with a piezo crystal, and/or by controlling the pump power of the laser.

Figure 2.4 shows a schematic of the basic LFC setup in the laser lab in Göttingen. The pulsed laser is denoted with Taccor 1GHz. The non-linear fiber is highlighted in the blue box with broken lines on the right and is named FemtoWhite. The self referencing loop to stabilize  $f_0$  is highlighted in the blue box with broken lines in the lower center part denoted f-2f interferometer.

Although LFCs are thought to be the ideal standards for future high-precision RV work, there are also some disadvantages with these advanced standards. Currently lasers cannot produce repetition rates that allow LFCs to produce a grid of lines that can be resolved by astronomical echelle spectrographs. With current technology, the frequency comb lines are packed too closely so that line filtering is required. To filter most of the comb lines and leave only a small percentage of them for calibration requires advanced and stabilized Fabry-Pérot filters. The setup is complicated, expensive and requires constant maintenance. As many RV instruments are located on small to medium class telescopes with modest budgets, the price tag becomes important. Many projects could thus benefit from a less complex and less expensive alternative to

## Experimental Setup

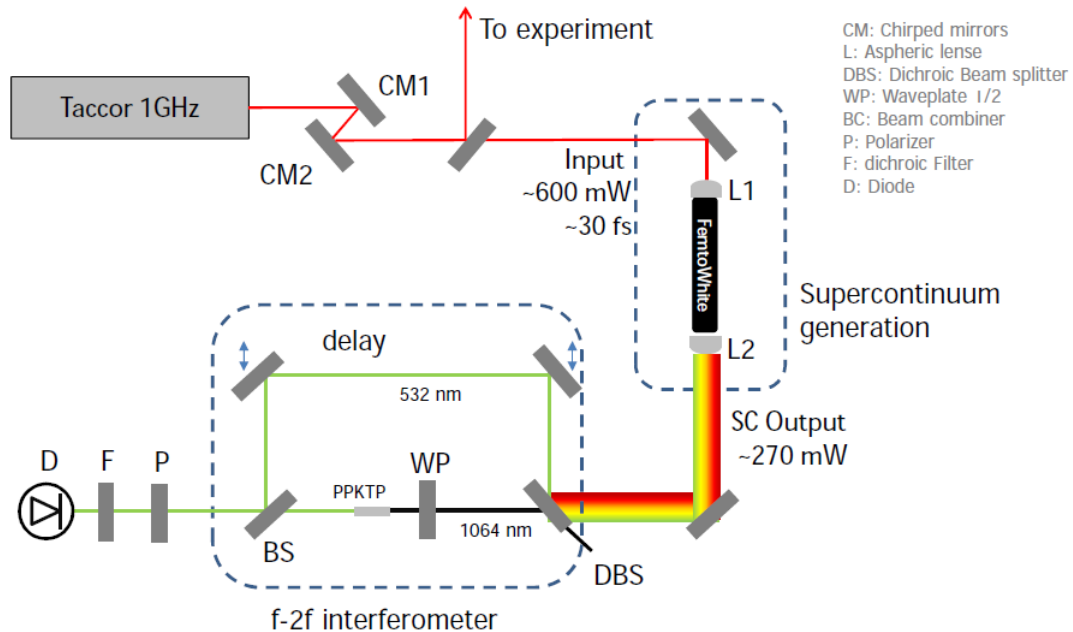


FIGURE 2.4: Schematic of the LFC set up in Göttingen

the LFC.

### 2.1.3 Fabry-Pérot interferometers

Fabry-Pérot interferometers (FPIs) are an alternative to LFCs. FPI devices are very simple and only consist of two plane-parallel, partly reflective mirrors (Born & Wolf, 1999). The interference condition of such a setup produces maxima at wavelengths that are integer multiples of the cavity width between the two mirrors. Similar to the LFC, FPIs offer a dense line grid which is very well suited for calibration. FPIs were discovered in the early 1980s for precision RV measurements. First, Cochran et al. (1982) used an FPI to superimpose well defined lines on the stellar spectrum that could be used for simultaneous calibration of a non-stabilized spectrograph. Later the same idea was used by Butler et al. (1996) with iodine cells in transmission; a technique that is still successfully used today. Kitt Peak hosted a high-resolution spectrograph dedicated to planet searches with a tunable FPI as the dispersing element and a demonstrated precision of 6 m/s (McMillan et al., 1994, 1988). After that, FPIs disappeared for about 15 years in planet hunts because their use in transmission results in extremely low instrument efficiency. However, FPIs were reconsidered for HARPS where such a device is currently used to monitor the spectrograph drift during observations (Wildi et al., 2010). Being able to use FPIs for wavelength calibration can be beneficial for many smaller RV projects because the setup is much cheaper than buying and maintaining a complex LFC.



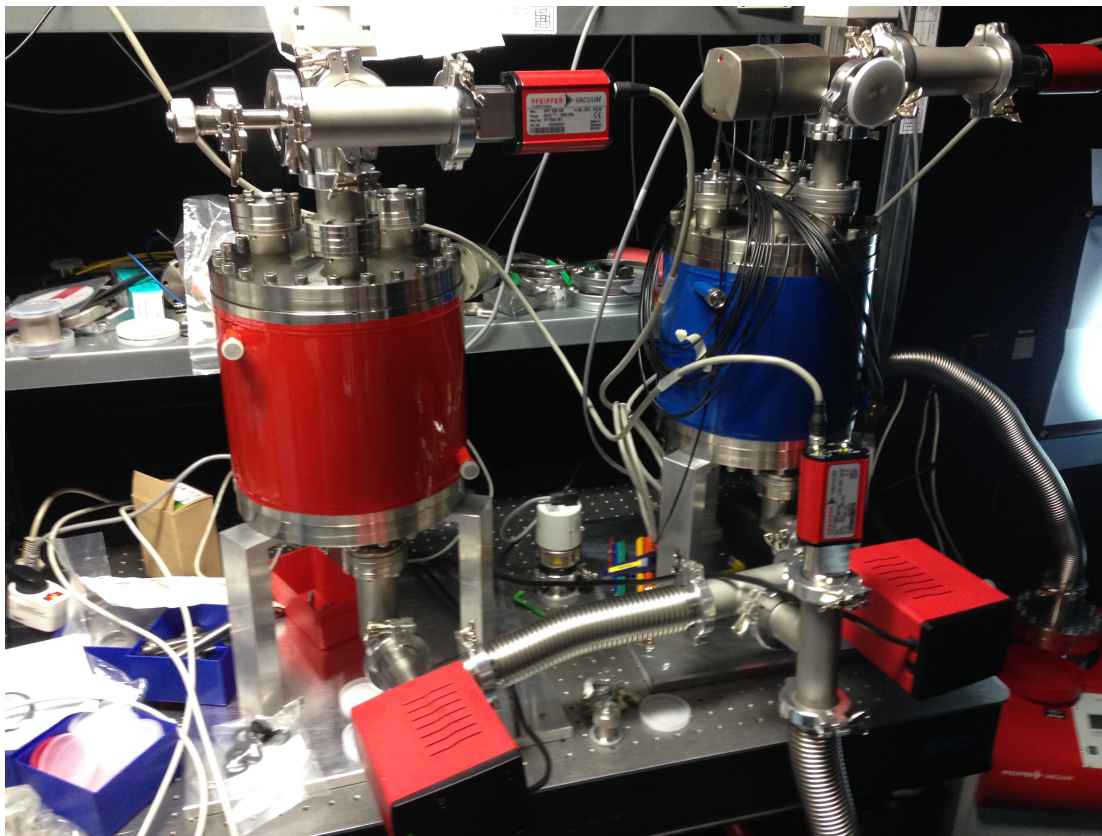


FIGURE 2.5: CARMENES Visible (blue) and near-infrared (red) FPIs inside their vacuum tanks. The setup as it is seen here was used for testing the FPIs in the optics laboratory in Göttingen and was later shipped to Calar Alto (Spain).

However, before CARMENES became operational in November 2015, there have been no attempts to use FPIs in astronomical spectrograph calibration. The main reason is that unlike HCLs and LFCs, FPIs are not absolute standards. Accurate line positions for calibration require exact knowledge of the cavity width and the interference order of the transmission maxima. The cavity width is especially critical since the accuracy with which this property can be measured is orders of magnitudes below the requirements for m/s calibration. Conventional distance measurements have access to  $\mu\text{m}$  accuracies, while for the purpose of using the FPI as an absolute reference the cavity width must be known to about  $0.1 \text{ \AA}$  accuracy. In addition, external influences like temperature or pressure changes let the etalon cavity drift. Temperature expansion of materials change the mirror distance and hence influence line positions. Furthermore both pressure and temperature influence the refractive index. Refractive index changes result in changes of the effective cavity width seen by the photons which in turn also influence the line positions. Schäfer & Reiners (2012) showed that FPI lines are stable to within  $10 \text{ cm/s}$  if the temperature change is kept below  $18 \text{ mK}$  and the pressure change is less than  $3 \cdot 10^{-6} \text{ bar}$ . To achieve the required temperature and pressure stability, FPIs are kept within a vacuum vessel that is actively temperature and pressure stabilized (Wildi et al., 2010; Schäfer & Reiners, 2012). As an example of how such an FPI setup looks like, Fig. 2.5 shows the setup with the vacuum tanks for CARMENES before the equipment was shipped to Calar Alto.

Recently etalons have been successfully referenced to Rubidium cells to achieve



even higher stabilities of 3 cm/s (Schwab et al., 2015). Because FPIs are relatively cheap, even including the vacuum tank and temperature control, they are already installed in many current spectrographs and are foreseen also for upcoming instruments, e.g. CRIRES+ (Follert et al., 2014) or SPIROU (Artigau et al., 2011).

As FPIs seem to become standard equipment for next generation instruments, this chapter introduces a method to use etalons in wavelength calibration and improve the RV precision. First, Sect. 2.2 presents a method to determine the effective line spread function of a spectrograph to enable high-precision line position fitting. In Sect. 2.3 a method to reference FPIs to absolute standards is presented and the principles of how to use the FPI comb pattern for high-precision wavelength calibration are shown. The high information content of the calibrated FPI is incorporated into the wavelength solution of HARPS and the results are compared to the standard wavelength solution available from the ESO pipeline. In Sect. 2.4 an overview of the CARMENES calibration procedure is given. Details about the fine tuning of the calibration plan are presented and results of using the FPI approach in daily calibration of the two CARMENES spectrographs are shown.

## 2.2 Line spread functions and high-precision line positions

### 2.2.1 Overview

Astronomical high-resolution spectrographs are typically implemented as cross dispersed echelle spectrographs. Light entering these instruments is diffracted by the grating and the resulting spectrum is imaged onto a detector. In echelle spectrographs the wavelength scale in the focal plane is not known and therefore it is not known which detector pixel records which wavelength. Hence, astronomical spectrographs need calibration that connects each pixel to the wavelength it sees. This wavelength calibration is done using wavelength standards which are well known. In the case of HCLs the wavelength of spectral lines are given in catalogs and are measured with high accuracy in laboratories (see e.g. Redman et al., 2011). For LFCs, the wavelengths of lines can be calculated. The positions of the spectral standard lines on the detector are then measured and a relation between pixel position and laboratory wavelength determines the wavelength scale of the detector. The first task in wavelength calibrating astronomical spectrographs is therefore always to determine the pixel position of spectral lines on the detector of the instrument.

As the goal of RV surveys is to detect small, Earth like planets, spectrographs need to achieve a precision on the order m/s or below. This requires that all calibration steps are done at least at the m/s level but fitting line positions with m/s accuracy or below is not trivial at all. In astronomical echelle spectrographs, calibration lines are typically sampled with about 6 to 12 pixels. A shift of a spectral line by 1 pixel is equivalent to about 1 km/s in radial velocity. Thus, line centers need to be determined to about 1/1000 of a pixel to reach the m/s requirement. Temperature, pressure, mechanical stress and other external effects influence the way a spectrograph system images the wavelengths onto the focal plane. Hence, the same wavelengths do not always fall onto the same detector pixels. Because of external influences also the focus of the instrument can change slightly resulting in changes in the full width at half maximum (FWHM) of the lines. These spectrograph drifts in combination with the poor sampling of resolution elements make line fitting a challenging task. Nevertheless, line position fitting needs to be as robust as possible because wavelength calibration is repeated on a daily basis and is thus crucial to achieve long-term precision in the order of m/s.

In this section a new technique called ‘effective line spread function reconstruction’ to measure accurate line positions on the detector is investigated. The concept of the effective line spread function is explained and the method how to reconstruct the effective line shape of calibration lines is shown in detail. Further, it is investigated how accurately line positions can be determined using different models for the line shape for different calibrator sources.

## 2.2.2 The concept of the effective line spread function (eLSF)

Whenever line positions have to be measured in a spectrum that is recorded by a detector, it is important to first understand what is actually recorded by the detector and then choose an appropriate model for the data. Echelle spectrographs image the entrance slit as a function of wavelength to the focal plane of the instrument. A monochromatic input beam would result in an image that reproduces the illumination conditions of the entrance slit. The image observed on a screen in the focal plane is what is called the instrumental profile (IP). The IP is similar to the concept of the point spread function (PSF) in imaging instruments. The PSF describes how a point source is broadened in the spatial domain by the optical system. The IP describes how a monochromatic beam is broadened in wavelength due to the finite slit width. Both the PSF and the IP have a two dimensional shape but in spectroscopy all photons perpendicular to the dispersion direction can be summed up so that in the extracted spectrum the IP is reduced to a one dimensional shape which is known as line spread function (LSF).

HARPS is fed by an octagonal fiber resulting in a Gaussian LSF. In other spectrographs, like for example CARMENES, the octagonal fiber is sliced into two halfmoon shapes stacked on top of each other. This configuration results in a Supergaussian<sup>1</sup> LSF. Figure 2.6 indicates how spectral lines are seen by HARPS and CARMENES. The top panels show the input into the spectrograph system. In case of the HCL or a LFC (panel 1a and 1b), the calibration lines are not resolved by the spectrograph. HARPS and CARMENES both use FPIs with a low finesse between 4 and 8 (Wildi et al., 2010; Schäfer & Reiners, 2012). Spectral lines of an FPI with such low finesse are typically resolved by high-resolution spectrographs (see panel 1c and 1d for an FPI with a finesse of 8). The second row in Fig. 2.6 shows how HARPS and CARMENES see the lines of the top panel if a screen is placed in the focal plane of the spectrograph. What is observed in the focal plane can be mathematically described as a convolution of the input line with the LSF of the spectrograph. HCLs or LFCs produce input lines that can be described as  $\delta$  functions. What is observed in the focal plane is thus simply the LSF of the spectrograph. For FPIs, however, the observed lines do not represent the LSF of the spectrograph because the broader FPI lines are convolved with LSF of the spectrograph.

For sharp, unresolved input lines, analytical models are often sufficient to represent the LSF and determine accurate line positions. The broader, resolved spectral lines of FPIs, however, are more difficult to model with analytical functions. Further, only ideal spectrographs producing ideal LSFs were considered until now. Misalignment of either optical components or the fiber (slices) at the entrance slit can distort a spectrographs’ LSF or produce asymmetries that vary over the focal plane. In practice, the LSF is unknown, making accurate position measurements prone to errors when analytic and symmetric models are used (Verschueren, 1995).

<sup>1</sup>Supergaussians are Gaussian functions with exponents larger or smaller than 2.

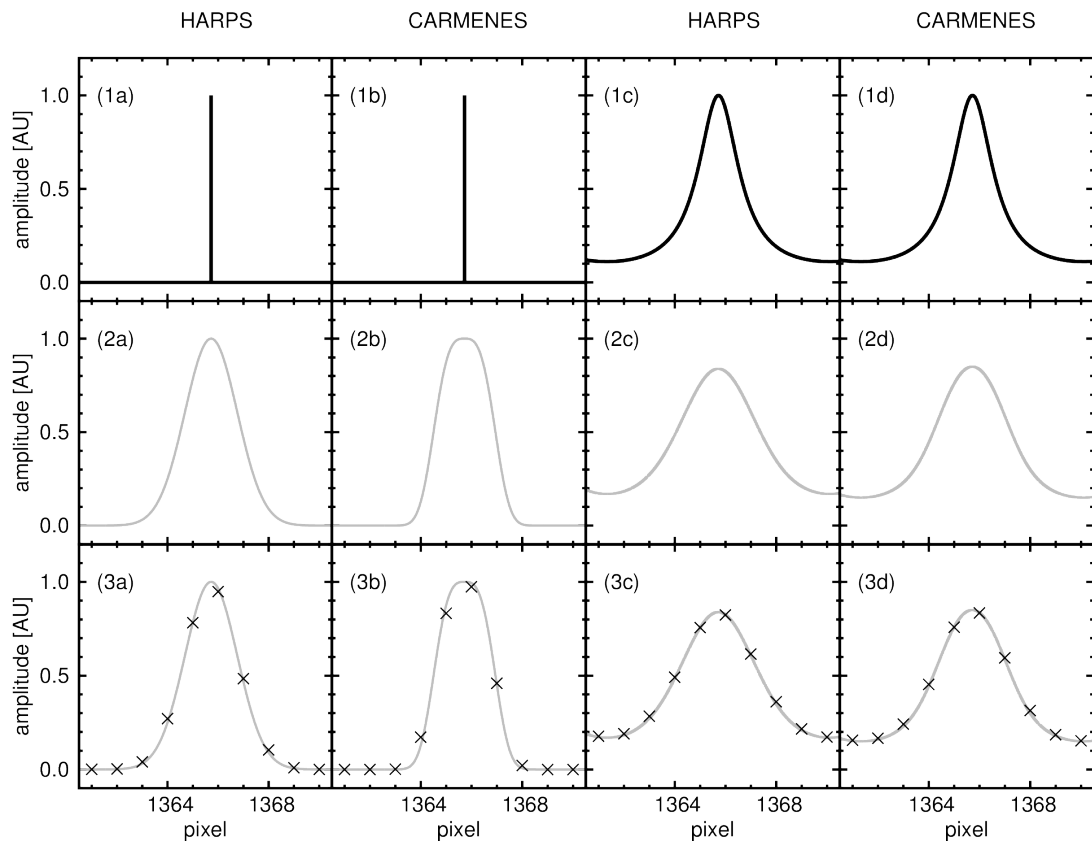


FIGURE 2.6: Schematic of how echelle spectrographs image calibration lines. Top row: input lines. Left panels (1a) and (1b) show a sharp, almost monochromatic line as produced by a HCL or a LFC. Right panels (1c) and (1d) show an FPI line with a Finess of 8. Middle row: Input lines convolved with the spectrograph LSF (Gaussian for HARPS and Supergaussian for CARMENES). Bottom panels: LSFs of middle panel (gray lines) convolved with a one pixel wide box function (black crosses) which gives the eLSF.

However, the LSF is not what is observed with astrophysical spectrographs because pixel detectors are used in the focal plane. Instead what is observed is the response of the detector pixels to the LSF. This idea was published in [Anderson & King \(2000\)](#) for high-precision astrometry but here it is applied to spectrographs rather than imaging instruments. Hence the discussion in this section follows closely the work of [Anderson & King \(2000\)](#) and the interested reader is referred to this paper and references therein for more detailed information. In detectors each pixel has a spatial extension. When placed in the focal plane each pixel observes an integration of the LSF over the wavelength range covered by the pixel. Pixels do not have the same sensitivity everywhere and generally the sensitivity of a pixel decreases towards its edge. Therefore the integration of the LSF over the wavelength range covered by the pixel is weighted with this intra-pixel sensitivity function. Since the LSF is never directly observed, the observed profile is denoted here as the effective line spread function (eLSF). The lower row of [Fig. 2.6](#) shows what a pixel detector placed in the focal plain would observe, assuming uniform pixel sensitivity. The black crosses are the observed pixel values and the gray line represents the LSF of the middle row for comparison. The changes from LSF to eLSF are very small and hard to notice. Because of the integration of the LSF over a finite wavelength range, the eLSF is slightly broader than the LSF. Even such small changes in the percent range can have impact on line position measurements if the aim is m/s calibration or better. Using proper models for the eLSF is challenging since it requires knowledge about optical distortions and intra-pixel sensitivity.

### 2.2.3 Reconstructing the eLSF

Because the eLSF can be a complex function it would be desirable to find a good approach to model the observed line shapes. It turns out that the eLSF can be reconstructed combining the information of all calibration lines when the LSF and eLSF are interpreted in a different way. The LSF describes the distribution of photons of monochromatic wavelength in the focal plane. It can be seen as an intensity profile that describes how much light is observed at a distance  $r$  from the line center. The same is true for the eLSF. It describes how many photons are collected by a pixel centered at a distance  $r$  from the line center. The line center can be located anywhere with respect to the pixel grid. Since the repetition rate of pixels and calibration lines are typically not the same, line centers are randomly distributed within pixels. This means that the phase of the lines are random and each line samples the eLSF at different points. Therefore the eLSF is also known between the discrete pixel values and the random pixel sampling of each line can be used to reconstruct the eLSF.

In an ideal case instruments have a constant LSF across the focal plain and a detector with constant pixel sizes and sensitivities. In this case the eLSF is constant everywhere on the detector. If all line positions are exactly known, all lines can be normalized and remapped to a common zero point. This results in a representation of the eLSF that is sampled also in between the discrete pixel nodes and that allows to construct a very accurate model of the eLSF using smooth functions.

In practice, no instrument is ideal, and the LSF varies across the focal plain. Gratings have the property to magnify or demagnify the slitwidth as a function of the incident and diffraction angle. This effect, called anamorphic magnification, stretches the slitwidth towards the red end of a spectral order, which increases the sampling of spectral lines and their apparent FWHM. For this reason also the eLSF changes from spectral line to spectral line instead of being constant. A second problem in the reconstruction method is that the eLSF profile and the line positions are interlinked. In order

to determine accurate line positions, an accurate eLSF model is required. However, to construct this accurate eLSF model, accurate line positions are needed. To resolve this problem, an iterative process can be used.

In practice the eLSF reconstruction process starts by finding a first guess for the line positions. This can be done by either fitting a roughly estimated eLSF (e.g. a Gaussian) to each line or by simply calculating the flux-weighted center (centroid) for each calibration line. The next step is to recenter all spectral lines. An aperture (typically 8 pixels wide) is chosen around each line. The first guess line center is then subtracted from all pixel values within this aperture, which recenters all spectral lines to a common zero point. Figure 2.7 shows the process with a simulation of one spectral order of the CARMENES FPI. In the upper panel all FPI peaks are remapped to a common zero point by subtracting the line centroid from an 8 pixel wide aperture around each line. The FWHM and the amplitudes of all FPI lines are changing across the order. In a real spectrograph, a similar behavior can be observed caused by the blaze function and anamorphic magnification. Because the eLSF is not constant across the order, the line number (from blue to red) is kept as a third axis which allows to model the eLSF changes. Assuming that all changes in the eLSF are small and smooth, a 2D b-spline (de Boor, 2001; Dierckx, 1995) is fitted to the surface. Thereby a local eLSF model is constructed for each FPI line, which is plotted in the middle panel of Fig. 2.7. This b-spline eLSF model can now be used to determine new line positions, which in turn can be used to create a new eLSF model. In principle, every iteration of the process produces a better eLSF model and in turn, better line positions. The procedure is stopped when the rms of the difference in line positions between two iterations is less than 30 cm/s.

Determining line positions with eLSF modeling requires a center definition for the line profile. The advantage of eLSF profiles is that the models are not necessarily symmetric and can capture distorted line profiles that are hard to reproduce with other models. However, in asymmetric line profiles, it is difficult to define the line center. For the eLSF procedure, the center definition of Anderson & King (2000) is adopted so that the line center is defined at the half distance between the points that define the FWHM.

#### 2.2.4 How accurate can we get? Comparing the performance of eLSF models

To test the method of eLSF reconstruction, four simulations were created. The spectra mimic one order of an LFC for HARPS and CARMENES, and one order of an FPI for HARPS and CARMENES. To make the simulations comparable, all spectra contain 425 emission lines at the same detector positions. The LFCs lines are very narrow and not resolved by both spectrographs. The FPIs have a finesse of 8 for both spectrographs and their lines are a convolution of the FPI lineshape with the LSF of the instruments. The LSF of HARPS is simulated as a Gaussian, and for CARMENES, a Supergaussian with an exponent of  $e = 3.15$  is used. To simulate anamorphic magnification, an increase of the LSF FWHM from 1.8 pixels at the blue order edge to 3.7 pixels at the red order edge is implemented. The detectors of both instruments are assumed to have a uniform pixel sensitivity. Photon noise is simulated, but any noise from the detector and its electronics itself is neglected. In total, eight spectra per case are simulated with a peak S/N between 50 and 400 in steps of 50. Examples of the line profiles simulated can be seen in the lower panels of Fig. 2.6.

Now the eLSF reconstruction method can be compared to fitting analytic model eLSFs. Although the eLSF is slightly different than the LSF, a Gaussian is fitted to the lines of HARPS, and a Supergaussian with the same exponent that was used for the

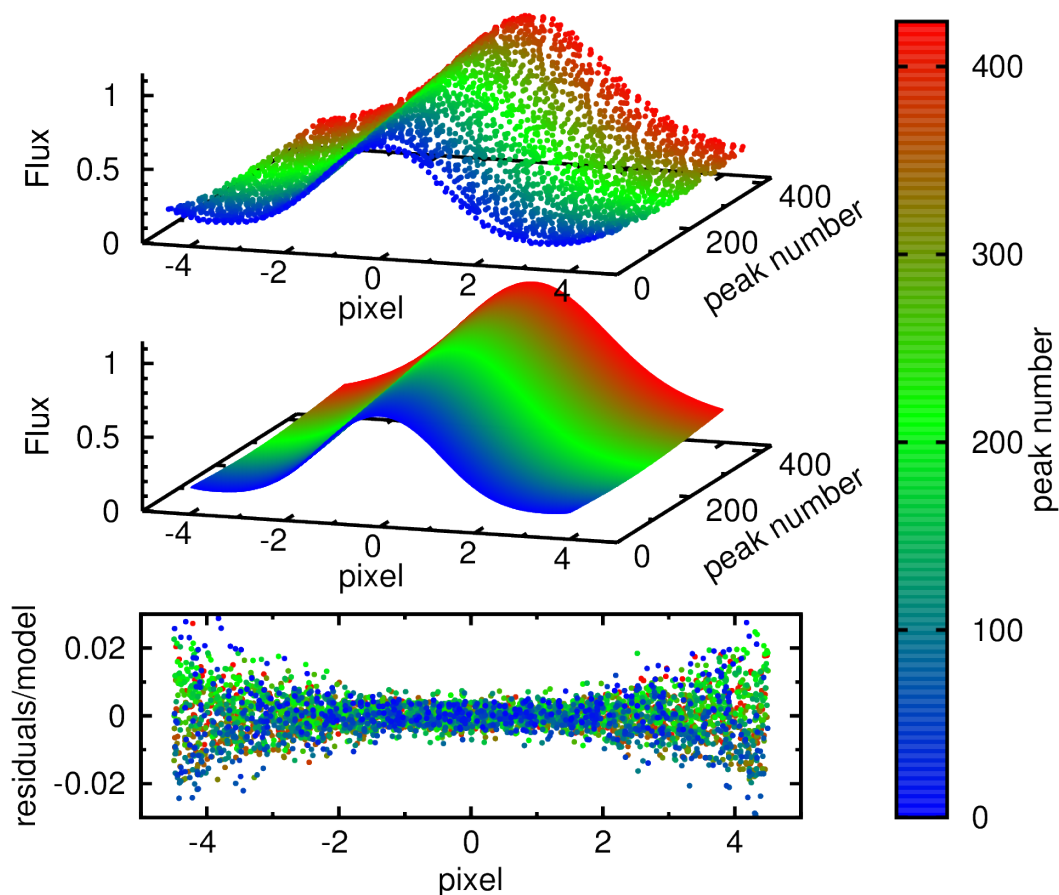


FIGURE 2.7: Example of the eLSF reconstruction procedure with a simulation of the CARMENES FPI. Top panel: Each FPI peak is remapped to a common zero point using the line centroid. The line flux is plotted as a function of remapped pixel location and increasing peak number from blue to red. Middle panel: The eLSF model is a 2D b-spline fit to the data. Lower panel: Normalized residuals of the fit



simulation,  $e = 3.15$ , is used for CARMENES. For the FPI, a Lorentzian function is used in both cases to fit the transmission lines because Taylor expansion of the FPI transmission equation and truncation after the quadratic term leads to a Lorentzian:

$$T = \frac{1}{1 + F \sin^2(u)} \sim \frac{1}{1 + F(u - u_0)^2}, \quad (2.3)$$

where  $u = 2d\pi/\lambda$  with  $d$  being the effective cavity width and  $F$  the finesse of the FPI.

All models have four free parameters: position, amplitude, width and offset from zero. Figure 2.8 shows the rms scatter between the fitted and true position of all lines as a function of the signal to noise ratio in the spectra.

From Fig. 2.8, three important conclusions can be drawn:

- In all simulated cases, the eLSF reconstruction method performs equal or better than the analytic model. Nevertheless, the benefits of using eLSF reconstruction rather than analytic models depends on the spectra under investigation. For a HARPS-like LFC and the two FPI spectra, the eLSF reconstruction does not improve the precision substantially. However, an LFC for CARMENES would be extremely sensitive to the profile function used. The Supergaussian model eLSF is only slightly different from the exact eLSF in the CARMENES simulation. Although it is the exact representation of the LSF for this simulation, using it to fit the eLSF data limits the precision to about 13 m/s.
- The rms scatter of the measured line positions around the true line position is inversely proportional to the S/N value. With increasing S/N of the spectrum the line positions are found more precisely. However, very high S/N values do not lead to substantially higher precision, so that in practice, line positions can be determined to m/s precision even without excessive S/N. Although this trend is seen in all simulations shown in Fig. 2.8, an appropriate eLSF model is still needed to overcome limitations given by eLSF mismatch (see LFC simulation of CARMENES in Fig. 2.8).
- Different eLSFs have different information content. If the precision of the line fitting is compared between simulations of the same S/N, it is found that in the case of the HARPS LFC the fitted line positions match the true positions best. This is related to the RV information contained in one line of a specific shape. However, the differences are very small between simulations and line position fitting is not the limiting factor in the calibration process as long as the S/N is high enough.

The rms is a way of testing the overall precision achieved with different eLSF models. Sampling and pixel phase of lines can influence the accuracy with which the positions can be determined. A detailed investigation is needed to reveal systematics driven by line parameters in combination with model mismatch. From astrometry, it is known that inaccurate PSF models can lead to phase errors in the fitted stellar positions (Anderson & King, 2000). As the true line positions in the simulations are known, the pixel phase,  $p$ , for all lines can be calculated by subtracting the pixel number from the line position;  $p = x - INT(x)$ . In the notation used here, a phase of 0 or 1 means that the line center falls in between two pixels, whereas in the case of phase 0.5, the line center falls directly onto the center of the pixel. Figure 2.9 shows the difference between the measured and true line position as a function of the line phase in the simulation for the CARMENES FPI. In the top panel the Supergaussian eLSF model is used. There is a

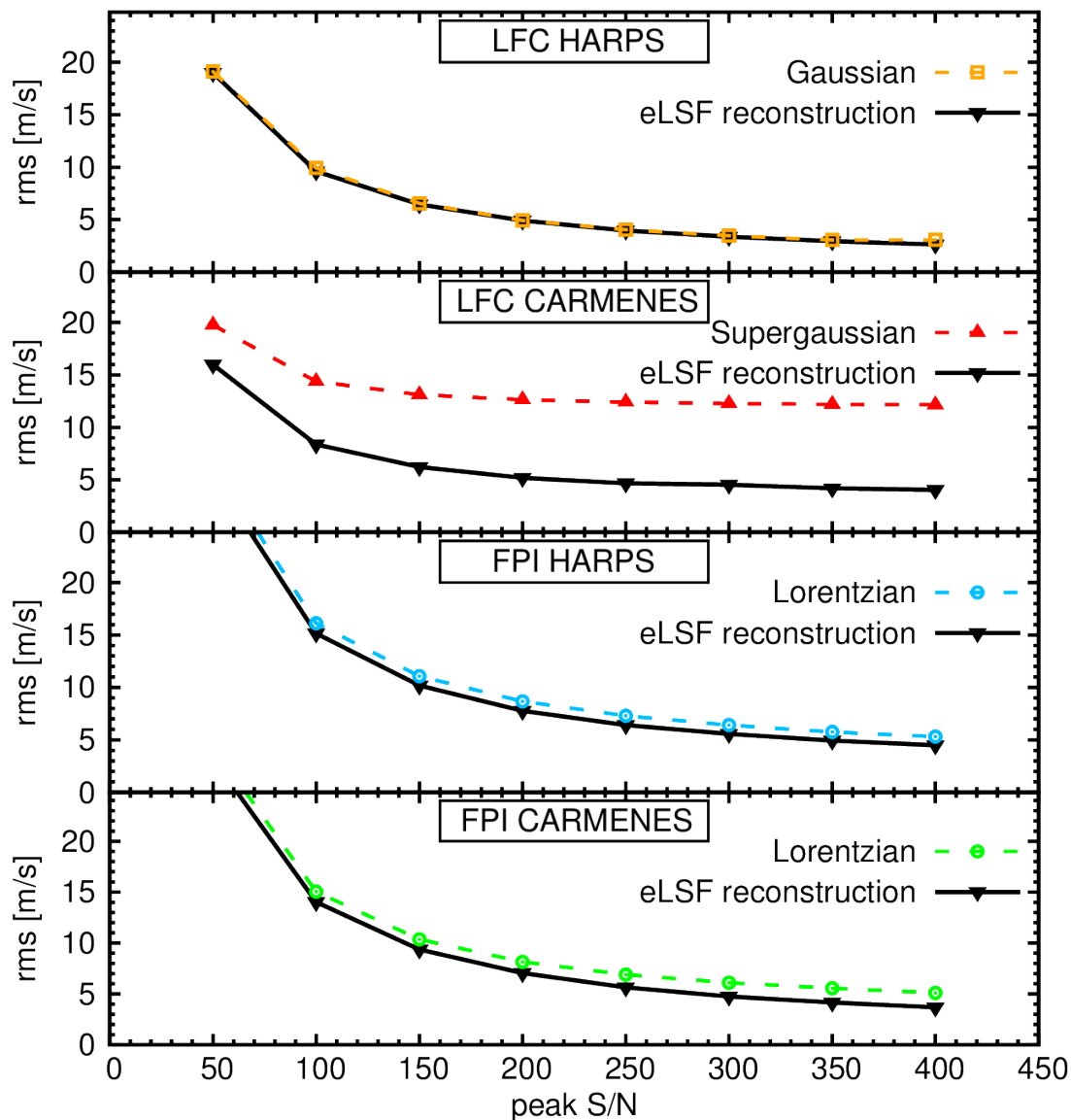


FIGURE 2.8: rms scatter between fitted and true position as a function of S/N value for different spectra and eLSF models. Solid black lines in all panels represent the eLSF reconstruction method. Broken colored lines represent the fit with analytic models. The upper panel shows the LFC simulation of HARPS. In the second panel the LFC simulation for CARMENES is plotted. The third panel visualizes the FPI simulation in HARPS and the lower panel shows the FPI simulation for CARMENES.



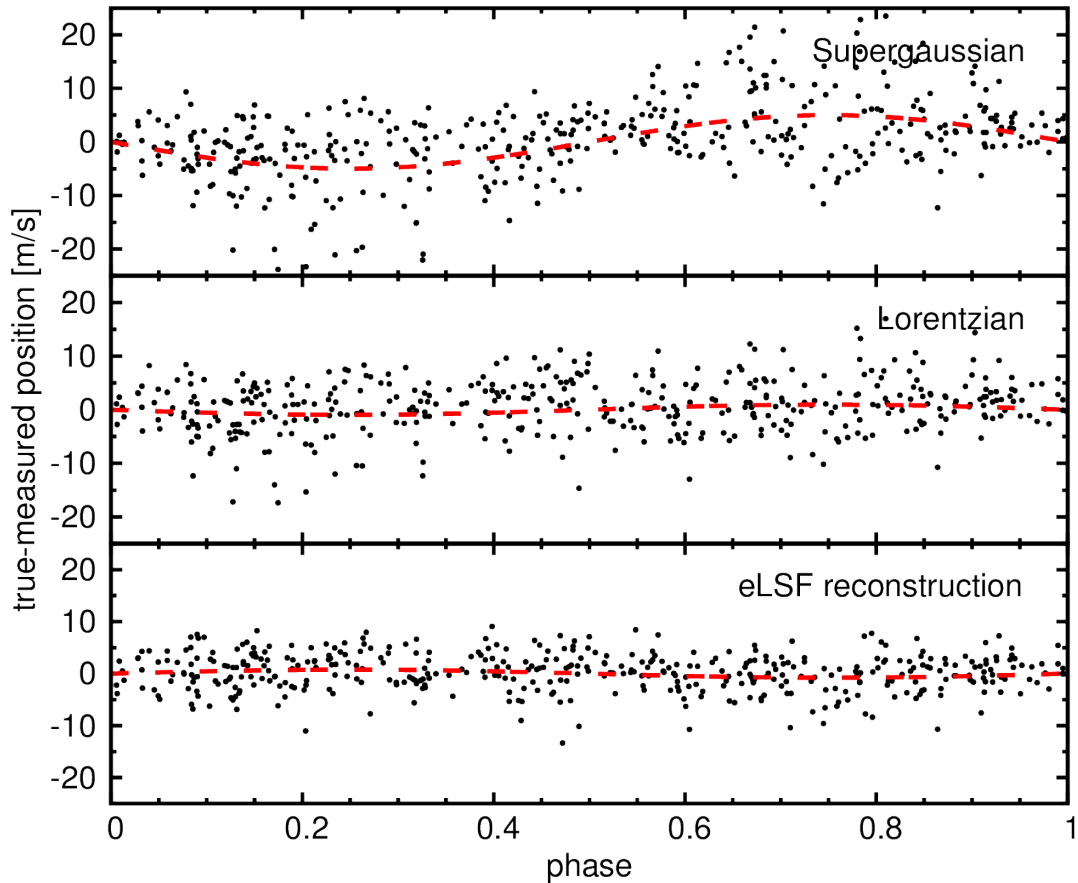


FIGURE 2.9: Pixel phase error in the CARMENES FPI lines for three different eLSF models. Black points represent measurements of single lines. Red lines are sinusoidal fits to the line position deviations. Upper panel: Line position error of a Supergaussian fit. Middle panel: Line position error of a Lorentzian fit. Lower panel: Line position error of the eLSF reconstruction method.

clear correlation visible between pixel phase and position error. A sinusoidal fit to the data reveals an amplitude of about 7 m/s. This clearly shows the necessity for good eLSF models. When a better model for the eLSF (in this case Lorentzian or eLSF reconstruction) is used, the systematics with pixel phase vanish and no significant amplitude is visible.

Another test that can be performed is to check the difference between measured and true line position as a function of FWHM of the spectral lines. Anamorphic magnification stretches lines towards the red end of a spectral order so that the sampling per resolution element is different for different lines across the detector. Figure 2.10 shows how different eLSF models perform for the changing FWHM in the CARMENES FPI simulation. For the Supergaussian and the Lorentzian models, seen in the upper and middle panel of Fig. 2.10 respectively, the precision increases with increasing FWHM. Towards the blue end of the spectral order, the scatter around the true line position increases, which in turn increases the uncertainty of a wavelength solution in the blue edge of the detector. If the eLSF reconstruction method is used, no trend for larger scatter in the blue is observed. Hence, in spectrographs close to undersampling, wavelength calibration can benefit from the eLSF reconstruction method.

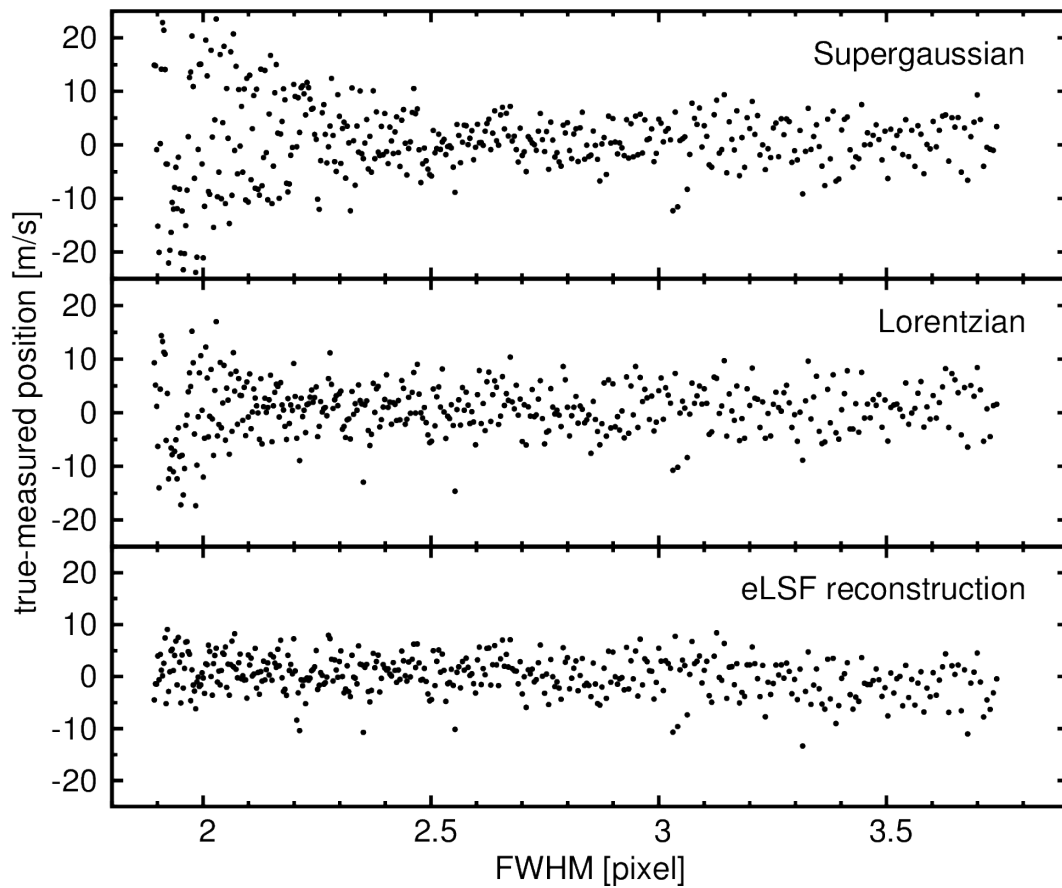


FIGURE 2.10: Position error in the CARMENES FPI lines for three different eLSF models as a function of FWHM of the calibration line. Top panel: Supergaussian model; middle panel: Lorentzian model; lower panel: eLSF reconstruction.

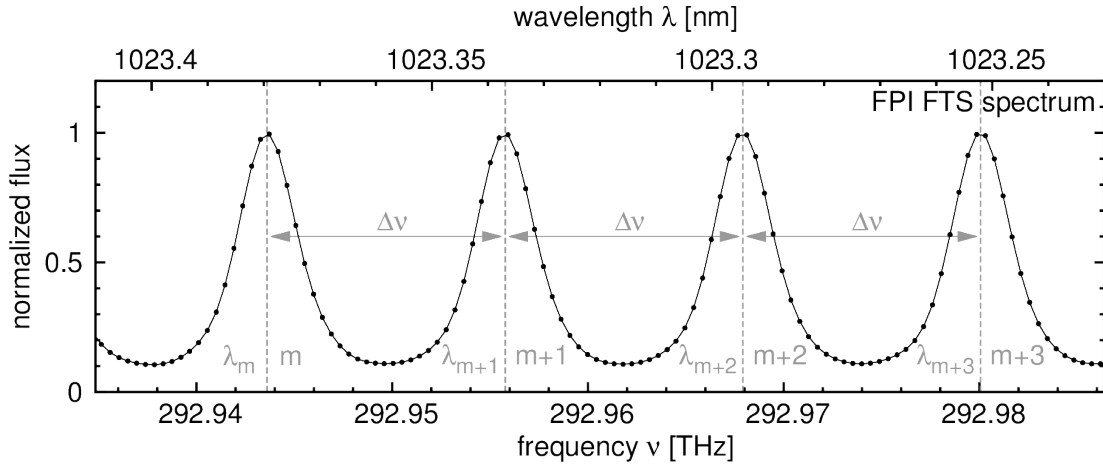


FIGURE 2.11: Spectrum of the CARMENES NIR FPI taken with the Fourier transform spectrometer in Göttingen.

## 2.3 Using Fabry-Pérot etalons for wavelength calibration

### 2.3.1 Overview

When measuring precise radial velocities, the wavelength solution is essential as it provides the physical scale for the spectra. For a long time hollow cathode lamps (HCLs) were used to calibrate astronomical spectrographs but their drawbacks (see Sect. 2.1) prevent further improvements in RV precision. Fabry-Pérot etalons are promising candidates for next generation calibrators as they offer a dense grid of lines. In this section we present a method to use Fabry-Pérot interferometers (FPIs) for wavelength calibration of astronomical echelle spectrographs. Although FPIs are no absolute standards we propose a method to anchor the numerous FPI lines with HCLs. We explore the feasibility of our method with the HARPS spectrograph and show that we are able to reproduce results obtained with the experimental LFC installed at the 3.6m telescope in La Silla. Further we show a simulation to estimate the consequences of a distorted wavelength solution on precision RV measurements. The following section was published in [Bauer et al. \(2015\)](#) in the *Astronomy and Astrophysics* journal as ‘Calibrating echelle spectrographs with Fabry-Pérot etalons’. The permission for republication in this thesis was granted by A&A.

### 2.3.2 Calibrating the Fabry-Pérot interferometer

#### Ideal FPI

The ideal FPI consists of two partly reflective plane-parallel surfaces ([Born & Wolf, 1999](#)). When light enters the device, it is reflected between the two surfaces and interference takes place between light rays, depending on the phase difference. With a white-light input source, a spectrograph will observe a pattern of transmission peaks that are spaced equidistantly in frequency  $\nu$  (Fig. 2.11). The transmission maxima fulfill the interference condition given by the effective FPI cavity width  $d$ :

$$m\lambda_m = 2d, \quad (2.4)$$

where  $m$  is an integer number called interference order, and  $\lambda_m$  are the transmission maxima wavelengths. The effective cavity width is  $d = D n \cos \theta$ , with the distance between the two mirrors  $D$ , the refraction index  $n$ , and the incident ray angle  $\theta$ .

Using the FPI as a calibration source requires knowledge of the exact wavelengths for all interference maxima observed by the spectrograph. As the wavelength  $\lambda_m$  of each FPI interference order depends only on the effective cavity width,  $d$ , this parameter must be known with the same accuracy as the wavelength requirement ( $\delta d/d = 3 \cdot 10^{-9}$  for 1 m/s). For the FPI this means that the spacing  $d$  must be known with an accuracy of about 0.15 (Reiners et al., 2014). Typically, the mirror distance  $D$  is only known to about 1  $\mu\text{m}$ , which is five orders of magnitudes above our requirements. In addition, we do not exactly know the incident ray angle  $\theta$ . Hence, we need to find a method for calibrating the effective FPI cavity width  $d$ . The interference condition of the FPI, Eq. (2.4), can be used to perform the cavity width measurement. This can be done with any high-resolution spectrograph and its internal calibration sources such as HCLs (Wildi et al., 2010). First, the wavelength solution of the spectrograph must be derived using only HCL lines. Then the line positions of the FPI can be measured on the detector and the HCL wavelength solution can be used to assign the corresponding wavelength to all interference peaks.

Now we need to identify the interference order  $m$  for all FPI peaks. From the observed spectra we can only obtain a relative numbering  $k = m - m_1$  by simply counting the lines, but the absolute interference order of the first (reddest) peak  $m_1$  is not known. To find  $m_1$ , the fact that the ideal FPI cavity width is a constant number can be used. A guess for  $m_1$  and the FPI wavelengths derived from the HCL wavelength solution can be used to obtain the effective cavity width  $d$  for each FPI line separately via Eq. (2.4). If the effective cavity width  $d$  is plotted as a function of relative peak number  $k$  using the correct  $m_1$ , the data points are expected to be distributed around a constant value. If the guess for  $m_1$  is incorrect, the data points will, however, produce a positive slope for  $m_1$  too small or a negative slope for  $m_1$  too high. Once the right value for  $m_1$  is found, the large number of individual effective cavity width measurements (typically on the order of  $10^4$  for high-resolution echelle spectrographs) can be averaged to obtain a good estimate of  $d$ .

We now only need this one parameter (the effective cavity width  $d$  of the FPI) to globally predict the wavelength of all interference peaks observed. The combined accuracy of all HCL lines anchors the FPI on an absolute wavelength scale. Paired with the high precision of the FPI, local imperfections of the wavelength solution can now be resolved and corrected. The numerous FPI features bridge regions lacking any spectral features of HCLs, and we can now determine whether the wavelength solution model is sufficient or if more detailed models are required.

### Real FPI

In practice, applying the concept of calibrating the effective cavity width using standard calibrators proves to be more difficult for real FPIs. The cavity width is not constant throughout the wavelength range covered by high-resolution echelle spectrographs. FPIs are usually soft coated, allowing photons of different energy to penetrate to different depths of the dielectric surface. Hence, photons of different wavelengths see different cavity widths. As an example, the penetration depth variation over the wavelength range of HARPS easily reaches a few hundred nm (Wildi et al., 2010). If we do not account for this, the computed FPI wavelengths are incorrect by several km/s. For a full characterization of the FPI spectrum, the task is therefore not only

to find the global mean effective cavity width  $\bar{d}$  of the FPI, but also to determine the wavelength-dependent function  $d(\lambda)$ .

We assumed that the penetration depth is a smooth function of wavelength and therefore also a smooth function of relative interference order  $k$ . Because the measured wavelength is uncertain but the relative peak numbering is noise-free, we modeled  $d$  as a function of  $k$  instead of  $d$  as a function of  $\lambda$ . We chose uniform cubic B-splines (de Boor, 2001; Dierckx, 1995) to fit the function  $d(k)$ .

The basis of our method is the assumption that the penetration depth variation is smooth. Hence we decided to test this with the CARMENES (Quirrenbach et al., 2011) near-infrared FPI (finesse  $\sim 8$ , Schäfer & Reiniers, 2012). We obtained 78 spectra with the Bruker IFS 125/HR Fourier transform spectrometer (FTS) with a resolution of  $R = 500\,000$  in our optics laboratory. The median signal-to-noise ratio of the coadded spectrum is 350 (Fig. 2.11). We used the internal wavelength calibration of the FTS as starting guess for all FPI peak wavelengths. This calibration is provided by a He-laser measuring the path length within the spectrometer. With the high-resolution of the FTS, the FPI peaks are well resolved in the spectra, and we measured individual line positions by fitting a Lorentz function. The FPI itself was not operated under vacuum conditions, thus water lines contaminated the spectrum. We excluded contaminated peaks from our cavity width measurement. We determined the mean effective cavity width  $\bar{d} = 12.338\,350 \pm 0.000\,040$  mm, which is consistent with what we expect from the mirror distance given by the manufacturer  $D = 12.334$  mm, the refractive index of air  $n \sim 1.003$ , and an incident ray angle of  $\theta = 0$  deg.

The penetration depth variation of the FPI as a function of the relative peak numbering and the measured FTS wavelength is shown in Fig. 2.12. Fitting a B-spline with 50 nodes results in an RMS scatter around the model of 0.326 nm, translating into 8 m/s. We found no discontinuities in  $d(k)$  within our measurement precision. Hence our assumption of a smooth penetration depth variation is valid. The RMS scatter of 8 m/s is about what we expect from photon noise.

After we determined the model parameters for the effective cavity width  $d$ , we searched for the absolute interference number of the first observed FPI peak  $m_1$  to compute the wavelength for all interference orders. This is no longer trivial if the effective cavity width  $d$  is not a constant.

Reiniers et al. (2014) suggested using the mean free spectral range  $\langle \Delta\nu \rangle$  to calculate the mean effective cavity width and then use Eq. (2.4) to calculate the absolute order numbering. Wildi et al. (2010) determined the absolute order numbering assuming the effective cavity width  $d$  to be exactly the absolute cavity width  $D$  specified by the manufacturer and rounding  $m$  to the closest integer number they obtained from equation Eq. (2.4) and the HARPS wavelength solution. Their approach resulted in a penetration depth curve with the least variation.

Other choices of  $m_1$  introduce steeper slopes in the effective cavity width  $d$ . We show that the choice of  $m_1$  influences the overall slope of the function  $d(k)$  for the CARMENES NIR FPI in Fig. 2.13. Because we do not know enough about the material properties of the coating, the penetration depth variation with wavelength and the alignment of the FPI, we do not know which  $m_1$  and which  $d$  curve are the true ones. We chose  $m_1 = 13604$  as reference value because it results in the  $d(k)$  curve with the least penetration depth variation. We note that this choice is not necessarily the true value of  $m_1$ .

We realized, however, that it is not necessary to know  $m_1$  exactly to calibrate the FPI wavelengths. As a result of the degeneracy in Eq. (2.4) between the absolute peak number  $m$  and the effective cavity width  $d(k)$ , identical peak wavelengths occur for

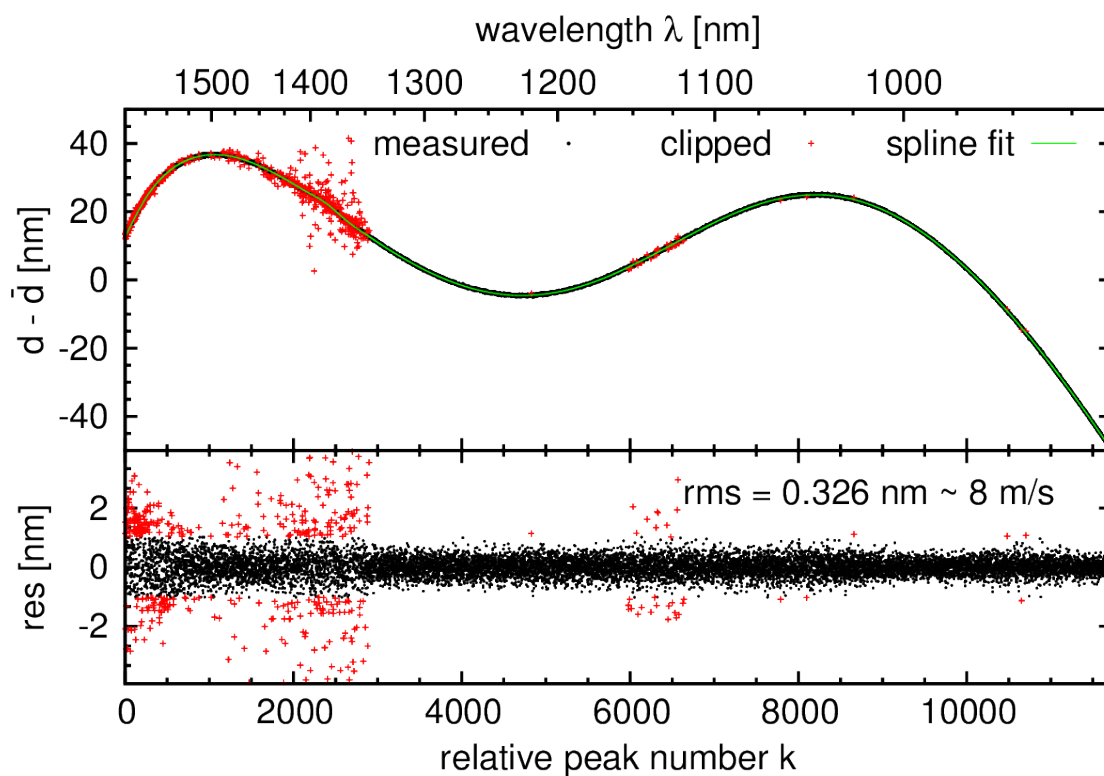


FIGURE 2.12: Upper panel: Effective cavity width measurement  $d(k)$  for the CARMENES NIR FPI. Lower panel: Residuals of the spline fit. We show the data (black dots), clipped lines due to atmosphere absorption features (red crosses), and the spline fit to the data (green line).

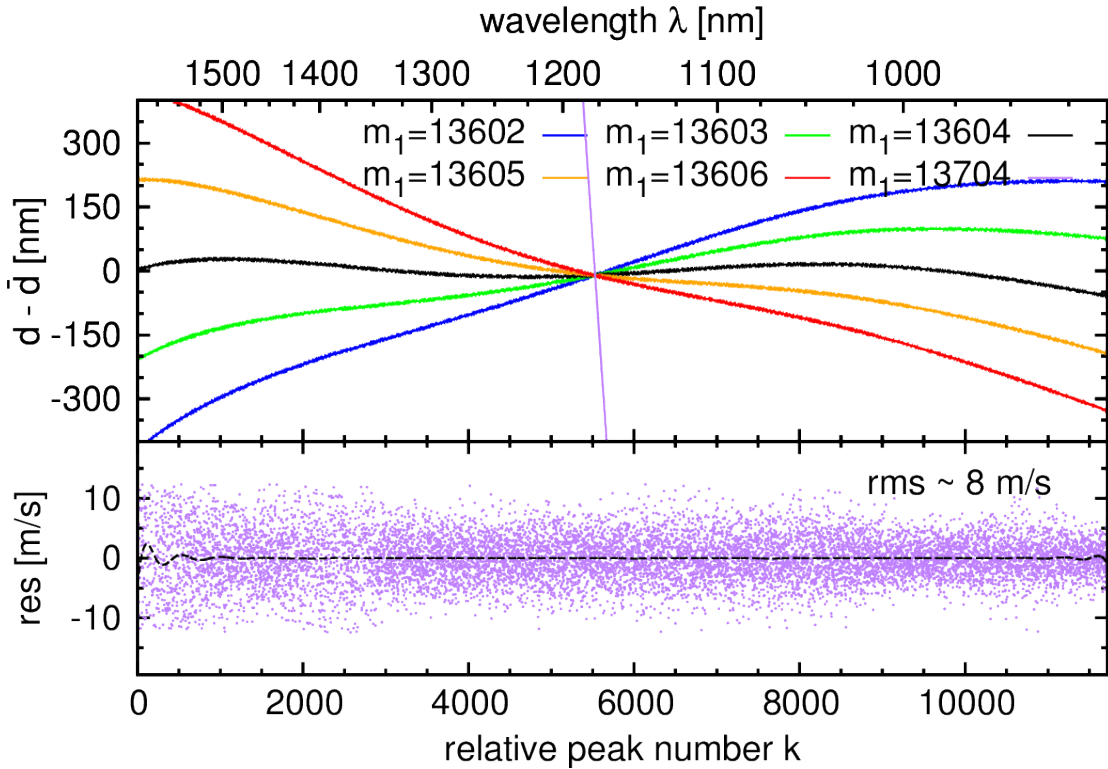


FIGURE 2.13: Upper panel: Effective cavity width  $d(k)$  obtained for different absolute interference order numbers  $m_1$ . The reference curve ( $m_1 = 13604$ , black) is shown in Fig. 2.12. Lower panel: Difference in modeled FPI peak wavelength  $\lambda_m$  using  $m_1 = 13704$  with the reference curve (dashed line). Residuals of the spline fit to the cavity width  $d(k)$  obtained with  $m_1 = 13704$  are shown as purple dots.

different combinations of  $m$  and  $d(k)$ . Hence, our method for deriving absolute wavelengths for individual FPI peaks is expected to be robust against choosing the wrong value for  $m_1$ .

To test this, we computed  $d(k)$  using two different values for  $m_1$  differing by 100 interference orders,  $m_1 = 13604$  and  $m_1 = 13704$ . For both cases we obtained the FPI wavelengths using Eq. (2.4). We calculated the difference of the calibrated FPI wavelengths line by line and show the result as a dashed line in the lower panel of Fig. 2.13. The largest difference between the calibrated peak wavelengths is 10 cm/s in the region between  $k = 2000$  and  $k = 10000$ . The difference increases up to 5 m/s at the edges because there are fewer data points to constrain the spline fit. Despite the large difference in  $d(k)$  (several hundred nm, see upper panel in Fig. 2.13), the difference between the modeled wavelengths is small (lower panel in Fig. 2.13).

### 2.3.3 Wavelength solution model

In this section, we explain our strategy to determine the wavelength solution of a high-resolution echelle spectrograph. First, we introduce our concept for the mapping between wavelength and pixels in the case of a monolithic detector with uniform pixels. We describe our model, the parametrization, and the regression type. Second, we proceed with real detectors and report how we account for uneven pixel sizes in the wavelength solution.



### Mapping wavelengths to detector coordinates

The general idea of wavelength calibration is to feed the spectrograph with a reference spectrum of known wavelengths  $\lambda_l$ , for example, the known atomic lines of a hollow-cathode lamp, and observe the response at the detector. The positions of the line centers  $x_l$  in the extracted spectrum are measured by fitting a model of the instrumental profile (e.g., Gaussian functions) to each individual spectral line  $l$ . For the set of wavelength-pixels pairs  $(\lambda_l, x_l)$ , a relation is established by a fit. Finally, this relation is used to predict the wavelength at the center of every detector pixel.

The wavelengths of standard lines are given by accurate and precise line lists with uncertainties in the range of  $10^{-6} - 10^{-7}$  ( $\sim 10$  m/s) (Redman et al., 2014). On the other hand, the positions of weak lines on the detector exhibit typical measurement errors of 0.1 pixels or about 100 m/s. Hence, in wavelength calibration the standard line wavelength,  $\lambda_l$ , represents the independent variable (or the cause) and the position on the detector,  $x_l$ , is the dependent variable (or the spectrograph response).

A common and simple procedure for wavelength calibration is to fit a continuous model for  $\lambda(x)$ . From the function  $\lambda(x)$  the wavelength for each pixel can be directly evaluated. In this approach the role between dependent and independent variable is reversed (Isobe et al., 1990), and we call this reverse regression. Reverse regression causes problems when weighting the data.

For this reason, direct regression can be used to fit the detector position of individual lines as a function of wavelength,  $x(\lambda)$ . To predict the wavelength for each pixel in this approach, the model  $x(\lambda)$  must be inverted  $\lambda(x) = x(\lambda)^{-1}$ .

The wavelength solution is usually modeled with polynomials ( $o\lambda = \text{poly}(x)$ ) in case of reverse regression, e.g., Baranne et al., 1996), which can be interpreted as a Taylor expansion of the grating equation

$$o\lambda = n\sigma [\cos \gamma_1 \sin \alpha_1 + \cos \gamma_2 \sin \alpha_2] \quad (2.5)$$

in terms of the diffraction angle  $\alpha_2$ , which is proportional to the position in the focal plane,  $x \propto \alpha_2$ . In Eq. (2.5),  $o$  is the diffraction order,  $n$  is the refractive index,  $\sigma$  is the grating constant,  $\alpha_1$  and  $\alpha_2$  are the angles of the incident and diffracted ray with respect to the grating normal, and  $\gamma_1$  and  $\gamma_2$  are the off-plane angles before and after the grating, respectively (Schroeder, 2000).

In direct regression with  $x(\lambda) = \text{poly}(o\lambda)$ , we basically approximate a re-arranged version of the grating equation

$$\alpha_2 = \arcsin \left( \frac{1}{\cos \gamma_2} \left[ \frac{o\lambda}{n\sigma} - \cos \gamma_1 \sin \alpha_1 \right] \right). \quad (2.6)$$

We note that polynomials of higher degree (typically +1) are needed with direct regression for a wavelength solution with similar quality compared to reverse regression. This is because the sine function in Eq. (2.5) is more amenable to a Taylor expansion than the arcsine function in Eq. (2.6).

A way of keeping both advantages (lower number of parameters and direct regression) would be to use  $x = \text{poly}(o\lambda)^{-1}$  as forward model, which we might call inverse regression. This requires nonlinear least-squares fitting and thus higher computational effort.

We chose direct regression below. This allows us to correctly weight all data points and to implement unequal pixel sizes in our model (next section). Moreover, instead



of simple polynomials for each order, we use 2D polynomials to couple the individual orders. This decreases the number of parameter and increases the robustness of the fit. We also use  $o\lambda$  instead of  $\lambda$  as variable, that is,  $x(\lambda, o) = \text{poly}(o\lambda, o)$ . The reparametrization is motivated by Eq. (2.6). This approach is a direct regression version of the algorithm implemented in the IDL REDUCE package of Piskunov & Valenti (2002).

### Unequal pixel sizes

Standard wavelength calibration procedures usually assume that detector pixels are equally spaced and of the same size. Irregularities in pixel size of CCD detectors are on the order of  $10^{-2}$  (Wilken et al., 2010), which is becoming a significant limitation for high-precision spectroscopic measurements. For example, Wilken et al. (2010) showed that inhomogeneous pixel sizes lead to discontinuities in the wavelength solution of HARPS on the order of several 10 m/s.

To account for irregular detector pixels, we included the pixel size in our model and distinguished now more strictly between detector coordinates,  $X$  (pixel), and focal plane coordinates,  $x$ . We denote the size of pixel  $i$  by  $a_{\text{pix},i}$  and assumed that there is no gap space between individual pixels. Thus, the transformation from pixel coordinates to focal plane coordinate is

$$x(X, \mathbf{a}_{\text{pix}}) = \sum_{i=1}^{[X]} a_{\text{pix},i} + a_{\text{pix},[X]} \cdot \left( X - [X] - \frac{1}{2} \right), \quad (2.7)$$

where  $[X]$  is the value of  $X$  rounded to the nearest integer (the pixel number). In Eq. (2.7) we integrate over the width of all preceding pixels (including the current one) and linearly interpolate (backwards) for subpixel position. The zero point of the coordinate system ( $x = 0$ ) coincides with the left border of the first pixel ( $X = 0.5$ ). If all pixels have the same size ( $a_{\text{pix},i} = 1$ ), the transformation is simply  $x = X - 0.5$ . The transformation can also be extended to include virtual pixels for the application to mosaic detectors with gap sizes corresponding to several pixels ( $a_{\text{pix},i} \sim 100$  pixels) between the detectors as, for instance, in CARMENES (Quirrenbach et al., 2011) or CRIRES+ (Follert et al., 2014).

Ideally, the sizes of all pixels are known prior to wavelength calibration, but in practice, the individual pixel sizes are not known precisely enough. For future detectors, it might be interesting to determine pixel sizes with a microscope prior to integration in the instrument.

Determining the sizes of all pixels is beyond the scope of ordinary wavelength calibration because the number of free parameters exceeds the number of comb lines, which are separated by a few pixels. But for a few significantly deviating pixels it is indeed possible to simultaneously determine pixel sizes and the wavelength solution if a spectral calibration source with dense line combs (FPI or the LFC) is available.

For example, the HARPS detector exhibits pixel irregularities every 512 pixels (Wilken et al., 2010), and we therefore chose the following pixel size model

$$a_{\text{pix},i} = \begin{cases} w_i, & \text{if modulo}(i, b \cdot 512 - 3) = 0 \\ 1, & \text{otherwise} \end{cases}, \quad (2.8)$$

where every 512th pixel was allowed to depart from unity, while all others were assumed to be homogeneous and were fixed to unity. The first pixel size irregularity,

$b = 1$ , is located in pixel 509, which causes an offset of three pixels in Eq. (2.8). The number of free parameters can be limited even more by using column sizes instead of individual pixels, as we show in Sect. 2.3.4 for HARPS.

To find the pixel sizes  $w_i$ , we performed a nonlinear least-squares fit by minimizing  $\chi^2 = \sum_l \frac{1}{\sigma_l^2} [X_l - X(x(\lambda_l, o_l, \mathbf{a}_{\text{poly}}, \mathbf{a}_{\text{pix}}))]^2$ . To evaluate the  $\chi^2$  with the observed data  $X_l$ , we must invert Eq. (2.7). The back-transformation of the focal plane coordinates  $x$  into the detector coordinates  $X$  is

$$X(x, \mathbf{a}_{\text{pix}}) = \frac{x - x_j}{a_{\text{pix},j}} + j + \frac{1}{2}, \quad (2.9)$$

where  $j$  is the pixel number into which  $x$  falls, which can be found via the condition  $x_j \leq x < x_{j+1}$  where  $x_j = x(j - 0.5, \mathbf{a}_{\text{pix}}) = \sum_{i=1}^{j-1} a_{\text{pix},i}$ .

The best-fitting model and coordinate transformation was then inverted numerically to obtain the wavelength for the center of each pixel. We found that a simple bisection method is appropriate<sup>2</sup>.

To accelerate the nonlinear least-squares fit, we separated it into a linear and a nonlinear least-squares fit. For a trial set of the nonlinear parameters  $\mathbf{a}_{\text{pix}}$ , the position of all calibration lines can be transformed into focal plane coordinates  $x_l(X_l)$  with Eq. (2.7). Then a smooth 2D polynomial was fitted with a linear least-squares fit. The smooth model was transformed back into detector coordinates with Eq. (2.9), and the  $\chi^2$  with the observed data  $X_l$  was computed. Testing different trial sets (e.g., with a downhill simplex; Press (2007)), the best-fitting pixel sizes  $\mathbf{a}_{\text{pix}}$  were found.

### 2.3.4 Calibrating HARPS using its FPI

#### Detector characterization and wavelength solution

To show the potential of Fabry-Pérot interferometers in wavelength calibration procedures for echelle spectrographs, we calibrated the HARPS spectrograph located at the ESO 3.6 m telescope in La Silla Chile (Mayor et al., 2003) with its etalon and our wavelength solution method. HARPS has a resolution of 115 000 (one resolution element is sampled with about 3.2 pixels), covers the wavelength range from 380 to 690 nm, and offers ThAr, FPI, and LFC spectra. The design of the HARPS FPI is described in Wildi et al. (2010). Briefly, the HARPS FPI has a cavity width of about 7.3 mm and a measured finesse of about 4.3. It is fiber coupled and temperature and pressure stabilized inside a vacuum tank.

We analyzed calibration data taken on March 16, 2011, offering both ThAr and FPI exposures. We reduced the calibration spectra using the IDL REDUCE package. The pixel position of the thorium lines were measured by fitting a Gaussian profile to the extracted 1D spectra. The FPI peaks are separated by 8 – 21 pixels (from blue to red orders). The median FWHM of the FPI lines is about 5 pixels or 1.6 resolution elements, which means that the FPI lines are moderately resolved in the spectra. Hence the line shape of the FPI deviates from the Gaussian-like instrumental profile of HARPS, and we fit a Lorentz function to each peak to obtain the pixel position.

The first step in our calibration procedure is to derive the wavelength solution from the thorium lamp alone using the line list of Palmer & Engleman (1983). Because only few thorium lines are available in single-calibration images, we used a 2D polynomial

<sup>2</sup>Methods involving first derivatives are not recommended by us because of convergence problems with discontinuous functions, as is the case here.

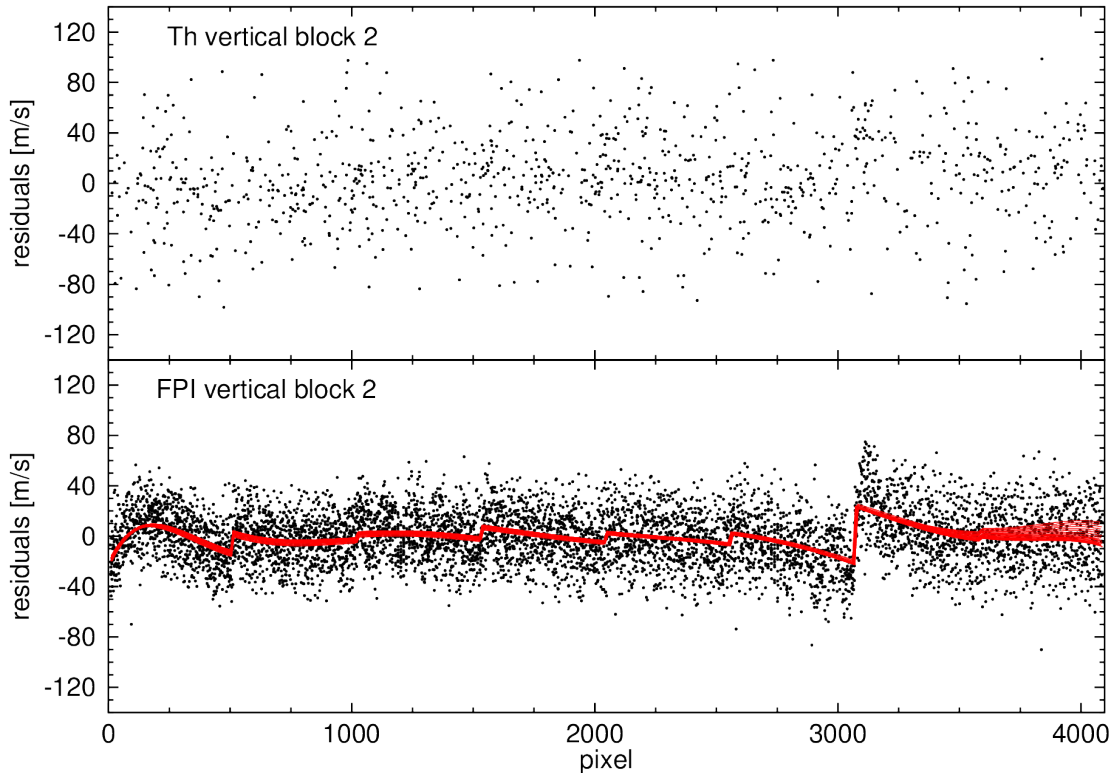


FIGURE 2.14: Residuals of Th and FPI positions with respect to a simple polynomial wavelength solution for HARPS orders 116 – 135 (black dots). The difference in the wavelength solutions between the polynomial model and the model with variable pixel sizes (red lines) shows the captured systematics.

in the direct regression (seventh degree in dispersion direction and sixth degree in cross dispersion) and uniform pixel sizes. The residuals of the best fit to the thorium lines in block 2 of the blue detector (echelle orders 116 to 135) are plotted in the upper panel of Fig. 2.14.

This HCL wavelength solution was used as input for the FPI calibration to assign wavelengths to all FPI peaks. We chose  $m_1 = 27526$  ( $\lambda_1 = 5304.24$ ), which resulted in the minimum penetration depth variation. The model for  $d(k)$  was a B-spline with 23 nodes in the blue and 13 nodes in the red HARPS regime. The effective cavity width measurement  $d_k = 2 \lambda_k (m_1 + k)$  and the according model  $d(k)$  are shown in Fig. 2.15.

The mean effective cavity width we measured for the HARPS FPI is  $\bar{d} = 7.30020 \pm 0.00015$  mm. (The spike seen around 430 nm is caused by a CCD defect in the blue detector that systematically influences the line position measurement in this region.)

From the B-spline model and the peak number we derived the calibrated FPI wavelengths. To check the quality of the thorium wavelength solution, we plot the residuals of the calibrated FPI lines included in block 2 of the blue detector (echelle orders 116 to 135) in the lower panel of Fig. 2.14. The numerous more lines offered by the FPI now enable us to see systematics that were hidden in the Th residuals simply because of the lack of lines.

These systematics are already known from laser frequency comb calibration and are due to a stitching effect (Wilken et al., 2010). The  $4k \times 4k$  HARPS detector consists of two  $4k \times 2k$  CCDs; each consists of  $8 \times 2$  blocks with  $512 \times 1024$  pixels.

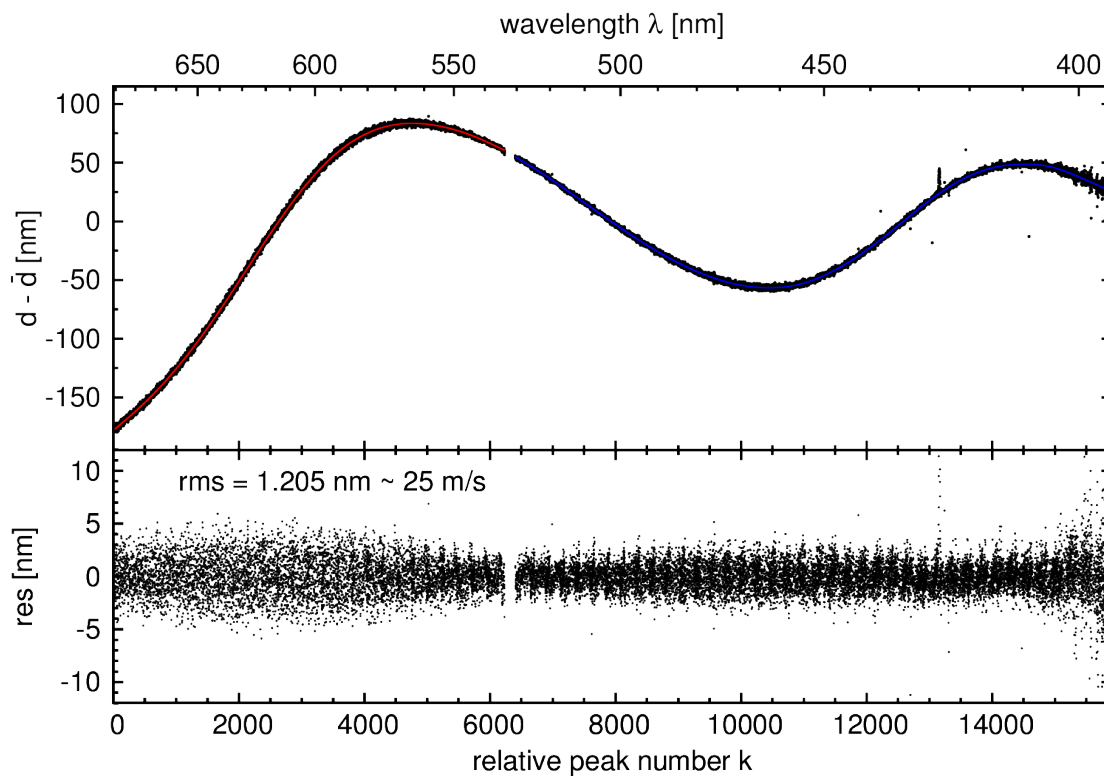


FIGURE 2.15: Effective cavity width measurement of the HARPS FPI using the thorium wavelength solution. Upper panel: Measured effective cavity width (black dots), spline fit to red and blue HARPS CCDs (red and blue line, respectively). Lower panel: Residuals of the spline fit. Orders 89 – 153 are shown here.

TABLE 2.1: Deviation of column sizes,  $a_{\text{pix},i}$  from unity in percent for HARPS. Estimated errors are typically 0.3% of a pixel ( $\sim 2.5$  m/s).

Column number $i$	509	1021	1533	2045	2557	3069	3581
Block row 1 ( $o = 135 - 161$ )	-4.3	1.1	-1.1	-3.5	0.5	-4.5	1.3
Block row 2 ( $o = 116 - 135$ )	-2.2	-0.8	-1.4	-1.0	-1.5	-5.8	-0.5
Block row 3 ( $o = 100 - 114$ )	-5.3	2.4	-3.1	-3.4	-0.5	-4.6	1.6
Block row 4 ( $o = 89 - 99$ )	-2.8	-1.0	-5.0	-0.4	0.2	-5.8	-0.2

To capture the effect caused by the CCD imperfections in our wavelength solution, we applied our pixel size concept (see Sect. 2.3.3) and modeled the size of every 512th pixel as indicated in Eq. (2.8). From the HARPS flat fields we see that the first pixel block is three pixels smaller than the rest and ends at pixel 509, while the rest continues in 512 pixel intervals. The edge of the detector is a small block of three pixels, which is not fitted. Furthermore, we assumed that the last pixels in the same column have the same size, which allowed us to limit the number of free pixel parameters to 14 per CCD. We found it necessary to increase the polynomial orders (eighth degree in dispersion direction and seventh degree in cross dispersion). This is possible because of the numerous FPI lines.

We overplot the model difference between the simple polynomial and the full model including inhomogeneous column sizes as solid red lines in the lower panel of Fig. 2.14. Our new wavelength solution handles most of the systematics that a simple polynomial model cannot account for. Using the FPI along with a wavelength solution that includes unequal pixel or column sizes therefore is a powerful method for characterizing the detector and calibrating the spectrograph at the same time.

Table 2.1 shows the deviation of column sizes from unity,  $a_{\text{pix}} - 1$ . The smallest column we found differs by 5.8% in size, which translates into about 46 m/s for HARPS.

### Comparison to DRS

To check the accuracy of our calibration using the HARPS FPI, we compared our wavelength solution to the standard calibration delivered by the HARPS Data Reduction Software (DRS). The difference between our wavelength solution and the cubic polynomial reverse regression of DRS is shown in the upper panel of Fig. 2.16. We observe distortions with amplitudes of up to 50 m/s that are repeated in every order. This is very similar to the results found in Wilken et al. (2010) and Molaro et al. (2013) using LFC data. These authors argued that the pattern is caused by DRS not taking into account different sizes of detector pixels. The difference between the LFC solution of Molaro et al. (2013) and the DRS solution yields an rms scatter of 25 m/s. We confirm the distortions in the DRS solution with the use of the HARPS FPI and find an rms scatter between our wavelength solution and the DRS solution of 22 m/s.

### Comparison to LFC

It is also interesting to compare our wavelength solution directly to the LFC solution. Unfortunately, there is no night in the HARPS archive with FPI and LFC spectra. Hence we computed the difference between the LFC wavelength solution of Molaro et al. (2013) for November 24, 2010 and our FPI wavelength solution for March 16, 2011 and corrected for 6 m/s, which is the mean drift between the two thorium solutions. The result is plotted in the lower panel of Fig. 2.16.



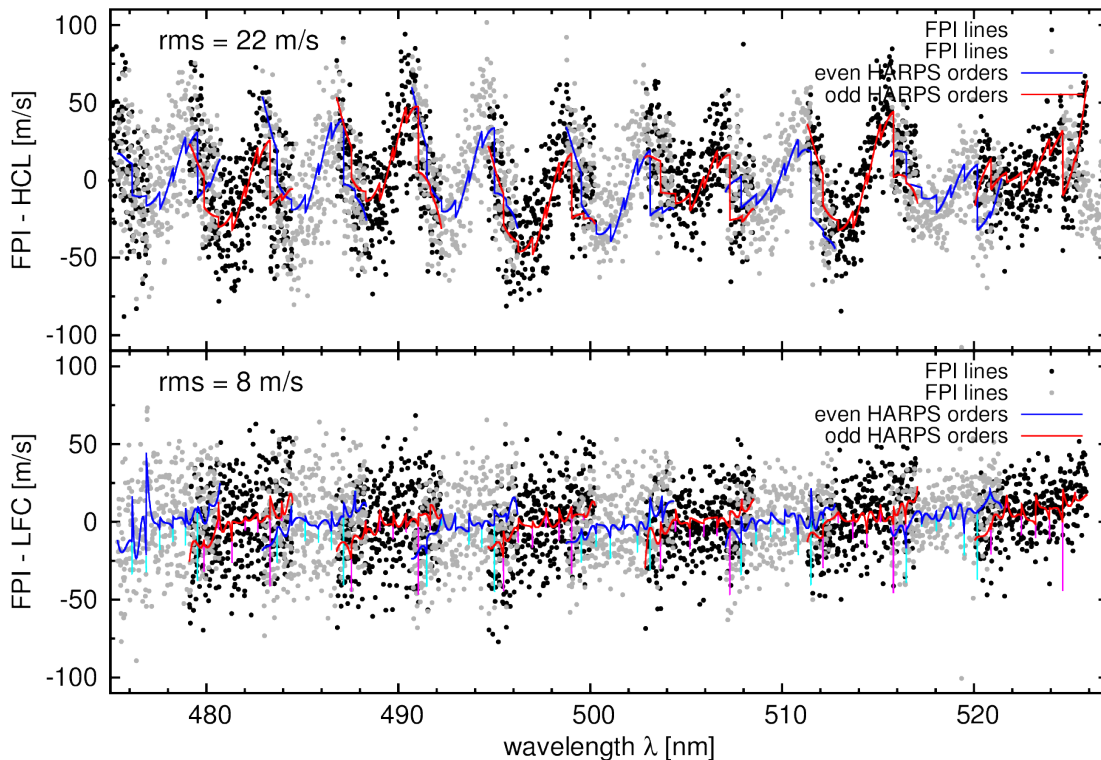


FIGURE 2.16: Upper panel: Difference between our FPI wavelength solution and the DRS wavelength solution (solid lines) with an rms of 22 m/s. Difference between FPI wavelengths obtained from the model  $d(k)$  and the wavelengths assigned to the FPI from DRS wavelength solution (points). Lower panel: Same as above, but using the LFC wavelength solution of [Molaro et al. \(2013\)](#). Orders 116 – 128 are shown here.

The deviations between our FPI wavelength solution and the LFC solution of [Molaro et al. \(2013\)](#) are smaller, only 8 m/s. Generally, no large amplitude distortions of 50 m/s are visible, as is the case with the ThAr solution.

We comment now on some remaining features. The peak around 477 nm is due to the low flux level of the LFC at this wavelength.

In six-pixel-wide regions at the block borders (marked in cyan and magenta in the lower panel of Fig. 2.16), the deviations can jump by 50 m/s. This is due to an indexing problem. We realized that the discontinuities (with a 512-pixel period) start at pixel 515 in the LFC solution of [Molaro et al. \(2013\)](#). We chose 509 as the start pixel to match the columns with an anomalous sensitivity in the flat echelle image that is most likely due to their pixel size.

In addition to this deviation, we observe a residual trend across the orders that is not well understood, but we note that four months passed between the two solutions. Differences with amplitudes of about 10 m/s are seen around this trend.

With the current data quality of the HARPS FPI spectra, we cannot completely resolve all the fine structure that is seen with the LFC. [Molaro et al. \(2013\)](#) individually fit a cubic polynomial for each block in each order, that is, their model has more parameters and is capable of handling more small-scale structure. We used a less flexible model because of the problems with the HARPS FPI spectra discussed in detail in Sect. 2.3.4. If the data quality improves, we might also be able to use more detailed models and to characterize the detector in more detail.

### Systematic high-frequency noise in the HARPS etalon

When we calibrated the HARPS FPI, the rms scatter around the B-spline model of  $d(k)$  was 25 m/s, see Fig. 2.15. This is about one order of magnitude more than what we expect from photon noise for signal-to-noise ratios of 150 to 300. We found that the residuals are not white but exhibit high frequency variations. The position residuals of lines present in adjacent and overlapping orders are correlated, which indicates a wavelength dependency. Frequency analysis revealed two significant periods. For order 131 we found  $P_1 = 1.5056 \pm 0.0017$  and  $P_2 = 0.6612 \pm 0.0005$ . Furthermore, periodic peak height variations of up to 10% are visible in the 1D extracted FPI spectra. Frequency analysis on the derived FPI amplitudes yields two significant periods that match those found in the position residuals. We assume that an additional low-finesse interference signal caused by plane-parallel optical elements in the path of the FPI (e.g. filters or the thickness of the FPI mirrors itself) superimposes on the HARPS FPI (Wildi et al., 2010; Pepe, private communication). The periods would correspond to FPIs with cavity widths on the order of millimeters. Furthermore, we find that the periods decrease towards blue and increase towards red orders, indicating that the effective cavity width of the optical element that causes the additional FPI effect changes with wavelength; this can be caused by a wavelength-dependent refractive index  $n$ . We assume that the slopes due to the flux variations shift the photocenter of the FPI lines, which in turn results in systematics in the line position measurement.

Without exact knowledge about the cause of this effect, our attempts to fit and correct the flux variations present in the spectra gave unsatisfying results. Because of this effect the FPI wavelengths are not known precisely enough to resolve the fine structure that is seen in the LFC. Our method involving the FPI in the wavelength calibration for HARPS is thus limited by this systematic high-frequency noise. Our laboratory experiment in Sect. 2.3.2 is limited by photon noise (8 m/s). With this data quality, we would be able to resolve more details in the wavelength solution systematics of HARPS. For future projects it is advisable to avoid optical elements in the light path of the FPI that might cause additional interference patterns and to characterize the FPI in advance with other instruments.

### 2.3.5 Influence of distortions in the wavelength solution on precise RV measurements

To estimate the influence of distortions in the wavelength solution on precise RV measurements with high-resolution echelle spectrographs, we simulated observations taken over the period of one year. The simulated spectrograph has 70 spectral orders that each cover  $57 \text{ \AA}$ , a delta peak instrumental profile, and a detector with 4096 homogeneous pixels. The observed star is perfectly quiet and has a zero RV signal. Its spectrum consists of randomly distributed delta peaks (100 lines per order). Furthermore, the star lies in the ecliptic, resulting in a 30 000 m/s RV signal from the barycentric motion of Earth. This oversimplified approach allows us to single out the effect of the wavelength solution alone on our observations without additional noise from any other sources.

In total we simulated 29 observations over one year. We initialized the stellar spectrum by computing random wavelengths such that 100 delta peaks fell on each of the 70 orders of the spectrograph. At time stamp 0 the barycentric motion of Earth toward the star is zero; hence this spectrum represents the rest frame. The spectra for all observation time stamps were computed applying only the Doppler shift corresponding to the barycentric motion of Earth around the Sun ( $\sim 30\,000 \text{ m/s}$ ). We ignored the effect

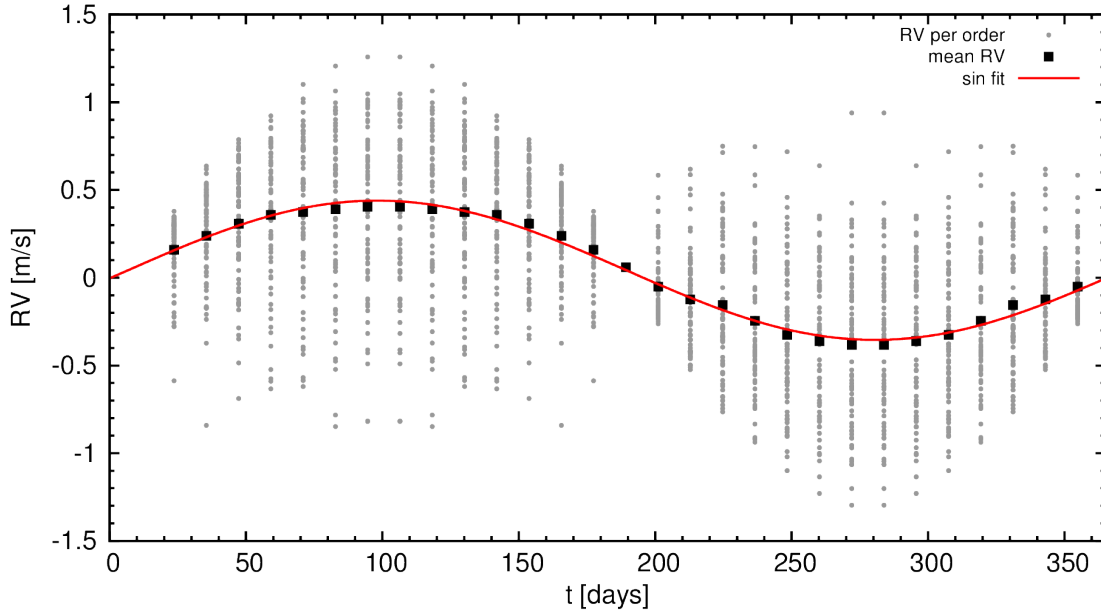


FIGURE 2.17: Simulation of the RV signal due to a disturbed wavelength solution. RV signal for single orders (gray dots), mean RV of all orders (black squares), and sinusoidal fit to the mean RV (red line).

of Earth's rotation because it is small compared to the motion of Earth around the Sun ( $\sim 500$  m/s). Spectra were taken every 12.175 days. We used the inverted wavelength solution of the spectrograph to place the stellar lines on the detector and to obtain our simulated observations.

When the true wavelength solution of the simulated spectrograph was used to convert the pixel value back to wavelength and the barycentric motion of Earth was corrected for, we obtained RVs of 0 m/s for all stellar lines within machine precision.

Next we disturbed the wavelength solution of the spectrograph. We wished to test the effect of small systematics in the wavelength solution on precise RV measurements. To do this, the difference between our thorium wavelength solution and our calibrated FPI wavelength solution in HARPS order 118 (Sect. 2.3.4, red line in Fig. 2.14) was added to the true wavelength solution of the simulated spectrograph.

Now we used the disturbed wavelength solution, converted the pixel values into wavelengths, corrected for the barycentric motion of Earth, and computed the RVs for all lines. We averaged the RV of single lines to compute the RV order by order for each observation (gray dots in Fig. 2.17). The RV per time stamp is the mean of the individual order RVs (black squares in Fig. 2.17).

We observe the following effects in Fig. 2.17. The RVs derived for individual orders (gray dots) show a scatter of about 1 m/s. This effect is a result of the random distribution of lines on the CCD and the systematics in the wavelength solution.

Depending on the location of a stellar line on the CCD, it will give a different systematically incorrect RV because the systematical error in the wavelength solution is different for different pixels. We can derive the expected RV signal  $\delta_v$  resulting from the wavelength solution systematics  $\delta_\lambda(x)$  by inserting the wavelength solution distortion in the Doppler equation:

$$\frac{[\lambda(x_i) + \delta_\lambda(x_i)] - [\lambda(x_0) + \delta_\lambda(x_0)]}{\lambda(x_0) + \delta_\lambda(x_0)} = \frac{v + \delta_v}{c}, \quad (2.10)$$



where  $x_0$  and  $x_i$  correspond to the line positions on the CCD at time stamps  $t_0 = 0$  (rest frame) and  $t_i$ , respectively. The barycentric motion of Earth shifts lines on scales of a few tens of pixels. On this scale, the change in  $\delta_\lambda$  is small and we can approximate the difference  $\delta_\lambda(x_1) - \delta_\lambda(x_0)$  with the derivative of the wavelength solution distortion multiplied by the pixel shift  $d\delta_\lambda(x)/dx \cdot \Delta x$ . The pixel shift  $\Delta x$  depends on the barycentric velocity of Earth  $RV_E$  and the pixel scale in terms of RV  $\kappa$ . For HARPS one pixel covers approximately 800 m/s/pix. Hence a barycentric velocity of  $RV_E = 30\,000$  m/s shifts the lines by 37.5 pixels on the detector. Finally, the small term  $\delta_\lambda(x_0)$  in the denominator can be neglected, and the velocity signal of a single line  $\delta_v$  that is due to the distortion in the wavelength solution  $\delta_\lambda(x)$  can be written as

$$\delta_v = \frac{c}{\lambda(x_0)} \cdot \frac{d\delta_\lambda(x)}{dx} \cdot \frac{1}{\kappa} \cdot RV_E. \quad (2.11)$$

Averaging over all the different  $\delta_v$  in one order results in a systematic RV that depends on the random distribution of stellar lines on the CCD. As the distortion is the same for all orders, the simulation results presented in Fig. 2.17 are equivalent to repeating the test 70 times for one order with different random sets of lines. This means that we observe scatter between the RVs of different orders that is solely caused by the line distribution on the detector. This scatter is mainly caused by lines experiencing different slopes in  $d\delta_\lambda(x)/dx$  and not by single lines moving across the jumps in the wavelength solution.

The slope  $d\delta_\lambda(x)/dx$  is negative almost everywhere on the CCD, however. Many lines therefore experience systematic RV shifts in the same direction. The mean RV derived from the individual orders accordingly shows a low-amplitude RV signal with a period of one year. The amplitude of this signal is about 0.5 m/s. Advancing to cm/s precision, we cannot afford systematic errors on the order of 0.5 m/s to 1 m/s. This clearly shows that accurate wavelength calibration is critical for reaching the aims with the next generation of instruments.

## 2.4 Calibrating the CARMENES spectrographs

### 2.4.1 The CARMENES survey

CARMENES stands for Calar Alto high-Resolution search for Mdwarfs with Exo-earths with Near-infrared and optical Echelle Spectrographs. The CARMENES instruments are two high-resolution spectrographs dedicated to planet hunting and designed to carry out an RV survey of low mass stars over a period of three years. Although open time is available for the community, the main focus of the CARMENES project is to monitor a sample of 355 M-stars. CARMENES focuses on targets with low and moderate activity to increase the chances of planet detections and minimize false positive rates. M-stars are interesting for planet searches because their low masses enable the detection of terrestrial planets in their habitable zones using an instrument with an RV precision of 1 m/s. For earlier type stars, planets down to one or two Earth masses are not accessible with current state-of-the-art instruments because the precision required is beyond reach of any currently existing spectrograph. The CARMENES survey is the first to provide enough statistics to get clear insights into the planet population around very low mass stars. The main scientific focus lies in stars with spectral types M4V and later because low mass companions are easier to detect. However, the CARMENES sample also includes earlier spectral types. This is because of two reasons:

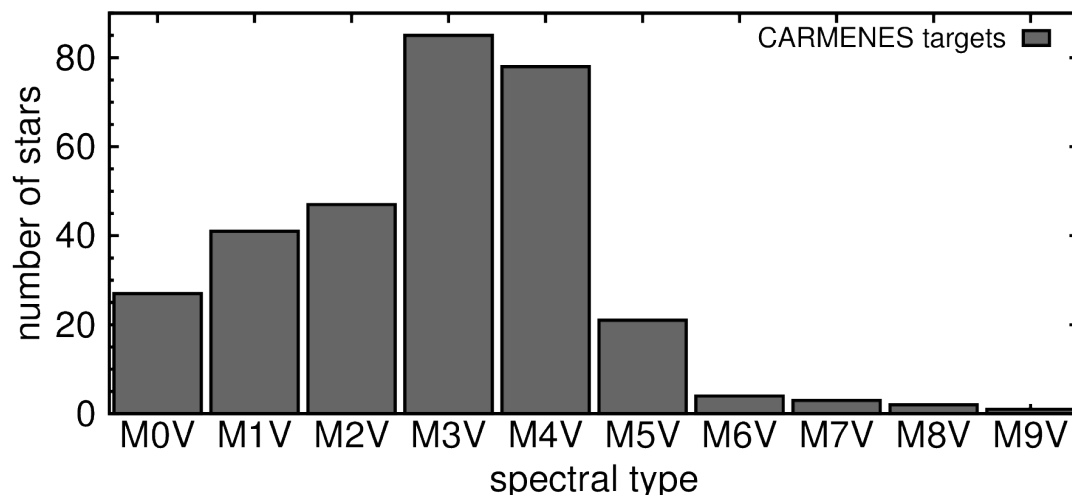


FIGURE 2.18: Histogram of the spectral type distribution of CARMENES targets.

- first, the amount of reasonably bright targets with late spectral types is limited, and
- second, brighter targets have to be available for observations under imperfect weather conditions including cirrus or bad seeing.

Figure 2.18 shows the distribution of the CARMENES sample in terms of spectral types. The distribution peaks around spectral types M3V and M4V because these targets are still bright enough for CARMENES while they start to become difficult to observe with spectrographs designed for bluer wavelengths (see Sect. 2.4.2). Bad weather targets are mainly early M-dwarfs. For stars with later spectral types, the distribution thins out because there are not many bright, very low mass stars in the sky. To fulfill the aim of CARMENES to detect small exoplanets, the survey will collect at least 60 measurements for each star in the target list over a period of 600 clear nights.

## 2.4.2 Instrument overview

The two high-resolution spectrographs of the CARMENES project are located in the Coudé room of the 3.5 m telescope at Calar Alto. The optical arm, CARMENES VIS, operates in the wavelength range between 0.52 and 0.96  $\mu\text{m}$  while the infrared arm, CARMENES NIR, covers the wavelength range from 0.96 to 1.71  $\mu\text{m}$  (Quirrenbach et al. 2016 subm.). Therefore, CARMENES is the first project aiming at a precision of 1 m/s in the optical and near-infrared domain.

To achieve the precision necessary to discover Earth-like planets around M-stars, both spectrographs are similar in design and were constructed for a resolution of  $R = 80000$ . After commissioning the instruments, CARMENES VIS turned out to surpass its requirements and achieve a resolution of about  $R = 93000$ . CARMENES NIR is well within the specification and offers a measured resolution of about  $R = 82000$  (Quirrenbach et al. 2016 subm.).

The optics of the spectrographs are housed in vacuum vessels. CARMENES VIS is not actively temperature controlled but kept at a moderate temperature of about 12 C° achieved by the air conditioning of the Coudé room. The near-infrared channel of CARMENES requires active cooling to reduce the thermal glow of the spectrograph

itself. A flow of liquid Nitrogen keeps the inside of the vacuum vessel at constant 140 K.

Despite the similarities in optical design and the vacuum tanks, the detectors of CARMENES VIS and CARMENES NIR are very different. CARMENES VIS uses a 4k x 4k e2V CCD. The CCD is not monolithic but instead 32 single blocks of 512 x 1024 pixels are stitched together with a stepper machine to form the detector. To reduce the read out time the CCD is operated with 4 amplifiers, each reading one quarter of the CCD. CARMENES NIR uses two 2040 x 2040 pixel Hawaii-2RG detectors to form an array of 4080 x 2040 pixels. However, between the two detectors is a gap, about 140 pixels wide, leading to intra-order gaps in the spectra. Because the visual and near-infrared detectors have different properties, the wavelength solution approach of both instruments is slightly different, see Sect. 2.4.8.

Both spectrographs have their own individual calibration unit. Each unit can carry up to eight lamps of which five can be turned on at the same time. A lamp can be selected by a rotating mirror with no danger of contamination by other light sources because the system is designed light tight. The advantage of turning on several lamps simultaneously is to minimize the time needed for calibration runs because all required lamps are heated up at the same time. In addition to the eight slots for lamps, there is an additional fiber allowing for an external FPI. The near-infrared calibration unit carries a flat field lamp and several UNe lamps. Additionally it can be supplied by an FPI especially designed for CARMENES NIR. The visual calibration unit also carries a flat-field lamp and two UNe lamps, but it also holds two UAr and two ThNe lamps and is supplied by a second FPI designed especially for CARMENES VIS. So in total, CARMENES uses 5 independent wavelength calibration sources.

As CARMENES is built as an ultrastable spectrograph system, both arms are fed via optical fibers. Each channel is supplied by the light of two fibers simultaneously. The first fiber is the science fiber used to guide starlight from the telescope to the spectrograph. The second fiber is the calibration fiber which can either be used to simultaneously record light from the calibration unit (usually the FPI) or to record the sky background for faint targets. What is special about CARMENES is that the two octagonal fibers (science and calibration) are sliced into two halfmoon shaped halves which are stacked on top of each other to form the slit.

Before CARMENES, all high-precision RV instruments focused on the visible spectral range. Observing in the visible has several benefits for technical and astrophysical reasons. A spectrograph designed for the visible spectral range does not require active cooling, while moving to the infrared requires complex thermal designs and advanced cooling systems. Detector technology is simpler and cheaper in the visible, while high-precision RV work has never been attempted in the near-infrared. For the last decades F, G and K stars have been main targets for planet searches because these hosts more closely resemble our Sun. These stars emit a large portion of their light in the visible spectral range, which is why spectrographs were built for this spectral region. Lately, however, the focus of RV surveys has moved towards later spectral types because smaller planetary companions are easier to detect around smaller hosts. M-stars emit more light in the infrared, and thus having a spectrograph observing in the infrared is a key for low mass planet detections around M-dwarfs. Figure 2.19 shows a Phoenix model spectrum (Husser et al., 2013) of an M4V star in black (left flux scale) and a G2V star in gray (right flux scale). The blue and red bars in the lower panel represent the wavelength coverage of CARMENES VIS and NIR respectively. The gray bar represents the wavelength coverage of HARPS. For G-stars HARPS nicely covers the maximum of the stellar emission curve. M-stars are faint in the visible and emit

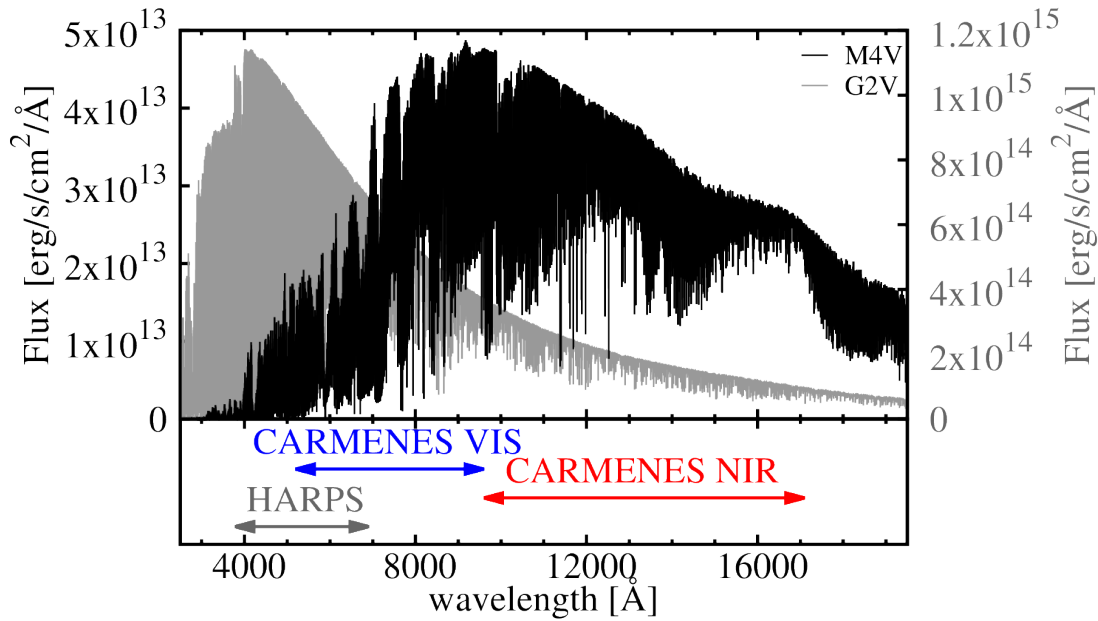


FIGURE 2.19: Illustration of the HARPS and CARMENES wavelength coverage. Upper panel: Phoenix spectrum of an M4V star (black) and a G2V (gray). Lower panel: wavelength coverage of HARPS (gray arrow), CARMENES VIS (blue arrow) and CARMENES NIR (red arrow).

most of their light red-wards of the HARPS spectral coverage. Therefore, these low mass targets are hard to observe with HARPS. Because most spectrographs designed for the visible spectral range have problems observing M-stars, CARMENES was designed especially for these low mass targets. As shown in the illustration in Fig. 2.19, CARMENES covers the spectral region where most of the stellar light from M-dwarfs is emitted. The ability to observe M-dwarfs with high efficiency makes CARMENES so important for planetary science.

For planet discoveries it is not only important to receive enough light and achieve high S/N ratios in the spectra, but it is also important to observe in spectral regions that have a rich RV content. [Reiners et al. \(2010\)](#) showed that the sweet spot for RV measurements of early M-stars is the V-band while M-stars later than M6 have most of their RV information in the Y-band. CARMENES covers the V, R and I-band with its visual arm and the Y, J and H-band in the near-infrared channel. Thus, CARMENES is perfectly designed for exoplanet searches around M-dwarfs.

Stellar activity has become a major concern for planet hunting because it has caused several confirmed false positives and many more cases still under debate (e.g. [Udry et al., 2007](#); [Vogt et al., 2010](#); [Robertson et al., 2014, 2015](#), [Anglada-Escudé et al., 2014, 2015](#)). Spots on the stellar surface can resemble planetary signals and stellar activity indicators are necessary to distinguish a signal of planetary nature from spots and plagues. CARMENES will be able to use the wavelength dependent spot contrast to verify the nature of a potential RV signal. Spots have higher contrasts in the visible spectral range and thus spots are expected to have larger RV amplitudes in the visual than in the near-infrared. A Keplerian signal is not wavelength-dependent and thus planets are expected to produce the same RV amplitude in both CARMENES arms. The large wavelength coverage pioneered by CARMENES will allow for the first time to disentangle spots from planets.

TABLE 2.2: Table of CARMENES calibration sources.

Instrument	ThNe	UNe	UAr	FPI
CARMENES VIS	✓	✓	✓	✓
CARMENES NIR	×	✓	×	✓

### 2.4.3 CARMENES calibration resources

As this part of the thesis focuses on improving the RV precision by improving the wavelength calibration, the CARMENES calibration procedure is explained in detail here. The wavelength calibration plan for CARMENES foresees a combination of HCLs and FPIs in both spectrograph arms. For the visual channel, ThNe lamps in combination with an FPI designed specifically for CARMENES VIS were planned. To calibrate the near-infrared arm a UNe and a second FPI designed for CARMENES NIR was foreseen. Neon was chosen as the lamp filling gas for the CARMENES NIR lamps because compared to argon, neon emits less lines that saturate the detectors in the near-infrared from 0.96 to 1.71  $\mu\text{m}$ . In the spectral range of CARMENES VIS, argon emits most strong lines at red wavelength above 0.7  $\mu\text{m}$  while neon shows strong lines mainly in the blue part of the spectrum below 0.7  $\mu\text{m}$ . To keep contamination in the more interesting red parts of the spectra low, neon was initially also chosen in the lamps for CARMENES VIS.

All HCLs were planned to be used with an operational current of 6 mA. According to the manufacturer, the life expectancy of the HCLs is about 4000 to 5000 mAh, which translates to a lamp lifetime of up to one year if the lamps are only used during the 2 hour daily calibration runs. However, an aging test in the optics laboratory of Göttingen revealed extended life times at least 35 percent above the expectations (Sarmiento et al. in prep.).

During commissioning in November and December 2016, the first lamp tests with CARMENES revealed a very low number of calibration lines in both spectrograph arms. In addition, the available thorium and uranium lines were very weak. Hence, the initial calibration strategy was changed. The first step was to enhance the line ratio between metal and gas lines by changing the operational current for all lamps from 6 mA to 12 mA. This change comes at the expense of a reduced lamp lifetime but the amount of usable calibration lines increases (Kerber et al., 2008). As there are 12 spare lamps of each kind, a shorter lamp lifetime is of no concern for the project.

After changing the operational current, the UNe lamp provided enough lines to calibrate CARMENES NIR reliably. For CARMENES VIS, the higher operational current also increased the number of thorium lines (about 1800 usable lines) but compared to HARPS (about 3900 lines), the amount of lines available for calibration is still a factor of 2.2 less. CARMENES VIS covers a redder spectral range than HARPS does, which is one reason for the lower number of calibration lines in the CARMENES spectra. Additionally, new lamp production techniques might have influenced the quality of the calibration spectra (Pepe; Sarmiento; private communication). To have more lines available for wavelength calibration, the strategy for CARMENES VIS was changed, making it the first RV instrument to use multiple calibration lamps of different kinds. Some ThNe lamps of the visual calibration unit were exchanged with UNe and UAr lamps so that the instrument uses four independent wavelength calibration sources in parallel. An overview of the wavelength calibration sources for CARMENES is given in Table 2.2.



#### 2.4.4 Line catalogs

For wavelength calibration, the standard lines detected in the echelle spectra have to be referenced to catalogs. In the infrared the uranium catalog of [Redman et al. \(2011\)](#) is used. Out of 5990 uranium lines listed for the wavelength range of CARMENES NIR, 2778 lines are detected in the daily standard calibration spectra. Because the bluest orders have a slight order overlap 56 lines are detected in two orders, so that 2834 lines are available in total to calibrate the 28 orders of the near-infrared arm.

In the visible spectral range, the catalog of [Redman et al. \(2014\)](#) is used for the ThNe spectra, and the catalog of Sarmiento et al. (in prep.) is used for both uranium lamps. In the daily ThNe calibration spectra, 1634 out of the 10432 listed thorium lines are detected. Because the echelle orders overlap in CARMENES VIS, 223 thorium lines are detected twice, resulting in 1857 thorium lines available for calibration. Sarmiento et al. (in prep.) compiled the only uranium line list available in the visible spectral range so far. The catalog was measured using the FTS in the optics laboratory in Göttingen and lists 1313 lines. Unfortunately, the HeNe laser of the FTS contaminates the regions between 617.8 and 650.5 nm so that no line positions are available for 5 spectral orders of the CARMENES VIS spectra. Out of 1313 lines listed in the uranium catalog, 1148 lines are found in the UNe spectra and 1155 lines are detected in the UAr spectra. In the UNe spectrum, 393 lines are present in two orders, so that in total, 1541 uranium lines are available from this lamp in CARMENES VIS. For the UAr lamp, 414 lines are detected in two orders, so that 1569 uranium lines are seen by CARMENES VIS.

Echelle spectrographs like CARMENES are much more sensitive than an FTS, and therefore a large number of unidentified lines are seen in the daily calibration spectra of the uranium lamps. Figure 2.20 shows a region of the raw spectrum of the UAr lamp together with lines listed in the uranium catalog of Sarmiento et al. (in prep.) marked in green. Overlaying the CARMENES spectra with the catalog shows that the lamp emits many more weak lines that could not be detected by the FTS.

As uranium lamps are not widely used to calibrate spectrographs in the visible, the emission lines seen with CARMENES are counted. This shows the calibration capabilities of uranium lamps for high-resolution spectrographs. The first step of line identification is to find all emission peaks exceeding a specific S/N ratio (typically 10) in the 1D extracted spectrum. In a second step the detected lines have to be cleaned from misidentifications of overspill stripes or Lyman ghosts of saturated strong lines.

Overspills are lines of saturated pixels emerging from bright gas lines. The overspill effect is created when a pixel carries more charge than its potential well can hold. Additional electrons then spill to neighboring pixels, contaminating or even saturating them. In extreme cases, saturated pixels can effect the readout register and thereby contaminate entire rows ([Neely & Janesick, 1993](#)). As an example the yellow arrows in Fig. 2.20 indicate some overspills. To clean the newly identified lines from overspills, the 2D images are used. The position of all line candidates on the image is found and all lines are cross correlated with the distinct, 2D double halfmoon shaped line profile of CARMENES. If the cross correlation coefficient is lower than 0.9, lines are flagged as overspill. With this method overspill contamination is identified efficiently because overspill stripes are narrower and show no double halfmoon shape like cathode emission lines.

Periodic errors in the grating lead to so-called Lyman ghosts. Lyman ghosts are shifted copies of the original spectral line and therefore have the typical halfmoon shape. Figure 2.20 shows a few Lyman ghosts marked with the red circles. Lyman ghosts are only visible for strong emission lines because they are orders of magnitude

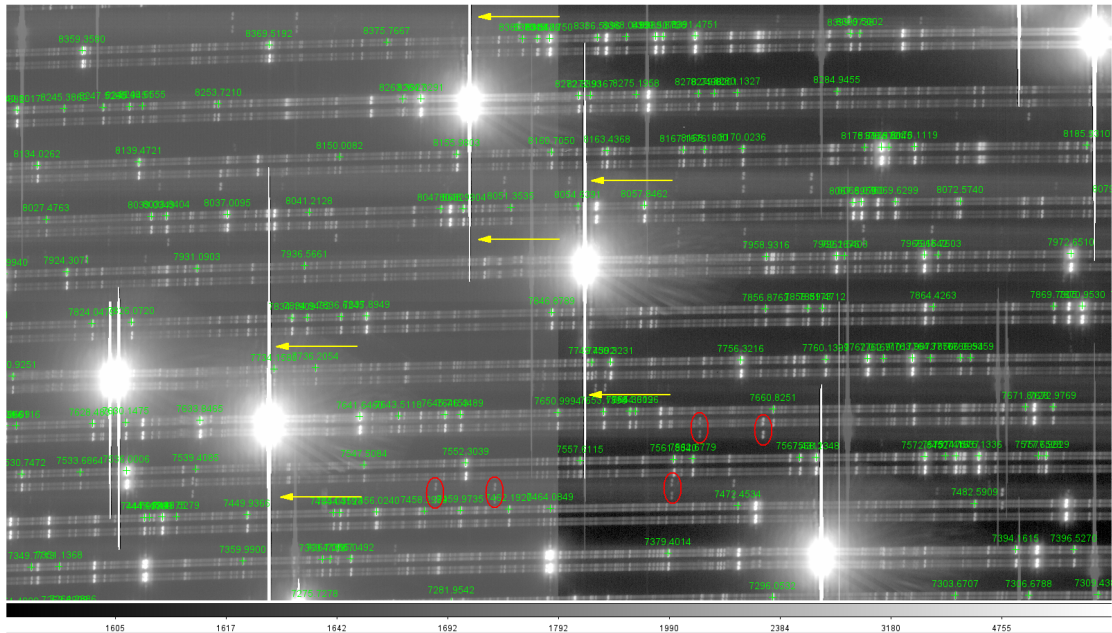


FIGURE 2.20: Raw spectrum of the UAr lamp taken with CARMENES VIS. Green crosses represent the catalog of Sarmiento et al. (in prep.). Examples of overspill stripes are marked with yellow arrows and some Lyman ghosts are marked with red circles.

fainter than the line they originate from. One property of these ghosts is that they do not follow the order trace. Hence, Lyman ghost lines are identified by measuring the distance of a line perpendicular to the order trace. Lines that are misplaced by more 1.5 pixels are sorted out as Lyman ghosts.

For the UAr lamp, a total of 6950 emission lines are identified, and for the UNe lamp, 5640 lines are seen in the CARMENES VIS spectra. These, however, are all identified lines without distinguishing between uranium, neon and argon lines. Nevertheless, most lines are expected to come from uranium because the low atomic number of argon and neon limits the number of spectral lines emitted. For argon, the NIST data base lists 138 lines with significant flux in the spectral range of CARMENES VIS. Neon contaminates the visual channel a bit stronger with 472 lines found in the NIST database. The number of potential uranium calibration lines is very high in the CARMENES VIS spectra. However, a detailed and deep catalog is hard to obtain with FTS spectra. Hence, a uranium catalog using CARMENES is in preparation (Klüter et al. in prep.).

## 2.4.5 Optimizing the hollow cathode lamp outputs

### Lamp exposure times

When choosing the exposure time for the calibration lamps, a trade-off has to be made between the number of metal lines detected and the contamination done by the strong carrier gas lines. In Fig. 2.21, two raw images of the ThNe lamp taken with CARMENES VIS are shown. On the left the exposure time was 1 s, and for the right image the exposure time was 5 s. Two changes between the images are noticeable. First, because of the longer exposure time, more weak lines stick out of the noise and become available for calibration. Second, the overspill of saturated gas lines on the CCD increases. The

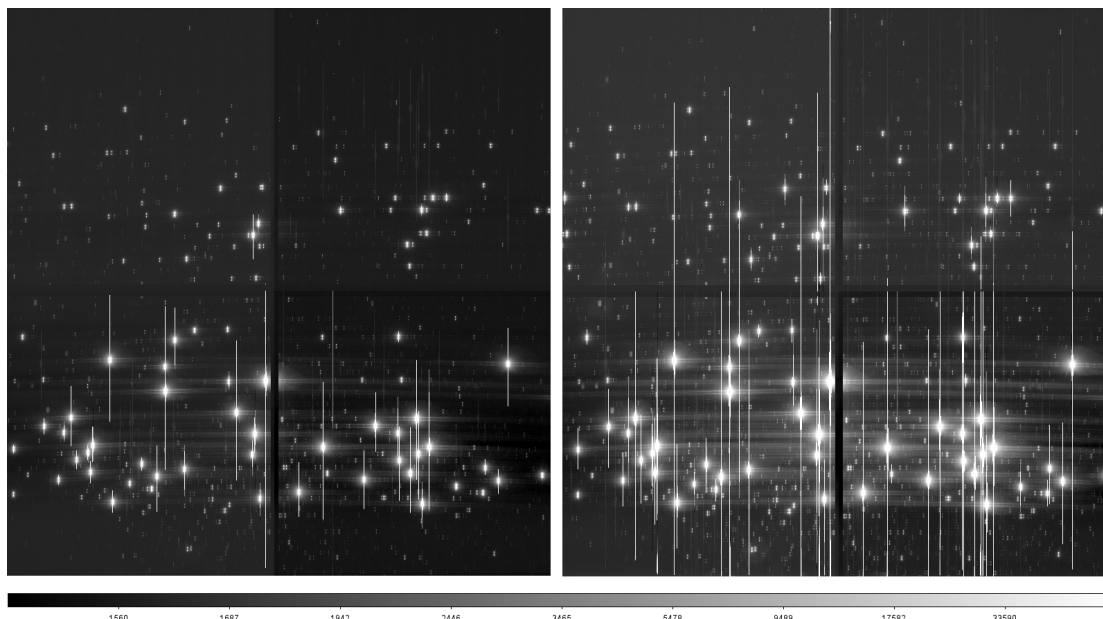


FIGURE 2.21: ThNe spectra taken with two exposure times. Left: integration time of 1s, right: integration time of 5s.

overspill effect grows with exposure time because more charge is accumulated in the already saturated pixels, which is then distributed to the neighboring pixels. In the visual arm, overflows are a big problem. Unfortunately, the filling gas lines of all lamps are so strong that overflow effects are easily created in CARMENES VIS and are rapidly growing with increased exposure time.

To find the best exposure time that maximizes the number of calibration lines, a series of lamp spectra were taken with different exposure times for all lamps. The ThNe lamp in CARMENES VIS is analyzed first. All CARMENES spectra are wavelength calibrated using a set of relatively strong thorium lines. Therefore, weaker thorium lines can be identified using the catalog of [Redman et al. \(2014\)](#). A line is flagged as present in the CARMENES spectrum if it matches the catalog within 500 m/s. This limit is chosen because it is wide enough to account for problems with the first guess wavelength solution of CARMENES but at the same time narrow enough to avoid mis-identification between close-by spectral features. Figure 2.22 shows the number of detected thorium lines in the ThNe lamp as a function of exposure time. On the one hand, it is observed that longer exposure times result in an increase of detected thorium lines but on the other hand, exposure times beyond 5 s do not provide a significantly larger amount of lines. The effect seen is a combination of the lamps' limited number of emission lines and the strong contamination of the spectra by bright lines with increasing exposure time. Doubling the exposure time to 1 s from 0.5 s yields an increase of detected thorium lines of about 20 percent, while doubling the exposure time to 10 s from 5 s results in less than 1 percent more identified lines. Hence the trade-off was to choose an exposure time of 1 s to keep contamination low but to record five spectra of the ThNe lamp in each calibration run, which results in an effective 5 s exposure when the spectra are averaged.

The exposure time experiment is also repeated with the UNe and the UAr lamps in CARMENES VIS to see if different lamps require different exposure times. Unfortunately the uranium catalog of Sarmiento et al. (in prep.) lists only 1313 lines in the



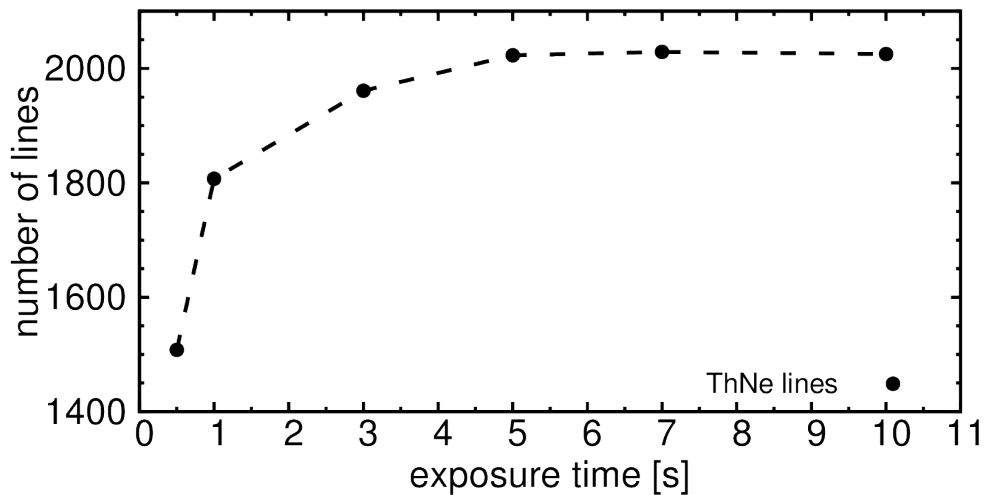


FIGURE 2.22: Number of detected thorium lines as a function of exposure time.

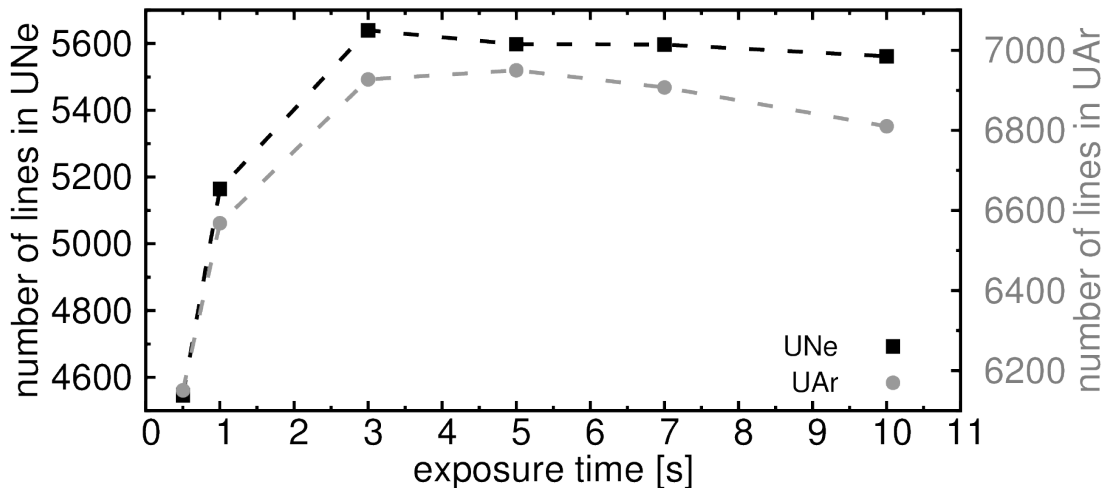


FIGURE 2.23: Number of detected lines in the UNe lamp (black squares) and in the UAr lamp (gray dots) with CARMENES VIS.

visible spectral range of CARMENES while many more potential uranium lines are visible in the CARMENES spectra (see Sect 2.4.4). This means that even for short exposure times the catalog is the limiting factor for uranium line identification. Therefore a slightly different approach is chosen to test the exposure time for the uranium lamps. Like in Sect. 2.4.4, the detected lines are counted after the candidate list is cleaned from contamination. In Fig. 2.23, the number of detected emission lines of the UNe and UAr lamp are presented as a function of exposure time. Like with the ThNe lamp, the biggest increase in the number of detected lines is happening when doubling the exposure time to 1 s from 0.5 s while longer exposure times yield a less significant increase of lines. In the 1 s exposure of the UNe lamp, about 14 percent more lines are seen as compared to the 0.5 s exposure. For UAr, the 1 s exposure reveals about 7 percent more lines than the 0.5 s exposure.

Exposure times longer than 3 s result in a decrease in the number of detected lines in both the UNe and UAr lamps. Like in the ThNe lamp, this effect can be explained

by the strong contamination by saturated neon and argon gas lines and the associated overspill stripes. As overspill in long exposures harm several neighboring orders, more and more lines are lost while less new weak lines appear. Because the uranium lamps have at least a factor of 3 more lines than the ThNe lamp, the chances that a line is affected by overspill are higher in the UNe and UAr spectra. Therefore, a slight decrease in the number of lines can be observed in the uranium lamps. Because the behavior of all lamps is similar, the exposure time was fixed to 1 s and five spectra per lamp and calibration run are taken with CARMENES VIS.

In CARMENES NIR, overspill is of no concern because the detector architecture is different. Nevertheless, the near-infrared detector of CARMENES exhibits electronic cross-talk between pixels in the same detector column. The two detectors of CARMENES NIR do not allow for exposure times less than 3 s. Because this minimum exposure time already provides a satisfying number of lines (2834), and in order to keep cross-talk effects low, the exposure time for the UNe lamp in the CARMENES NIR was chosen to be 3 s. Like in CARMENES VIS, 5 UNe spectra are taken in each calibration run.

### Fiber position

Operational current and exposure time are two parameters to maximize the number of calibration lines seen by a spectrograph. As CARMENES has fiber-fed spectrographs, there is a third parameter to improve the number of calibration lines. This third parameter is the fiber position in the focal plane of the lamp. As mentioned in [Huke et al. \(2016\)](#) the cathode of HCLs is a hollow cylinder. Therefore light is emitted as a ring rather than a central point source. Depending on where in the focal plane of the lamp lens system the pick up fiber is placed either more light from the glowing metal cathode or more light from the surrounding glowing buffer gas is observed. To maximize the line ratio between gas lines and metal lines, a fiber position test was performed with CARMENES. The fiber position was varied from 0 degrees (center) to a 0.5 degrees offset angle and three spectra of the ThNe lamp were taken for each position. The first ThNe spectrum was taken at the center of the focal plane and defined as reference spectrum. Thorium lines were identified from the catalog of [Redman et al. \(2014\)](#) and neon lines were identified from the NIST data base ([Reader et al., 2012](#)). The line amplitudes were measured with a Gaussian fit for each line. In each of the following ThNe spectra, the same line list was used to fit the amplitudes of all lines. For each spectrum, the line ratio relative to the reference spectrum is calculated by

$$R = \text{mean} \left( \frac{A_{Th} \cdot A_{Ne,ref}}{A_{Ne} \cdot A_{Th,ref}} \right). \quad (2.12)$$

Figure 2.24 presents the results of the fiber positioning test. As the fiber distance increases from the center of the focal plane, an increase in the line ratio between thorium and neon lines is seen at first. Hence the thorium lines increase in amplitude relative to the neon lines. The improved line ratio can be interpreted as a result of the fiber being moved inside the image of the glowing cathode ring. Increasing the fiber distance to the center further results in a decrease in relative strength between thorium lines and neon lines. It can be concluded that the fiber is moved outside the glowing ring and into the surrounding region of strong neon gas glow. Therefore neon lines are seen stronger than the thorium cathode lines. According to the measurements the best fiber position is located around a distance of 0.1 to 0.2 degrees from the center of the focal

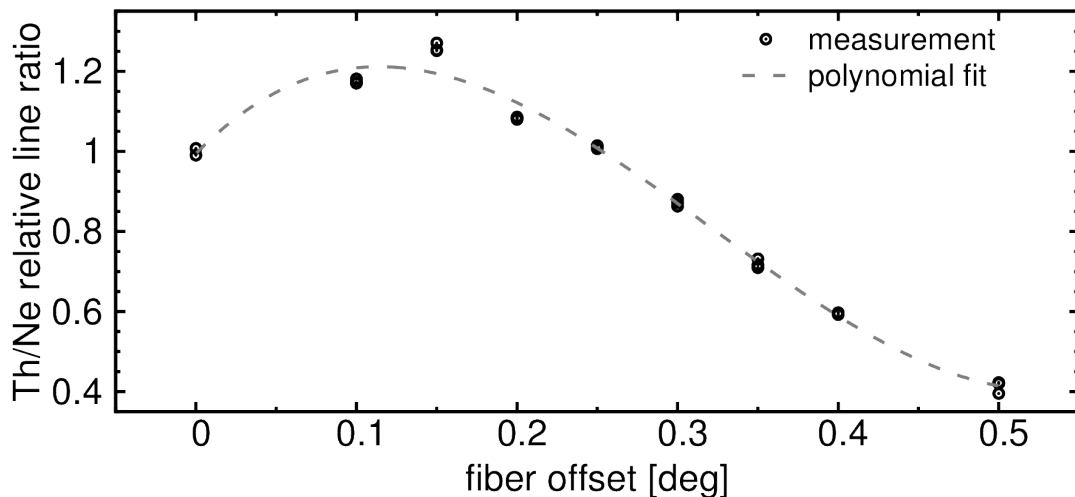


FIGURE 2.24: Relative line ratio between thorium and neon lines as a function of the fiber offset. Black circles mark the measurement and the gray line represents a polynomial fit to the data.

plane and the thorium lines gain approximately 20 percent in amplitude relative to the neon lines.

#### 2.4.6 Line catalog cross-referencing

With all lamps optimized for the number of calibration lines available, it has to be ensured that the lamps and line lists used for calibration of CARMENES are in agreement with each other. Especially for CARMENES VIS, it is important that there are no offsets between the lamps or the catalogs because three different lamps and catalogs of two authors (Redman et al., 2014; Sarmiento et al. in prep.) are combined. Offsets between the lamps or the catalogs will bias the wavelength solution. Depending on which lines are selected for wavelength calibration, offsets can result in uncontrollable, untraceable RV drifts from night to night. Hence cross-calibration and lamp monitoring is crucial to ensure the high-precision performance of CARMENES.

CARMENES VIS uses three kinds of lamps: UNe, UAr and ThNe. An offset between the uranium lines in the UNe and UAr lamp implies that there is an issue (possibly aging) with one of the lamps because the catalog used is the same for both lamps. The situation is more complex if there is an offset between the uranium lamps and the ThNe lamp. Because two different catalogs are used, the offset could originate either from differences in the absolute wavelength scale of Redman et al. (2014) and Sarmiento et al. (in prep.) or, in case the two uranium lamps are in agreement, from an issue with the ThNe lamp. If an offset between the uranium and thorium lamps is observed, the only way to test if the lamps or the catalogs are responsible is to measure more than one set of lamps. If the same offset occurs in several measurements of different thorium and uranium lamps, there is most likely an inconsistency in the catalogs. If the offset vanishes for different ThNe lamps, there is a bad lamp in the batch. The calibration unit for CARMENES VIS carries two sets of lamps. The first set is used for daily calibration and the second set of lamps are master lamps to trace lamp drifts over time. Thus the configuration in the VIS calibration unit allows distinguishing between all mentioned scenarios.

To cross-check the three HCLs, the UNe lamp is selected as the reference. The wavelength solution of CARMENES VIS is computed first only from uranium lines of the UNe lamp using the uranium catalog of Sarmiento et al. (in prep.). In the next step, the uranium wavelength solution is used to measure the wavelength of the uranium and argon lines in the UAr spectrum and thorium and neon lines in the ThNe spectrum. The difference between the wavelengths measured with CARMENES and the wavelengths listed in the catalogs are then computed.

On January 3 2016, the first calibration run using both sets of lamps was carried out. When cross-calibrating the UNe, UAr, and ThNe spectra of the daily calibration lamps, the mean offset between the two uranium lamps is found to be 4.7 m/s. For the UNe and ThNe lamp, a mean offset of only 0.2 m/s is measured. For the master lamp set, a mean offset between the UNe and UAr lamp of 4.9 m/s is obtained, and between the UNe and ThNe lamp, a difference of 1.9 m/s is measured. The error estimate for all these measurements are on the order of 35 m/s, and therefore none of the measured offsets is significant. Hence, it can be concluded that all three lamps and both distinct catalogs agree with each other within the measurement precision of 35 m/s. This is a remarkable result when we keep in mind that the catalogs used here were measured with two different instruments, and that three different lamps made with different cathode material and buffer gas are used.

## 2.4.7 The CARMENES Fabry-Pérot etalons

### The CARMENES Fabry-Pérot etalons in the optics laboratory

Apart from the HCLs, CARMENES also uses FPIs in the daily calibration of both spectrograph arms. The two CARMENES FPIs were extensively tested in the optics laboratory in Göttingen because they are used as drift references during the night. Before shipping the FPIs to Calar Alto, a drift test was performed with both FPIs inside their fully assembled, temperature-stabilized vacuum tanks for 24 hours. During these tests 504 spectra of the visual FPI and 550 spectra of the near-infrared FPI were taken. All spectra are averaged to compute a high S/N template for the visual and near-infrared FPI. From this template the effective cavity width and its variation over the CARMENES wavelength range is measured using the internal wavelength solution of the FTS.

The effective cavity width measurement of both FPIs is fitted with a smooth, natural b-spline (de Boor, 2001; Dierckx, 1995). For the visual FPI, 30 nodes are required to fit  $d$  from 510 to 1100 nm and for the near-infrared FPI, 45 nodes are needed to fit  $d$  in the region between 900 and 1700 nm. The  $d$  curves of both FPIs are displayed in Fig. 2.25 for the visual FPI and in Fig. 2.26 for the near-infrared FPI. The black points represent the effective cavity width  $d$  measured for individual FPI lines. In the upper panels of Fig. 2.25 and Fig. 2.26, the green solid curve is the b-spline fit to the data. For orientation, the gray shaded areas mark wavelengths beyond the wavelength coverage of the CARMENES spectrographs. The FTS measurements of both FPIs reveal no sudden changes in  $d$  so that a smooth model is justified, see Sect 2.3.2. In the lower panels of Fig. 2.25 and Fig. 2.26, the residuals of the b-spline fit to the FTS data are plotted. The rms scatter around the fit to the visual FPI  $d$  curve is 0.25 nm or 7.5 m/s. For the rms calculation, data points below 550 nm are excluded because the soft coating of the mirrors is not specified for these wavelengths. The transmission of the coating below 550 nm is above 70 percent, resulting in very low finesse values below 2.5 and high

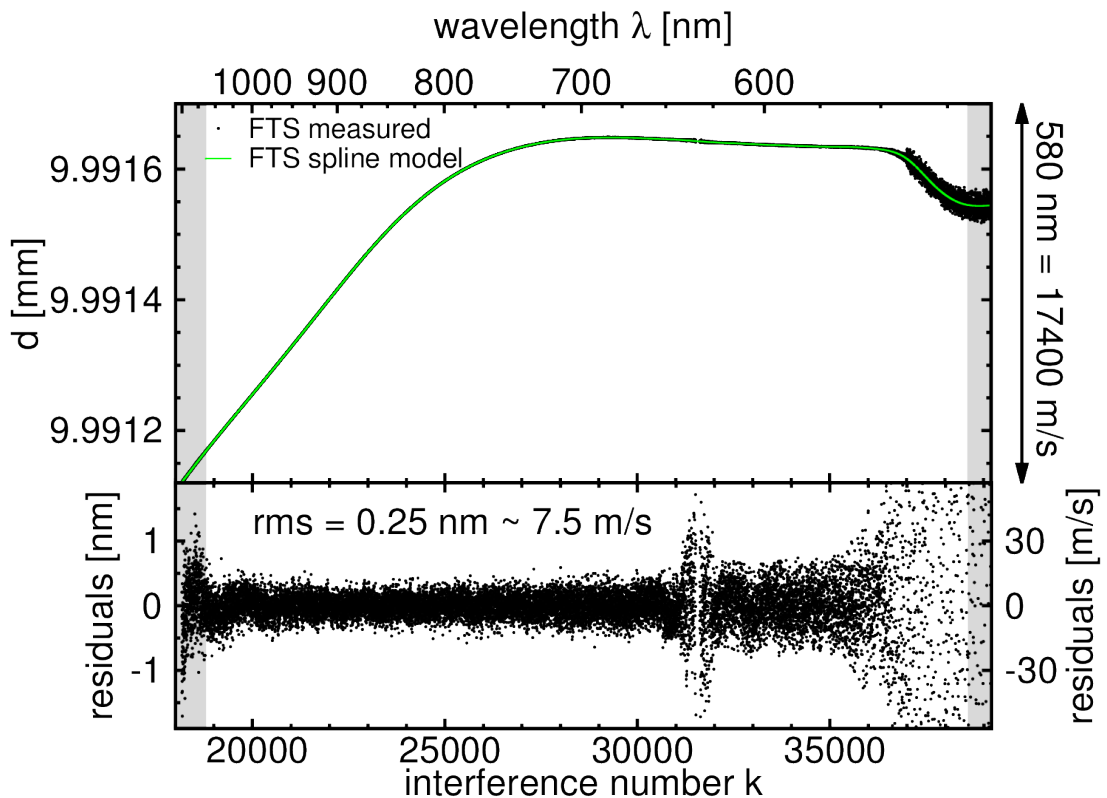


FIGURE 2.25: Cavity width measurement of the visual FPI with the FTS in the optics laboratory in Göttingen. Upper panel: Cavity width measured with individual FPI lines (black dots) and fit with a b-spline model (green line). Lower panel: residuals of the b-spline fit to the data (black dots). Gray areas mark wavelength regions outside the CARMENES VIS wavelength coverage.

uncertainties in the fitted positions. Additionally, the halogen lamp used for the experiment becomes faint for these blue wavelengths, which further increases the uncertainties in the line positions. For the near-infrared, an rms scatter of 1.8 nm, translating to 21.5 m/s, is measured in the region between 950 and 1600 nm. Like for the visible FPI, the coating transmission bluewards of 950 nm is high, resulting in increased uncertainties in the line positions. As the near-infrared coating is only specified up to 1350 nm, increased uncertainties of the line positions towards the red end of the spectrum are also observed. Nevertheless, the near-infrared FPI works fine up to 1600 nm. Beyond 1600 nm, however, the contrast drops fast, resulting again in large uncertainties of the fitted line positions.

As compared to CARMENES, the FTS measurements cover a larger spectral range and do not suffer from spectral gaps (like CARMENES NIR). Direct comparison between high quality FTS measurements and the CARMENES spectra can provide valuable information and help to improve the wavelength solution of both CARMENES channels. When using FPIs in wavelength calibration, measuring  $d$  is inevitable. As explained in Sect. 2.3.4,  $d$  can be measured using the HCL wavelength solution of the echelle spectrograph. Nevertheless, systematic errors in the input wavelength solution derived from HCLs may propagate through the entire  $d$  measurement process and go unnoticed. Up to now, there was no comparison data set to investigate this issue. Further, the b-spline fit to the  $d$  curve becomes more uncertain towards the wavelength

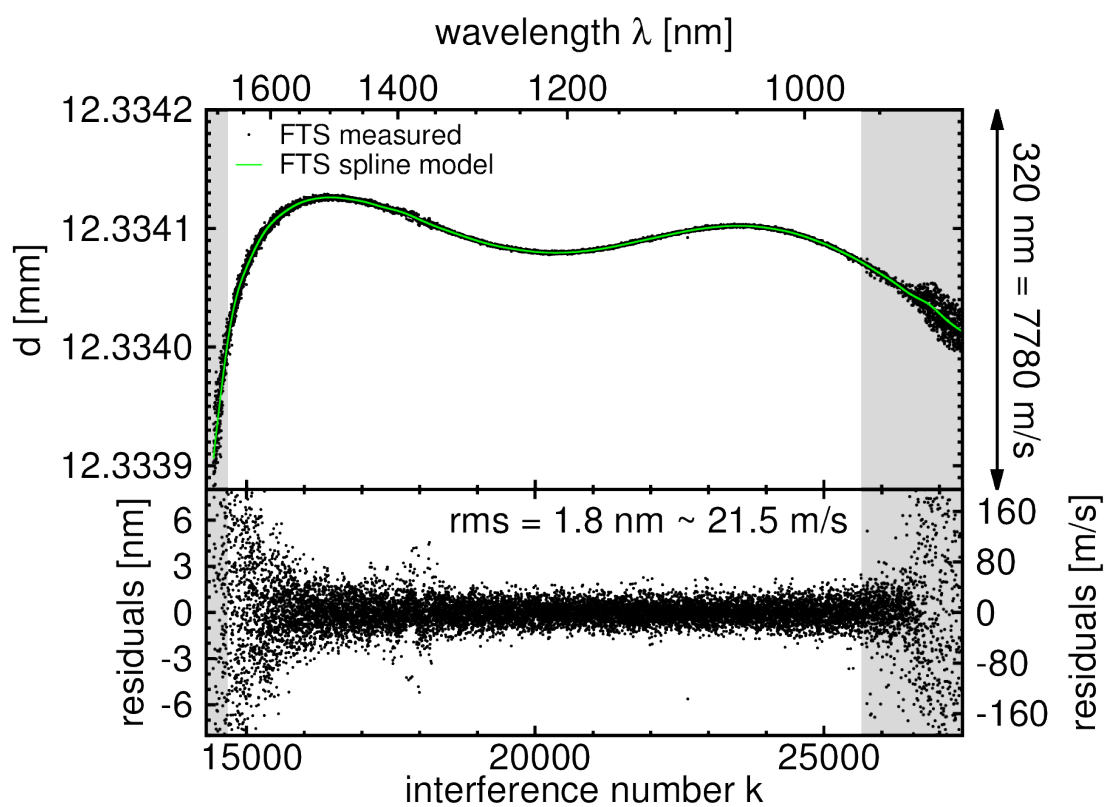


FIGURE 2.26: Cavity width measurement of the near-infrared FPI with the FTS in the optics laboratory in Göttingen. Upper panel: Cavity width measured with individual FPI lines (black dots) and fit with a b-spline model (green line). Lower panel: residuals of the b-spline fit to the data (black dots). Gray areas mark wavelength regions outside the CARMENES NIR wavelength coverage.



region edges because of the smaller number of data points constraining the model. As the data sets provided by the FTS measurements cover a larger wavelength range than the CARMENES spectrographs (see gray shaded areas in Fig. 2.25 and Fig. 2.26), this problem can be tackled. For CARMENES, we are in the unique position of having a comparison measurement done at very high-resolution. For the first time it is possible to compare the effective cavity width of the FPIs measured by two instruments and hence systematics potentially introduced by the echelle spectrograph and its optics can be identified.

### Comparing the FPIs in CARMENES and the FTS

As shown in Sect. 2.3.4 the wavelength solution of echelle spectrographs can exhibit systematics that could not be seen if only HCLs are used for calibration. Because both CARMENES FPIs have been tested with the FTS spectrograph in Göttingen we go one step further here and inter-compare the FPI measurements taken with both spectrographs to reveal systematics in the CARMENES wavelength solution.

First the effective cavity width  $d$  is measured in both spectral arms of CARMENES. This is done by following the basic steps in Sect. 2.3.4. For both CARMENES channels a first guess wavelength solution is computed based only on the HCL lines (for details of the wavelength solution see Sect. 2.4.8). Using this first guess wavelength solution translates the pixel positions of the FPI lines into wavelengths to calculate the effective cavity width curve  $d$  for the etalons. Figure 2.27 and Fig. 2.28 show the cavity width curves measured with CARMENES VIS and CARMENES NIR respectively.

In the upper panel of Fig. 2.27, the effective cavity width  $d$  of the visual FPI measured with CARMENES VIS is shown as black points. The green solid line represents the b-spline fit obtained from the FTS data taken in Göttingen. Globally, the effective cavity width obtained with CARMENES VIS and the FTS agree well. For the visual FPI a mean  $d$  offset between the CARMENES and FTS measurements of  $-3.5$  nm or  $-104.3$  m/s is measured. Typically the wavelength scale of the FTS spectra have to be corrected (Griffiths & de Haset, 2007) by measuring accurately known thorium or uranium standard lines (Degraffenreid & Sansonetti, 2002). The FTS data taken for the FPIs were not corrected with these standards so the error in the absolute wavelength scale of the spectra can be on the order of 100 m/s. Hence the global offset between the  $d$  curve measured with the FTS and the  $d$  curve measured with CARMENES VIS can be explained by the uncertainties in the FTS wavelength solution.

However, subtracting the  $d$  curve model obtained with the FTS measurements from the  $d$  curve obtained by CARMENES VIS results in residuals that do not scatter around a constant value, but instead exhibit a wavelength dependency (see lower panel of Fig. 2.27). The amplitude of this wavelength dependent pattern is 3.5 nm or 104.3 m/s. This is a hint for alignment changes in the FPI system that most likely happened during the shipment of the device to Calar Alto. The soft coating of the FPI mirrors consists of different layers of material reflecting different wavelengths. A change in illumination angle of the FPI changes the path of the light through the soft coating of the mirrors. Hence, the light entering sees different coating layer thicknesses. Because light of different wavelengths penetrates into different depths of the coating, a change in illumination angle is expected to result in the observed wavelength dependent change of the cavity width.

For the near-infrared FPI, the changes in the  $d$  curve after shipping the device to Calar Alto are larger than in the visible FPI. The CARMENES measurements of the



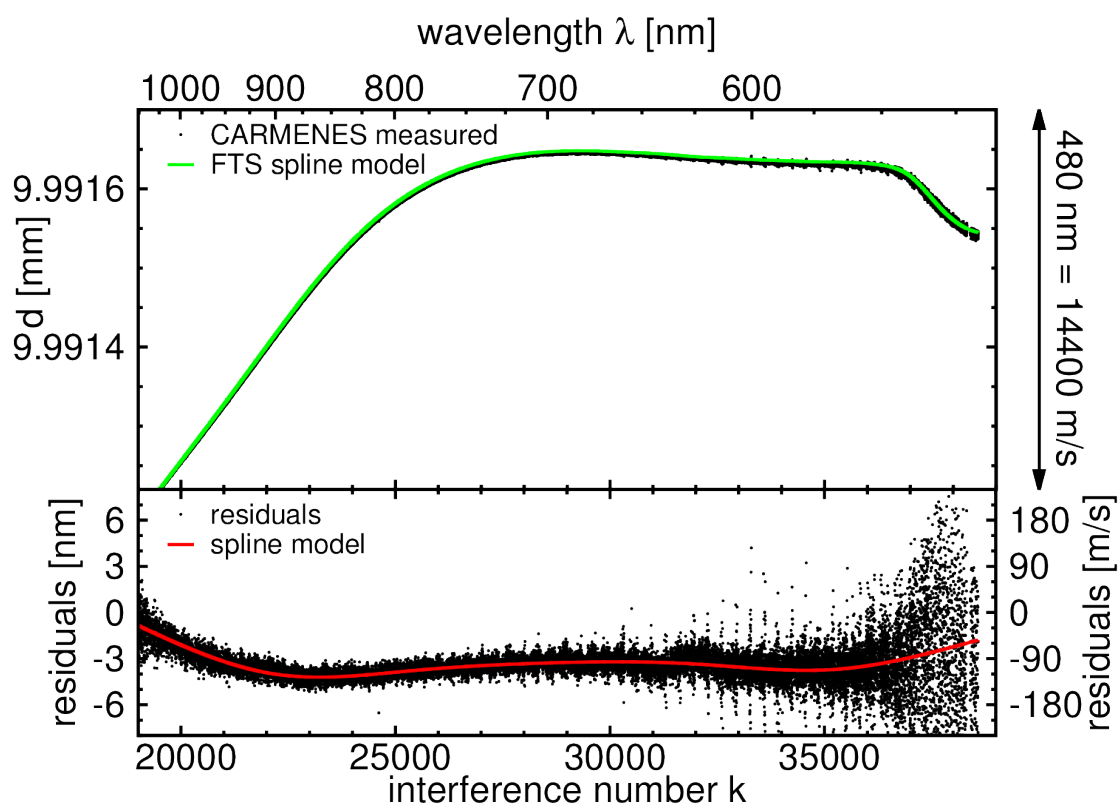


FIGURE 2.27: Cavity width of the visual FPI measured with CARMENES VIS. Upper panel: cavity width measured from individual FPI lines (black dots). B-spline model from the FTS measurements of the visual FPI in the optics laboratory in Göttingen (green line). Lower panel: residuals of the FTS b-spline model (black dots). Low order b-spline fit to the residuals (red line).

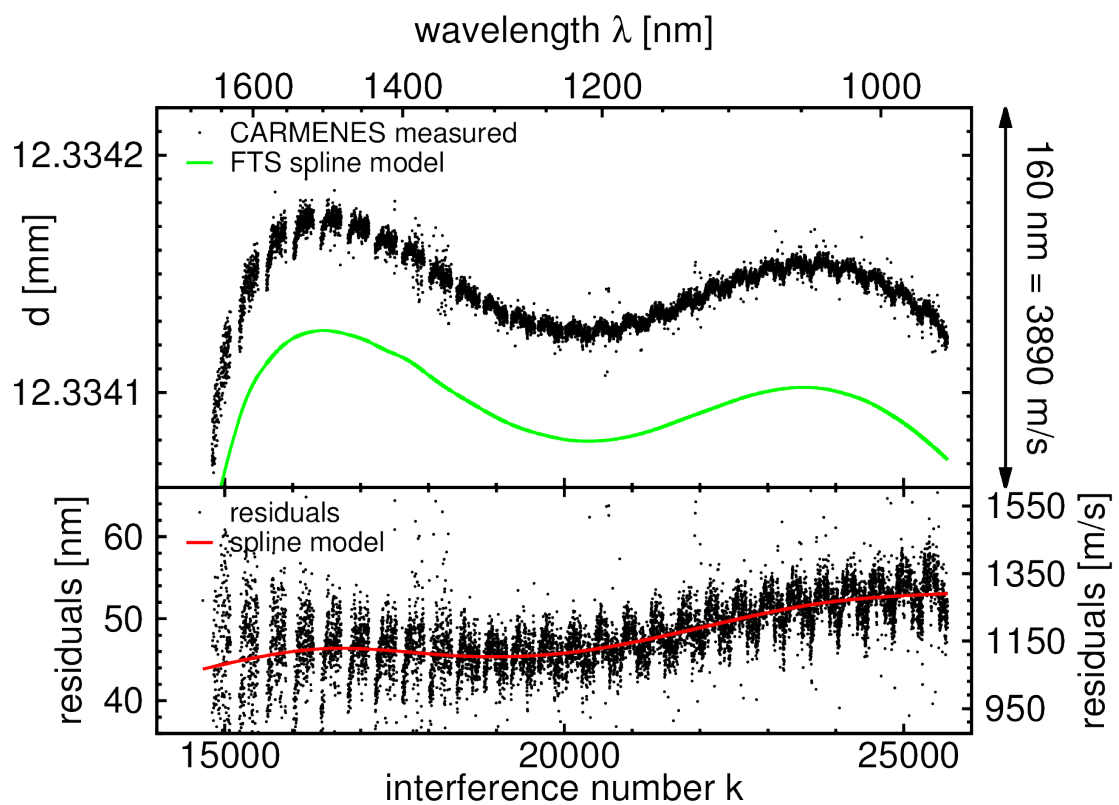


FIGURE 2.28: Cavity width of the near-infrared FPI measured with CARMENES NIR. Upper panel: cavity width measured from individual FPI lines (black dots). B-spline model from the FTS measurements of the near-infrared FPI in the optics laboratory in Göttingen (green line). Lower panel: residuals of the FTS b-spline model (black dots). Low order b-spline fit to the residuals (red line).

near-infrared FPI can be seen in Fig. 2.28. For the near-infrared FPI, the mean offset between the CARMENES  $d$  measurements and the in FTS  $d$  model is 48.5 nm or 1179 m/s. During commissioning of CARMENES the fiber of the near-infrared FPI broke. The fiber was changed and the FPI had to be realigned. The large change in the effective cavity width  $d$  is thus a result of a different illumination angle of the FPI after the system was opened. Like in the visual FPI, the residuals of the near-infrared FPI shows wavelength-dependent residuals (see lower panel of Fig. 2.28). The amplitude of this wavelength-dependent pattern is 8 nm or 210 m/s. Like for the visual FPI, the alignment change is suspected to be the cause of the wavelength-dependent residuals.

To include the FPIs in the wavelength solution of CARMENES the FTS model of the effective cavity width  $d$  measured in Göttingen is used. The FTS models for the visual and near-infrared FPI describe the global  $d$  curve well, but both models have to be adjusted to account for the wavelength dependent differences observed after shipment of the FPIs to Calar Alto. Instead of fitting the entire effective cavity width  $d$  with 33 spline parameters in the visual and 45 spline parameters in the near-infrared, a low order b-spline with 7 parameters is enough to adjust the FTS model for the wavelength-dependent differences observed after the FPIs were moved to Calar Alto (see red lines in the lower panels of Fig. 2.27 and Fig. 2.28). Thus, the high-precision measurements taken in Göttingen can be combined with the data of CARMENES to reduce the number of free parameters. The wavelengths of all FPI peaks are then computed from this adjusted model of the effective cavity width  $d$  and used in the following to find the wavelength solution of CARMENES.

#### 2.4.8 Systematics in the wavelength solution of CARMENES

As already demonstrated with HARPS, FPIs can improve the wavelength solution of echelle spectrographs and thereby potentially improve the RV precision of the instrument, see Sect. 2.3.4. In this section, details about the wavelength solution approaches for CARMENES VIS and NIR are given and the wavelength solutions computed with HCLs and with FPIs are compared.

For CARMENES, the wavelength calibration procedure follows three basic steps:

- A first guess wavelength solution only from the HCLs is computed.
- The FPIs are anchored to the HCLs by using the first guess wavelength solution to compute and model the effective cavity width  $d$  (see Sect. 2.4.7)
- The final wavelength solution is then calculated based on the HCLs and the FPI lines.

In the following each of these three wavelength solution steps are described.

For CARMENES VIS, the first guess wavelength solution is done with a smooth 2D polynomial reverse regression model. This polynomial has 7 orders in the dispersion direction, 6 orders in the cross-dispersion direction, and allows for cross terms up to 7<sup>th</sup> order (e.g.  $x^3 \cdot y^4$ ). In total, 35 free parameters are needed to describe the wavelength range from 513 to 1063 nm in 61 orders of CARMENES VIS. Combining the spectral lines of the UNe, UAr and ThNe lamps, and clipping outliers deviating more than 250 m/s from the wavelength solution, leaves approximately 5000 calibration lines for CARMENES VIS. However, one has to keep in mind that two of the lamps are uranium cathode lamps which means that most of their lines are located on the same detector pixel. Therefore, uranium lines appearing in both lamps do not add much information

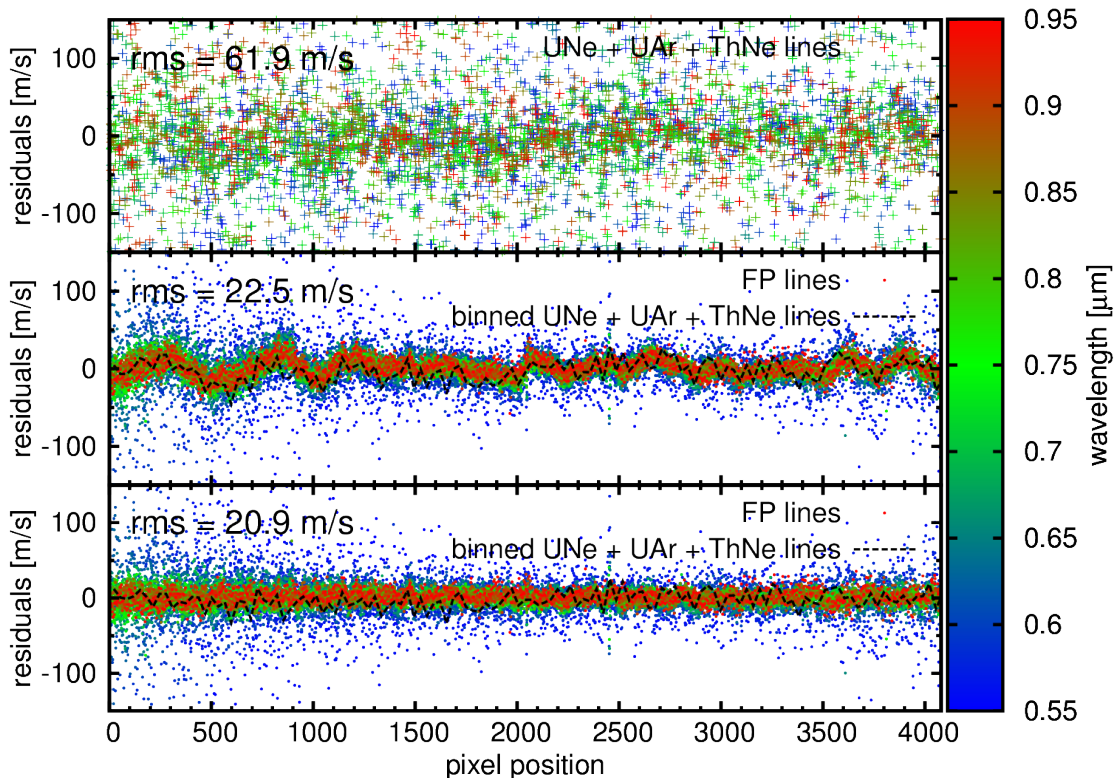


FIGURE 2.29: Residuals of the CARMENES VIS wavelength solution. Color coded is the wavelength of the lines used to compute the wavelength solution. Upper panel: residuals of the HCL lines when using the first guess wavelength solution. Middle panel: residuals of the FPI lines when using the first guess wavelength solution (colored points) and binned HCL lines (black broken line). Lower panel: residuals of the FPI lines when using the final wavelength solution model for CARMENES VIS (colored points) and binned HCL lines (black broken line).

in the calibration process, as for the wavelength solution it is beneficial if calibration lines are well distributed.

The upper panel of Fig. 2.29 presents the residuals of the first guess wavelength solution for CARMENES VIS using only the HCLs. The color code of the points corresponds to wavelength of the lines. In CARMENES VIS, an rms scatter around the smooth polynomial model on the order of 60 m/s is found. Residuals on this order of magnitude are common even in high-precision RV machines, and the numbers obtained with CARMENES compare well with what is found with HARPS. As possible CCD imperfections or other systematics have amplitudes below the rms scatter of the HCL lines, they are not revealed. Hence the simple but fast reverse regression is sufficient to start with.

For CARMENES NIR, a more advanced model for the first guess wavelength solution has to be used. The near-infrared arm of CARMENES uses two 2040 x 2040 pixel Hawaii-2RG detectors. These two detectors are separated by a physical gap of about 140 pixels. However, the gap is not constant in width because the two detectors are tilted with respect to each other so that the physical separation narrows to 139.6 from 141.8 pixels from blue to red orders. To compute the first guess wavelength solution an inverse regression approach (see Sect. 2.3.3) is used. The gap between the two detectors is modeled with 2040 virtual pixels (one per row). However, the size of each

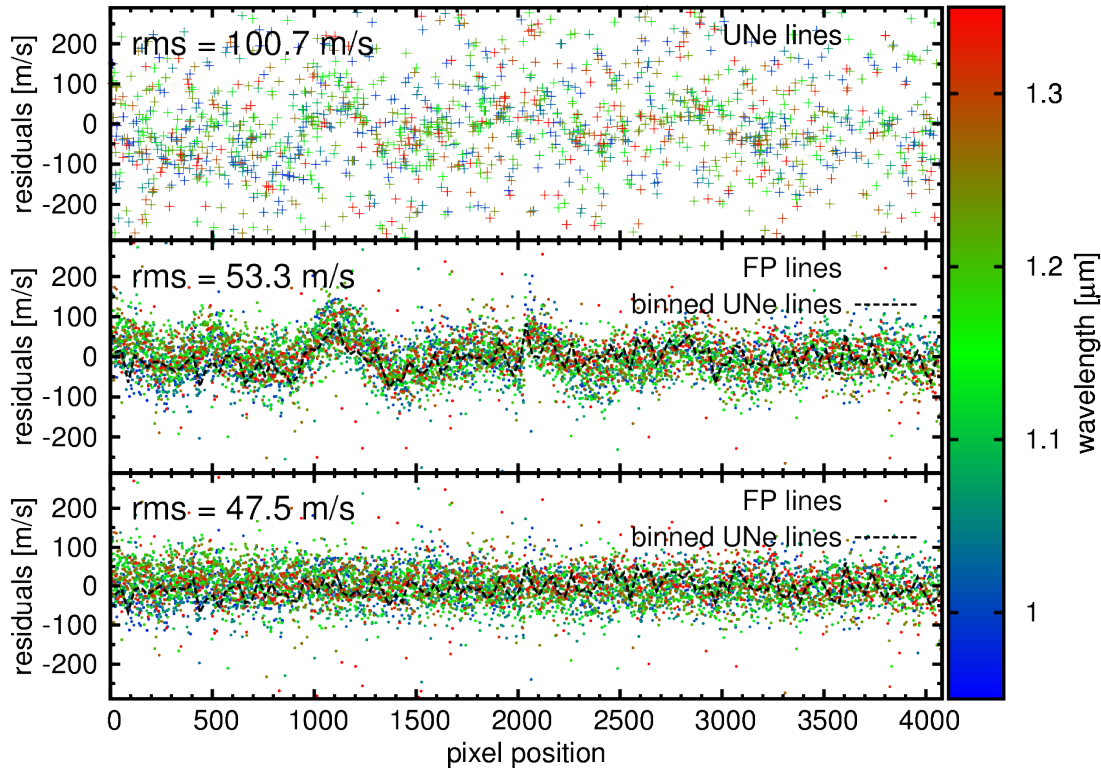


FIGURE 2.30: Residuals of the CARMENES NIR wavelength solution. Color coded is the wavelength of the lines used to compute the wavelength solution. The wavelength range plotted corresponds to the wavelength range used to compute near-infrared RVs. Upper panel: residuals of the uranium lines when using the first guess wavelength solution. Middle panel: residuals of the FPI lines when using the first guess wavelength solution (colored points) and binned uranium lines (black broken line). Lower panel: residuals of the FPI lines when using the final wavelength solution model for CARMENES NIR (colored points) and binned uranium lines (black broken line).

virtual pixel is computed by a linear interpolation between the lower most and upper most detector pixel so that the gap is effectively described by only two parameters. After using the virtual pixels to remap the detector pixels  $X$  to focal plane coordinates  $x$  a smooth 2D polynomial is fitted to the data. The model has 4 orders in dispersion direction, 5 orders in cross dispersion direction and allows for 4<sup>th</sup> order cross terms. In total 15 + 2 (polynomial + gap) free parameters are needed for the wavelengths from 960 to 1712 nm in 28 orders. The UNe lamp for CARMENES NIR provides 2056 lines for the first guess wavelength solution after clipping down to a maximum deviation from the model of 450 m/s. The residuals of the first guess wavelength solution of CARMENES NIR can be seen in Fig. 2.30. Also in Fig. 2.30 the color code corresponds to wavelength of the lines. The wavelength range plotted corresponds to the range specified by the CARMENES consortium to measure RVs for stars. For CARMENES NIR, an rms scatter around the fit of about 100 m/s is measured.

From experience with HARPS, it is known that systematics with amplitudes of a few 10 m/s go unnoticed with the low number of HCL lines used to compute the first guess wavelength solution, see Sect. 2.3.4. To reveal hidden systematics in the wavelength solutions of CARMENES VIS and NIR, the calibrated FPIs (see Sect 2.4.7)



are used. The first guess wavelength solution obtained only from the HCLs is subtracted from the calibrated FPI wavelengths. These FPI residuals are plotted as color coded points in the middle panel of Fig. 2.29 for CARMENES VIS and in Fig. 2.30 for CARMENES NIR.

Unlike the HCL lines the residuals of the FPI lines show a non random distribution in both spectrograph arms. The residuals found in CARMENES are not similar to what was found in HARPS. In CARMENES VIS distinct sharp discontinuities every 512 pixels caused by the pixel block stitching were expected (see Sect. 2.4.2 for a description of the CARMENES VIS CCD). However, no signs of this imperfect block stitching are found. This is probably related to a hardware change in the CCD production. The stepper machines putting together the pixel blocks have improved in precision (e2V; private communication) so that discontinuities caused by the block stitching effect are well below our detection abilities and thus should not be of any concern for RV measurements with CARMENES VIS. The only discontinuity found in CARMENES VIS is located at 2048 pixels and is not caused by pixel block stitching. Instead this discontinuity is a result of the readout mode with four amplifiers and has an amplitude of about 12 m/s. Also in CARMENES NIR, a discontinuity at 2040 pixels is seen. Because there are not many HCL lines available to calibrate the near-infrared arm, the uncertainties of the virtual pixel sizes describing the gap between the detectors is large. With the FPI data, the residuals of the virtual pixel sizes that were determined slightly wrong in the first guess wavelength solution are seen.

Despite the missing block stitching discontinuities, there is a second unexpected feature visible in the wavelength solution residuals of the FPI. In both CARMENES channels a similar wave-like pattern is observed in the FPI residuals. To rule out that these wave-like structures are intrinsic to the FPIs, the HCL residuals are binned in 32 pixel wide chunks and checked for a similar pattern. The binned HCL residuals are plotted as black broken lines in Fig. 2.29 and Fig. 2.30. Although the binned HCL data points are still noisy, the same pattern of systematics as in the FPI residuals of both CARMENES channels is visible. Observing the same pattern with four different sources (UNe, UAr, ThNe and the FPI) in CARMENES VIS and two sources (UNe and FPI) in CARMENES NIR leads to the conclusion that the wave-like structures in the residuals originate from the spectrograph.

It is important to note that in each spectrograph, the shape of the systematics is the same for all orders. Thus, the effect is either a function of the pixel grid or a function of the diffraction angle from the grating. As the wave-like systematics are observed in both spectrographs but their detectors vastly differ, a reason other than pixel size variations is suspected as the cause of the problematic pattern. Both spectrographs use a copy of the same grating and have a similar optical layout. Observing a similar pattern in both CARMENES channels is thus strong evidence for the systematics being intrinsic to the spectrograph optics. However, it is difficult to find the exact cause of the observed systematics. A plausible explanation could be self-contamination of the spectra. In this theory a portion of the light is reflected by the camera to the entrance slit and back to produce a weak, slightly displaced copy of the original spectrum. The superposition of the two spectra in the focal plane leads to line asymmetries and thus to apparent systematic shifts in wavelengths of lines.

Because the wave-like pattern is observed with several calibration sources, it is very likely to also affect the stellar spectra. This issue with the wavelength solution can limit the RV precision of CARMENES. Thus the systematics in the wavelength solution are corrected by allowing for small perturbations in the smooth polynomial that is fitted to the focal plane coordinates  $x$ . The final wavelength solution for both CARMENES VIS

and NIR is an inverse regression. Instead of only fitting a 2D polynomial after pixel coordinate transformation, the wave-like perturbations are included as a 1D b-spline with 50 nodes that is added to the polynomial. The wavelength solution model can be written as  $x(\lambda, o) = [\text{poly}(\lambda o, o) + \text{spline}(\lambda o)]^{-1}$ . For CARMENES VIS one discontinuity is used to fit the amplifier jump at 2048 pixels. For CARMENES NIR the linear gap model from the first guess wavelength solution is also used for the final solution. The FPI residuals for the model including the wave-like distortions are plotted in the lower panel of Fig. 2.29 for CARMENES VIS and in Fig. 2.30 for CARMENES NIR. It can be seen that a 1D distortion model is enough to correct for the wavelike residuals in both spectrographs. After applying the new distortion method, the rms scatter around the final wavelength solution is about 21 m/s for CARMENES VIS and 48 m/s for CARMENES NIR. Like in HARPS, the FPIs of CARMENES demonstrate to be beneficial for wavelength calibration of high-precision optical and near-infrared RV instruments because low amplitude systematics can be discovered and corrected.

## 2.5 Summary

To conclude this chapter the most important points are summarized here. For wavelength calibration the first step is always to determine the line positions of calibration lines on the detector. This requires modeling the line shape. What is observed with a detector in a spectrograph is not directly the LSF of the spectrograph but a convolution with the pixel sensitivity function which is generally not known. Hence what is observed instead with a pixel detector is the LSF integrated over the wavelength range covered by one pixel. We call this the effective line spread function (eLSF). The eLSF can be complex and analytical functions might be inaccurate models in some cases. To determine better models for the eLSF, a method called eLSF reconstruction is proposed. Every calibration line samples the eLSF at different positions because the line centers are located at random positions within pixels. This means there is information about the eLSF also between the discrete pixel values if all calibration lines are used simultaneously. An iterative method to find both line positions and the eLSF with high accuracy is proposed. First all calibration lines are remapped to a common zero point by subtracting a first guess pixel location from all lines. These stacked calibration lines are then modeled by a smooth b-spline function and new, better line positions can be measured.

To test the approach of eLSF reconstruction, simulations of calibration spectra of an LFC and an FPI for HARPS and CARMENES are created. As long as the eLSF is symmetric, analytic models are sufficient to fit line positions with precisions of a few m/s in most cases. However, small inaccuracies in the eLSF can lead to pixel phase errors or produce large errors in narrow lines close to undersampling. With the eLSF reconstruction method, accurate models for the line shapes are obtained so that phase errors and narrow sampling become less problematic. The simulations show that for current state-of-the-art instruments and calibration sources, traditional line fitting methods are sufficient. Nevertheless, instruments with complex LSFs due to sliced fibers, or future instruments that require higher calibration precisions, can benefit from the eLSF reconstruction method.

When calibrating echelle spectrographs, HCLs have been the common standard for decades. However the limited number of lines, uneven line spacing, and saturation problems have triggered the development of new calibration sources. LFCs are



promising machines but their complexity, price tag and a few unresolved issues postpone their use as a standard calibration source. FPIs offer a dense grid of calibration lines similar to the LFC. However, the wavelengths of FPI lines are only determined by the cavity width of the device which is not known to the precision needed for wavelength calibration. In addition, the cavity width is not constant but a (smooth) function of wavelength so that FPI lines are not completely equidistantly spaced. With our method, the cavity width and its wavelength dependent variation is modeled and the FPI lines are anchored with absolute standards like a HCLs. The global wavelength scale of a spectrograph will thus still be determined by the absolute standard but the numerous FPI lines improves the wavelength solution on small and medium scales.

The many calibration lines offered by the FPI allow for more detailed models of the wavelength solution of echelle spectrographs. Instead of fitting with reverse regression  $\lambda(x)$ , a direct regression approach  $x(\lambda)$  is suggested. Direct regression is motivated by the fact that the pixel position is the measured quantity while the wavelength of calibration lines is the independent variable. The direct regression model is a 2D polynomial instead of the order by order approach commonly used (e.g. in the HARPS DRS pipeline). Using a 2D polynomial reduces the amount of parameters needed to describe the wavelength solution and enhances the robustness of the fit. It is found that parameterizing the 2D polynomial according to the grating equation  $x(\lambda, o) = poly(o\lambda, o)$  works better than using  $x(\lambda, o) = poly(\lambda, o)$ . Although the wavelength solution of the spectrograph is thought to be a smooth function in the focal plane of the instrument, pixel size variations can lead to discontinuities in the observed apparent wavelength solution. In order to model discontinuities caused by the detector, a coordinate transformation is implemented that maps detector coordinates  $X$  to focal plane coordinates  $x$  before direct regression is performed.

When using the FPI of HARPS for wavelength calibration, pixel size variations of up to 5.8 percent are found in CCD. These results compare well to what was found earlier by the experimental LFC setup in La Silla. When the HARPS FPI wavelength solution is directly compared to the wavelength solution provided by the LFC, only small differences with an rms of 8 m/s are found.

The impact of a distorted wavelength solution can be estimated by simulating observations of a perfectly quiet star over one year. Distortions in the wavelength solution are found to lead to false signals and additional noise in the RV measurement on the order of 1 m/s. Hence a progress towards higher RV precisions with the next generation of instruments requires better calibration sources. As shown with HARPS, FPIs can be considered as one possibility to achieve better wavelength calibration.

Because of the progress made with the HARPS FPI, CARMENES uses the FPIs originally foreseen only for drift measurements also in daily calibration. To anchor the FPI in the visible channel of CARMENES, a ThNe, UNe and UAr lamp are used. In the near-infrared a UNe lamp is the only absolute standard. To maximize the number of calibration lines offered by the HCLs, tests were performed with each lamp. The operational current of all lamps is 12 mA. For CARMENES VIS, the best exposure time for all lamps was found to be 1 s, while for CARMENES NIR, the UNe lamp is exposed for 3 s, which is the minimum exposure time possible for the CARMENES NIR detector. To further enhance the line ratio between buffer gas and metal lines, the fiber picking up the light from the lamps is slightly displaced from the center of the focal plane. This places the fiber more into the image of the glowing cathode ring and enhances the metal to gas line ratio by about 20 percent.

The FPIs used by CARMENES were extensively tested in the optics laboratory in Göttingen using a high-resolution FTS. Data of this kind was not available for the

HARPS FPI and it gives us the opportunity to compare two independent measurements of the cavity width of the FPIs done with two instruments. It is found that the cavity width of both CARMENES FPIs has changed after moving the equipment from Göttingen to Calar Alto. We suspect that mechanical changes are the reason for the slight changes in the effective cavity width curve. Nevertheless, it is possible to use a slightly modified version of the cavity width model based on FTS measurements to describe the cavity width observed with CARMENES. Hence the number of free parameters required to model the wavelength dependent cavity width can be reduced from 33 for the visual FPI and 45 for the near-infrared FPI to just 7 parameters in both cases. Anchoring the cavity width model with FTS measurements further improves the stability of the wavelength solution.

Using the FPIs in daily calibration of CARMENES has revealed a number of systematics in the wavelength solution that is only based on the HCLs. In contrast to HARPS, however, no prominent pixel size variations are found in the residuals of the wavelength solution because the stepper machines used to produce CCDs seem to have improved. Instead, a wave-like pattern is observed in the wavelength solution residuals. In order to rule out the FPIs themselves as cause of the systematics, the HCL residuals are checked for similar patterns. Binning the HCL residuals results in the same systematics, which means that the problem does not originate from the FPIs. Further similar systematics are found in both CARMENES spectrographs which is a hint for possible issues in the optical design of the instruments. As both spectrographs have a very similar layout, the systematics are suspected to be self-contamination of the spectra caused by reflections between the camera and the slit.

A new wavelength solution model is implemented, capable of taking into account the wave-like systematics and correcting them. For CARMENES VIS, this model achieves an rms of 21 m/s when the FPI is used. In CARMENES NIR, the rms around the final wavelength solution is 48 m/s. Hence, using FPIs in daily calibration of CARMENES has improved the wavelength solution. FPIs are successfully used already in two high-precision RV instruments and are foreseen for many upcoming instruments. Thus FPIs can be considered a valuable investment and economical alternative to the LFC in small and medium class RV projects.

## 2.6 Future outlook for further improvements

Although this thesis demonstrates that FPIs have the potential to improve the wavelength solution of high-precision RV instruments, there are still a few open questions concerning the long term stability of the FPIs or the stability of the wavelength solution. Because the CARMENES spectrographs are the first to use FPIs in daily calibration, these open issues can be addressed in the future once enough data is collected to draw definite conclusions.

The first issue that should be addressed is the stability of the effective cavity width  $d$ . So far there have been no efforts to track changes in the wavelength dependent properties of the coating. Up to now it is not clear whether the cavity width curve  $d$  is stable over time or if it is susceptible to changes because of aging. A stable cavity width curve  $d$  would mean that the number of free parameters to model the cavity width could be reduced only to a global length change caused by thermal expansion or a change in refractive index of the medium between the two mirrors. Reducing the number of free parameters to fit the FPI cavity width could possibly result in a further

improvement in the stability of the wavelength solution and thus improve the precision of CARMENES.

Systematics in the wavelength solution of CARMENES are of great concern for the RV precision of the instrument. Up to now the wavelike pattern observed in the FPI residuals is fitted night by night. An investigation of the stability of this pattern over time could help to further stabilize the wavelength solution of CARMENES. If the wavelike pattern turns out to be stable over time, a precise map of these systematics can be determined and used in daily calibration. This would greatly reduce the number of free parameters in the wavelength solution and reduce uncertainties introduced by the complex model.

The origin of the complex systematics in the wavelength solution of CARMENES is not totally clear yet. As a result the question arises if these systematics should be taken into account in the wavelength solution at all. Computing RVs of a stable star with a different wavelength solution approaches can help clarifying this question. However, performing such an RV test requires a large number of spectra of the same star being taken over a long period of time. As CARMENES is finishing its first survey year by the end of 2016, enough data of a few known RV stable stars should be collected soon to address this critical question.



## Chapter 3

# Simulating active stars

The following chapter is accepted for publication in the *Astronomy and Astrophysics* journal as 'The influence of convective blueshift on radial velocities of F, G, and K stars' (Bauer et al., 2017 in press). The permission for republication in this thesis was granted by A&A.

### 3.1 Introduction

For about two decades the RV method has been successfully used to detect extrasolar planets. Since the first planet detection by Mayor & Queloz (1995), the RV technique has undergone major improvements (Latham et al., 1998; Fischer et al., 2016). Current instruments like HARPS (Mayor et al., 2003) can reach a precision of a few m/s in the optical. Several instruments like CARMENES (Quirrenbach et al., 2011), SPIROU (Artigau et al., 2011), IRD (Kotani et al., 2014), and HPF (Mahadevan et al., 2010) aim to extend this precision to the near-infrared. Furthermore, ESPRESSO (D'Odorico & CODEX/ESPRESSO Team, 2007) aims to achieve an unprecedented precision of 20 cm/s at optical wavelengths. With this next generation of RV machines, instrumental noise will no longer be the limiting factor in detecting exoplanets similar to Earth. However, astrophysical noise sources such as stellar activity will become a concern and hinder planet detections (Fischer et al., 2016).

Current exoplanet surveys focus on solar-like stars of spectral type F, G, K, and M to detect rocky planets in the habitable zone of their host stars. Unfortunately these stars exhibit a variety of intrinsic signals as well which are often referred to as stellar noise or jitter in the exoplanet community. Asteroseismic oscillations and granulation cause signals on the timescales of minutes to days (Butler et al., 2004; Dumusque et al., 2011; Meunier et al., 2015), and magnetic activity, manifested as spots, produces signals associated with the rotation period of the stars (Radick et al., 1983; Benedict et al., 1993) in the range of hours to months. Because of the similar timescales, the signals of small spots are often hard to disentangle from planetary companions in RV searches. This has led to several examples in the literature where proposed planets have been identified later as stellar activity or are still under debate today (e.g., Udry et al., 2007; Vogt et al., 2010; Robertson et al., 2014; Anglada-Escudé et al., 2014; Robertson et al., 2015; Anglada-Escudé et al., 2015).

Hence, new tools are needed to understand and possibly correct for activity signals of the host star to be able to detect small exoplanets reliably in the future. In this context, many attempts have been made to model the activity modulation of RV curves from stellar spots. All of them have in common that the RVs are estimated from the disk integrated spectrum which is constructed from a grid of quiet and spotted regions. The spectra of quiet and active regions differ from each other because of temperature and magnetic field, and therefore the choice of the model spectra influences the resulting

RVs. As an overview we summarize here the approaches that have been used by other authors in their studies of activity-related RV simulations.

Saar & Donahue (1997) started to simulate the spot induced RV jitter by measuring the apparent line shift of atmospheric models of a single FeI line around 6000 Å. They used a solar-like star with an effective temperature of  $T_{\text{eff}} = 5750$  K and set the spot temperature to 0 K. In this way the spot is totally dark and no spot spectrum has to be considered. Hatzes (2002) used a similar technique to derive RV amplitudes due to spots. In his case the spot was 1200 K cooler than the quiet stellar photosphere which corresponds to the mean between temperatures measured in the umbra and penumbra of the Sun (Berdyugina, 2005). As in Hatzes (2002), spots are not completely dark; they also contribute to the disk-integrated stellar spectra. Although the RV amplitudes should be sensitive to spot temperatures, the results in Hatzes (2002) and Saar & Donahue (1997) are seemingly consistent. As a next step, Desort et al. (2007) used the full spectral range of HARPS to simulate RV amplitudes of dark spots for an F, G, and K star using Kurucz (1993) models. Desort et al. (2007) showed that using different single spectral lines can lead to different spot RV amplitudes. Thus the consistency in the spot RV amplitudes despite the spot temperature difference between Hatzes (2002) and Saar & Donahue (1997) could be the result of different lines used in both works. In order to use broader spectral regions while saving computational time when computing spot RVs, Boisse et al. (2012) suggested to directly use a Gaussian cross correlation function (CCF) instead of a full spectrum in each grid point. However, all of the above mentioned works have been done taking into account only the contrast of active regions. Lagrange et al. (2010), Meunier et al. (2010), Lanza et al. (2010), Jeffers et al. (2014), Dumusque et al. (2014) and Borgniet et al. (2015) realized the importance of the convective blueshift for activity-induced RVs: as magnetic fields hinder convection inside active regions, spectra of spots are expected to appear redshifted compared to the surrounding disk. This apparent redshift of the active region changes the RV signature of spots and gives rise to stronger RV amplitudes of bright plages compared to simulations only taking into account the flux effect.

Nevertheless, measuring the convective blueshift proves to be difficult even for the Sun. Absolute values of the convective blueshift of the Sun differ significantly throughout the literature (e.g., 200 m/s in Meunier et al., 2010; or 500 m/s in Lanza et al., 2010). This large range of convective blueshift values found in the literature can at least partly be attributed to the fact that the measured convective blueshift depends on the depth of the lines used to derive it (Gray, 2009; Meunier et al., 2017). As the convective blueshift is even difficult to measure in the Sun, only recently has there been first attempts to measure convective blueshifts in other stars (e.g., Meunier et al., 2017). The best guess of the convective blueshift in stars we have today therefore comes from simulations (e.g., Beeck et al., 2013b). Convection patterns might be different in other stars and extrapolation of activity models based on solar parameters might lead to under- or overestimation of activity effects. With the help of simulated line profiles we can explore which differences arise in the RV curves of spots and plages in stars other than the Sun. Hence in this work we investigate the effects of convective blueshift on the RV curves of spots and plages not only in the Sun but also in other stars. Instead of relying on measured spectra, we chose to use magneto-hydrodynamic (MHD) simulations (Beeck et al., 2013a,b) where convective blueshift can be simulated for the Sun and other stars.

While so far only the projection effect of the convective blueshift has been taken into account, Herrero et al. (2016) also included a more complex variation of the convective blueshift with  $\mu$  (where  $\mu = \cos \theta$  with  $\theta$  being the angle between the line of sight and the normal to the stellar surface, which is what we call the limb angle) by using

simulation results from [Ludwig et al. \(2009\)](#). By using the line profiles of [Beeck et al. \(2013a,b\)](#) we automatically include variations of convective blueshift in a consistent way.

The magnetic fields that hinder convection in active regions and lead to the importance of including convective effects in the simulations of activity RV jitter also alter the line profiles in spots and plages (e.g., by Zeeman broadening; [Reiners et al., 2013](#); [Reiners, 2014](#)). Strong magnetic fields present in active regions can be taken into account by using observed spectra of the quiet photosphere and a sunspot as done by [Dumusque et al. \(2014\)](#). While this approach might hold for G-type stars, expanding the simulations to fit stars of spectral types F, K, or M might not work well. The line profile simulations of [Beeck et al. \(2013a,b\)](#) enable us to include the line profile altering effects of the magnetic field into our activity simulation. In this work we also investigate the influence that the line profiles in active regions have on the resulting RV curves.

In this work we investigate activity RV curves from spots and plages on stars other than the Sun. In Sect. 3.2 we describe the fundamental effects causing active regions to produce apparent RV jitter in stars. In Sect. 3.3 we present our model including the properties of the simulated stars. In Sect. 3.4 we show how magnetic fields, convective blueshift, line profiles and spot temperatures influence the RV signals of spots and plages. The bisector span has become an important parameter in exoplanet studies during the last years. However, it is beyond the scope of this work to investigate bisectors in detail, which will therefore be done in a separate article.

## 3.2 Effect of active regions on stellar RVs

When active regions appear on the stellar disk they can cause apparent variations in RV measurements of the star. The parameters of the active region play an important role in determining which signatures are observed. In this section we summarize the most important effects influencing the RV curves of active regions.

### 3.2.1 Temperature

On the Sun we can observe active regions as dark spots or bright faculae and plages. Spots are regions in which a strong magnetic field hinders the plasma moving perpendicular to the field lines. This suppresses the supply of new, hot material into the active region and causes a cool down in the vicinity of a strong magnetic field. On the Sun a dark spot is about 500 – 1700 K cooler than the rest of the stellar surface ([Lagrange et al., 2010](#); [Berdyugina, 2005](#)).

If a dark spot rotates into view it disturbs the flux balance between the approaching part of the disk that rotates towards the observer and the receding part of the disk that rotates away from the observer. As a result more light from the receding part of the stellar disk reaches the observer when the spot first comes into view resulting in line profile distortions (e.g., [Reiners et al., 2010](#)) and an apparent redshift of the disk integrated spectral lines. When the spot proceeds with stellar rotation across the disk it blocks light from areas with different projected rotational velocities. The result is a variation of the RV during one rotation period of the star. This effect is commonly referred to as the flux effect which produces symmetric, sinusoidal RV curves (while the spot is visible and if the size of the spot is constant throughout its lifetime). The upper



panel of Fig. 3.1<sup>1</sup> indicates a schematic path of a dark active region on the equator of the stellar disk and the middle panel shows the characteristic RV variations caused by the dark spot on an active star.

Compared to dark spots, bright faculae and plages behave differently. As explained in Foukal (2008) faculae are conglomerates of narrow flux tubes with diameters in the order of 100 km. These flux tubes are much thinner than spots and the strong magnetic field inside sustains it against the outside gas pressure. The walls of the flux tube are hot and radiation is entering the tube horizontally. If the tube is thin enough the horizontal flux reaches the center and the interior of the tube is heated which results in an increase of vertical flux turning the tube bright (Ortiz et al., 2002). Plages show a significant limb brightening effect. While the disk dims towards the limb (at small  $\mu$ ), the hot walls of the plages seen by the observer increase the plages contrast. Meunier et al. (2010) published an empirical law for the temperature of plages on the Sun as a function of  $\mu$ :

$$\delta T_p = 250.9 - 407.7\mu + 190.9\mu^2 [\text{K}]. \quad (3.1)$$

Because of the smaller temperature difference between quiet stellar photosphere and active region, plages have a much lower contrast than dark spots and consequently the flux effect expected from faculae and plages leads to lower RV amplitudes. In the lower panel of Fig. 3.1<sup>1</sup> the RV curve due to the flux effect of an equatorial plage with the same size as the dark spot can be seen. Because the plage is bright its RV curve shows a reversed behavior as compared to the dark spot. Nevertheless, RV amplitudes caused by the flux effect of plages are much lower than for dark spots in the scenario shown here. Although in our similar size scenario plages produce lower RV signals than spots, plages occupy larger surface areas on the Sun than spots do. Therefore the total RV amplitudes from plages are similar to those of spots (Meunier et al., 2010).

The very different RV amplitudes of spots and plages caused by the flux effect show the importance of active region contrast. While the dark spot is 550 K cooler than the rest of the disk, the bright plage is only 250 K hotter than the quiet photosphere at maximum. The contrast of active regions can be estimated by dividing the Planck curves from spot and plage by the Planck curve of the quiet photosphere (this is a simple estimate without the line variations studied later in Sect. 3.4.3). At a visible wavelength of 6200 this results in a spot contrast of 0.68 and a plage contrast of 1.18 for the Sun. Hence, larger temperature differences,  $\Delta T$ , between active region and quiet photosphere lead to a larger contrast and thereby to larger flux effect in the RV signal. However, spot temperatures are not necessarily the same in all stars which gives rise to the assumption that RV curves of spotted stars can be very different depending on the active region contrast. Therefore, we investigate the behavior of spot RV curves with changing spot contrast in Sect. 3.4.4 in more detail.

### 3.2.2 Convection

Convection is present in the outer layers of late-type stars of spectral types F, G, K, and M. Uprising material in the outer convective layers of these stars is hotter, brighter, and covers a larger surface area than the down-flowing material. Hence, lines formed in the

<sup>1</sup>The RV curves in Figs. 3.1 and 3.2 are derived with our method explained in Sect. 3.3. The spot parameters are similar to what is observed on the Sun. The spot covers 1 % of the visible disk and is 550 K cooler than the surrounding photosphere. The plage has the same size as the spot but its temperature is derived from Eq. 3.1. The star rotates with a projected velocity of  $v \cdot \sin(i) = 2$  km/s. The line profile used is a Voigt profile and the value for the convective blueshift in Fig. 3.2 is constant at 300 m/s.

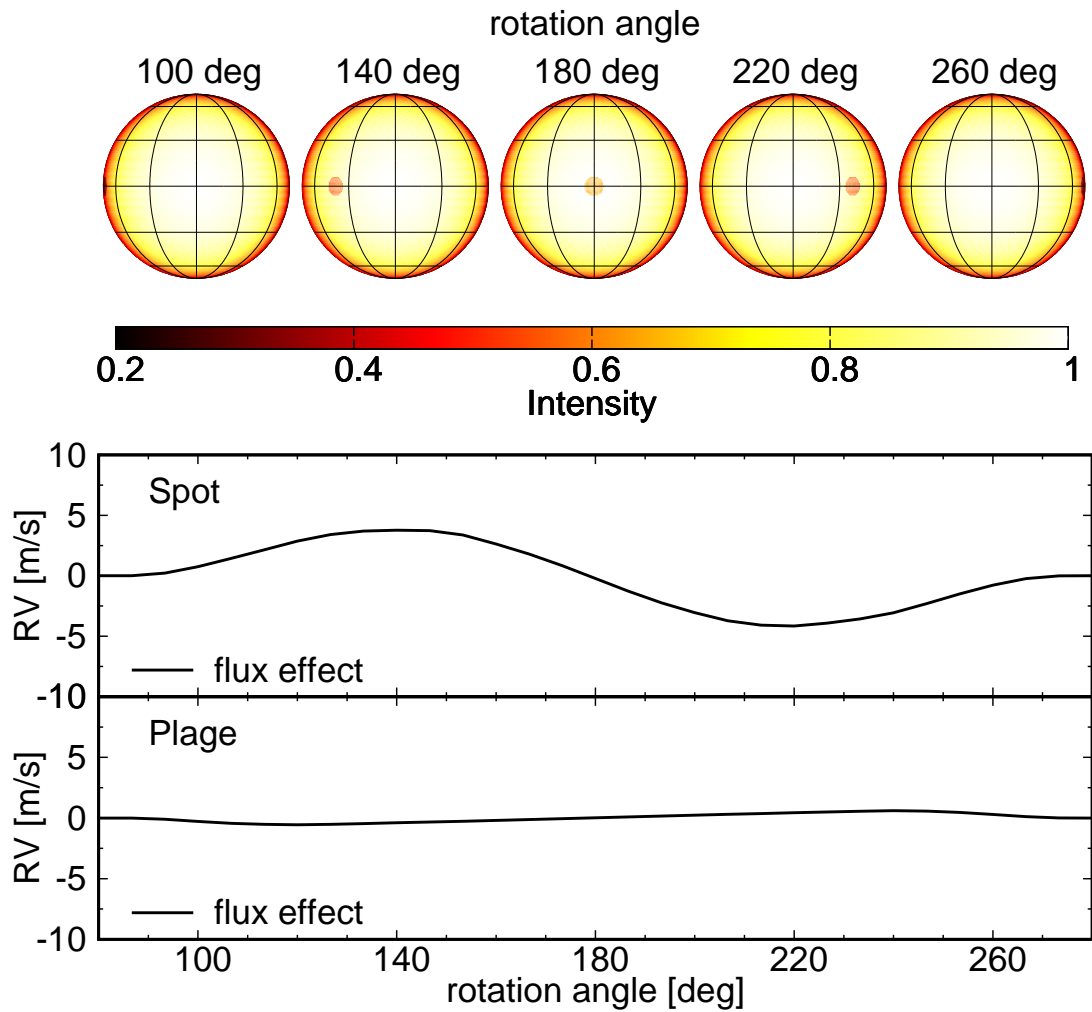


FIGURE 3.1: Upper panel: Schematic view of the G2 star with a dark spot rotating across the projected stellar disk. The color code indicates the intensity across the disk. Middle panel: RV curve produced by the flux effect of a dark spot. Lower panel: RV curve produced by the flux effect of a bright plage. The stellar effective temperature in this models is  $T_{eff} = 5780$  K. The spot is 550 K cooler than the quiet stellar photosphere and the plage temperature is derived from Eq. 3.1.

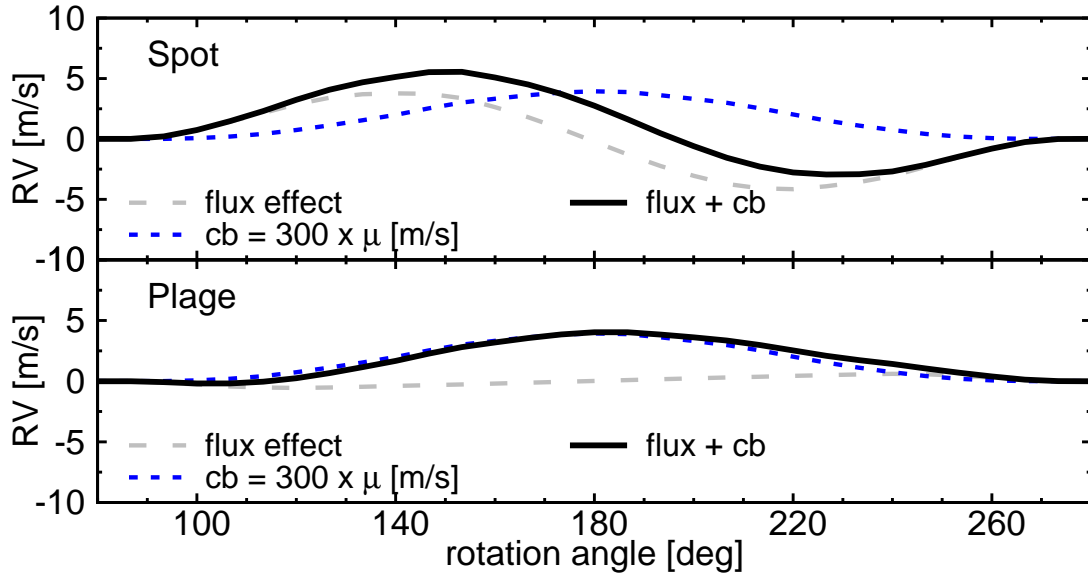


FIGURE 3.2: Upper panel: The spot RV curve produced by the flux effect only (gray dashed line), the RV curve produced by the active region by only suppressing the convective blueshift (blue dashed line) and the spot RV curve produced by taking into account the flux effect and the suppression of convection (black solid line). Lower panel: RV curves for the bright plage plotted in the same way as for the dark spot in the upper panel.

unperturbed quiet photosphere are blueshifted due to the rising convection cells. The strong magnetic fields that create dark spots and bright plages suppress the supply of new material into the active regions. As a consequence, convection inside dark spots and faculae is reduced. Lines formed in active regions are thus less blueshifted and the entire active region appears redshifted as compared to the surrounding quiet photosphere. For the Sun, the inhibition of the convective blueshift within active regions is estimated to be in the range between 200 and 500 m/s (Meunier et al., 2010; Lanza et al., 2010; Dumusque et al., 2014). This large range of values found in the literature is not only the result of the technical challenges faced when measuring the absolute convective blueshift; in fact, the convective blueshift also depends on the lines and wavelength range used to measure it (Gray, 2009; Meunier et al., 2017). For a quantitative description of the effect of reduced convection on RVs of active stars we choose here to use 300 m/s as a value for the apparent active region redshift. This value corresponds to the mean of the estimated range for the Sun.

The effect of reduced convective motions changes what we observe in RV curves of spotted stars. In addition to the flux balance being perturbed by spots or plages, the velocity field also becomes disturbed by magnetic fields. The RV curves of spots and plages incorporating different mechanisms are shown in the middle and lower panels of Fig. 3.2<sup>1</sup>. The different lines in Fig. 3.2 represent different cases. The broken gray line is our reference and represents the RV curve we get only from the flux effect of the spot and the plage.

In addition to the RV curve caused by the flux effect we also plot the RV curve caused only by the suppression of the convective blueshift as blue dashed line. The temperature of the spot and the plage are the same as the quiet photosphere (contrast of 1) but the spectra inside the active regions are redshifted by 300 m/s. Hence the

blue dashed line involves no flux effect. In this case the RV shift caused by the active region is entirely produced by the spectral lines being redshifted relative to the quiet photosphere. This constant apparent redshift of 300 m/s causes the active region to produce an RV curve only showing a redshift.

When both the flux effect and the suppression of convection are combined, the result is the black solid line. Compared to the RV curves of the spot and the plage involving only the flux effect, the RVs including the convective blueshift become biased towards positive Doppler velocities. However, there is a significant difference between what happens to the spot and the plage RV curves when the suppression of convection is considered. A dark spot emits less light than the rest of the disk and the contrast is high. Thus the redshift of this region is less important for the RV curve and the differences between RV curves with pure flux effect and flux effect in combination with convection effects in Fig. 3.2 are rather small. For a bright plage, however, the contrast is low and therefore the RV curve changes considerably from the case where only the flux effect is considered to the case where the suppression of convection is included. In the case of the plage, the RV curve is almost entirely governed by the reduction of convection while for the dark spot the flux effect plays the major role. The brightness is the key parameter here. RVs of stars with bright active regions are more sensitive to the reduction of convective motions than stars with dark spots because dark regions contribute less to the integrated disk spectrum.

Parameters that influence the strength of convection effects in spots and plages are the speed of the convective motions and the magnetic field strength. The faster the convective motions are in the quiet photosphere, the greater the impact on RV curves if convection is suppressed, because the apparent redshift of the active region becomes larger. The magnetic field strength plays a role because the ability to prevent convection grows with increasing magnetic field strength. A weak magnetic field might not suppress convection efficiently; therefore only the flux effect is observed. Strong magnetic fields can stop convection, and then apparent redshifts of active regions become important.

### 3.3 Methods

Observations of stars other than the Sun only give us access to their disk-integrated spectra. Active regions have different spectra than the quiet photosphere. Hence the presence of active regions on the stellar surface alters the integrated disk spectrum and leads to observable effects, for example, in radial velocity. In order to model the effects of activity on RVs we must reconstruct the unresolved surface of the star and mimic observations by integrating contributions of single surface elements. In this Section we explain step by step how we built our model.

#### 3.3.1 Model grid

We start by defining a grid of surface elements for the star. We define the number of grid points at the equator and adjust the number of grid points for each latitude so that all surface elements roughly cover the same surface area. This has the advantage that an active region of given size is sampled with the same number of grid points anywhere on the sphere. Typically we use 500 grid points on the equator so that the star is sampled with about 80000 grid points in total.

TABLE 3.1: Stellar parameters of stars simulated in this work.

Star	F3	G2	K5
T [K]	6900	5780	4300
$a_1$	1.110	1.180	1.067
$a_2$	-0.961	-0.751	-0.442
$a_3$	0.429	0.275	0.164
$\delta\Omega$ [rad/d]	0.112	0.078	0.043

When observing stars, we only see a projection of the stellar disk. Therefore the second step is to project the three-dimensional (3D) sphere onto a two-dimensional (2D) disk. With our grid definition the projected area of grid points at the limb is smaller than in the center. Towards the edges changes in projected velocities and intensity are large and smaller grid points avoids covering large velocity or intensity ranges with only one grid point.

The last step is to fill the plain grid with information. Each grid cell holds information about its effective temperature, intensity (limb darkening), and projected rotational velocity. In this work we present results for three stars of spectral types F3, G2, and K5. In what follows we describe their properties, which are summarized in Table 3.1:

- We estimated the effective temperatures of the F3 and the K5 star from their spectral types by fitting a 4<sup>th</sup> order polynomial to the spectral type - temperature data of [Gray et al. \(2006\)](#). For the G2 star we adopted the solar value. We use 6900K, 5780K and 4300K for the F3, G2, and K5 star respectively throughout this paper.
- The center to limb brightness variation for the F, G, and K stars was computed by [Beeck et al. \(2013b\)](#). We fit their data with a cubic limb darkening law of the form:

$$\frac{I(\mu)}{I(1)} = \sum_{k=0}^3 a_k \cdot \mu^k, \quad (3.2)$$

where  $a_0 = 1 - \sum_{k=1}^3 a_k$ . The derived coefficients for all stars are listed in Table 3.1.

- In this work we only consider slow rotators because current exoplanet surveys focus mainly on these stars as the RV content of the stellar spectra are higher when line profiles are narrow ([Bouchy et al., 2001](#)). The rotational velocity of our stars was chosen to resemble the Sun and was set to 2 km/s at the equator. However, stars are not rigid bodies and we use the relation between effective temperature and differential rotation derived by [Küker & Rüdiger \(2011\)](#) to include the differential rotation into our model. The differential rotation parameter  $\delta\Omega$  for all stars are also given in Table 3.1 and range between  $\frac{\delta\Omega}{\Omega} = 0.18$  and 0.45

### 3.3.2 Line profiles

After the stellar parameters are inserted to the grid the next ingredient for our model is the spectral contribution of each cell. There are different approaches to choosing which spectrum to use in the grid points. Some authors used atmospheric models of single lines (e.g., [Saar & Donahue, 1997](#); [Hatzes, 2002](#)), full model spectra in the

wavelength ranges of different spectrographs (e.g., [Desort et al., 2007](#); [Herrero et al., 2016](#)), or directly the CCF (e.g., [Boisse et al., 2012](#)).

In our implementation we can choose between two spectral contributions. The first one is a Voigt profile with solar parameters which is similar to the CCF approach of [Boisse et al. \(2012\)](#). The second one uses single line profiles of full MHD simulations of [Beeck et al. \(2013b, 2015b\)](#). From now on we refer to this line profile as the MHD line profile. We give here a short summary of the line synthesis but for more detailed information we refer the reader to the paper series covering the simulations used in this work ([Beeck et al., 2013a,b, 2015a,b](#)).

In this work we consider the FeI lines at 6165 and 6173 . Both are relatively isolated (no blends) and their effective Landé factors are  $g_{\text{eff}, 6165} = 0.69$ ,  $g_{\text{eff}, 6173} = 2.5$ . Hence, the FeI line at 6165 is rather insensitive to magnetic fields while the FeI lines at 6173 can be considered magnetically sensitive. Differences in the RVs derived from both lines can give us an idea as to whether or not Zeeman splitting is important in our simulations. The FeI lines were synthesized in simulations of the stellar surface convection with parameters matching those of F3, G2, and K5 stars. The 3D stellar convection process was simulated with average magnetic field strengths of 0 G (no magnetic field), 20 G, 100 G and 500 G. However, local magnetic field strength can exceed several kG in the surface simulations. Although no spots or faculae form in these simulations, areas of local strong magnetic fields exhibit significant changes in temperatures and convection patterns. Because the MHD simulations incorporate a realistic treatment of surface convection and magnetic fields, the resulting line profiles include the suppression of the convective blueshift by magnetic fields. Therefore the MHD line profiles capture one main mechanism believed to influence RV curves of active stars that is not yet observationally accessible.

The MHD line profiles capture intensity and convection effects not only globally but also locally because they are available for different  $\mu$ , from 0.1 to 1. Therefore any spatial changes in convective blueshift, line shape, or line depth due to changes of the optical surface as a function of  $\mu$  are simulated in our models. All three quantities influence the RV signals from active regions, as is discussed in Sect. 3.4.3.

To involve the spatial information from the MHD line profiles in our grid defined in Sect. 3.3.1 we have to interpolate the line profiles between the discrete  $\mu$  values and also extrapolate the line profiles to  $\mu$  values below 0.1. Although the portion of the disk with  $\mu$  values below 0.1 is negligible and is expected to have no influence on our derived RV curves we still fill these grid points with extrapolated MHD line profiles to avoid numerical discontinuities at the disk edges. The extrapolation is done by fitting a low-order polynomial to the line intensities as a function of  $\mu$ . We do this for individual wavelengths so that our model can be written as  $\hat{S}(\mu, \lambda_j) = \text{poly}(S(\mu, \lambda_j))$ , where  $\hat{S}$  is the polynomial fit and the index  $j$  denotes the individual wavelength points of the line profile. This polynomial model is then used to extrapolate the line profile to  $\mu = 0$ . When we have gained line profiles ranging from  $\mu = 0$  to  $\mu = 1$  we interpolate between the discrete  $\mu$  sampling steps of 0.1 using 2D B-splines ([de Boor, 2001](#); [Dierckx, 1995](#)).

### 3.3.3 Disk integrated line profiles and radial velocities

To derive the disk integrated spectra we need to sum the contribution of all grid points (quiet or active) visible to the observer. In the following Section we explain the three steps we take to derive the disk integrated stellar spectrum.



### Step 1: immaculate star

To compute the disk integrated spectrum is computationally expensive. However, we can save computation time if we compute the disk integrated spectrum of quiet disk elements only once and not for every active region configuration again. Therefore we start by computing the surface integrated disk spectrum of the immaculate star first (no spots or plages on the surface). This spectrum is what we call the quiet spectrum  $S_{\text{quiet}}$ . For the quiet spectrum we use line profiles computed without magnetic field. All grid points on the visible disk are assigned the corresponding interpolated line profile,  $S(\mu, \lambda, B = 0)$ . The line profiles,  $S(\mu, \lambda, B = 0)$ , are then Doppler shifted by the projected rotational velocity of the grid points including differential rotation. Interpolation of the Doppler shifted line profiles to a fixed wavelength grid is necessary to enable summation. We use spline interpolation here and denote the interpolated profile with  $\tilde{S}$ . Then the contribution of all grid points is weighted by the Planck function,  $P(T_{\text{eff}}, \lambda)$ , and the cubic limb darkening law fitted to the data of [Beeck et al. \(2013b\)](#). Finally the spectra of all grid points are summed to get the disk integrated quiet spectrum  $S_{\text{quiet}}$ :

$$S_{\text{quiet}}(\lambda, T_{\text{eff}}, B = 0) = \sum_{\text{allpixels}} \frac{I(\mu)}{I(1)} \cdot P(T_{\text{eff}}, \lambda) \cdot \tilde{S}(\mu, \lambda, B = 0). \quad (3.3)$$

### Step 2: spectra of active regions

After having computed the immaculate star spectrum we compute the spectrum integrated over active regions only. The spectrum resulting from integrating all active regions is denoted  $S_{\text{act}}$  here. Spots or plages have lower or higher temperatures than the quiet stellar disk. Consequently the Planck weight for these grid points are different. Along with the temperature change there is also a strong magnetic field present in active regions. We use the MHD line profiles of  $B = 0$  G,  $B = 20$  G,  $B = 100$  G and  $B = 500$  G and calculate the integrated spectrum for grid points defined to be active by the user (active pixels) for each B value. The integrated active region spectrum is then:

$$S_{\text{act}}(\lambda, T, B) = \sum_{\text{activepixels}} \frac{I(\mu(x_{\text{act}}, y_{\text{act}}))}{I(1)} \cdot P(T, \lambda) \cdot \tilde{S}(\mu, \lambda, B). \quad (3.4)$$

We note that we also compute the integrated active region spectrum using the  $B = 0$  G line profile and calculate  $S_{\text{act}}$  twice using once the temperature for active region  $T = T_{\text{act}}$  and once with the temperature of the quiet photosphere  $T = T_{\text{eff}}$ . We call the latter case the quiet photosphere spectrum of active elements. This is necessary to compute the full disk integrated spectrum of a star with spots or plages on the surface in the next step.

### Step 3: full disk integrated spectrum

To derive the full disk integrated spectrum of a spotted star we add the integrated active region spectrum (from step 2) to the immaculate star (from step 1). Thus we need to subtract the quiet photosphere spectrum of active elements to effectively replace the quiet disk elements with active disk elements. The full disk integrated spectrum including active regions is then simply:

$$S_{\text{tot}} = S_{\text{quiet}}(\lambda, T_{\text{eff}}, 0) - S_{\text{act}}(\lambda, T_{\text{eff}}, 0) + S_{\text{act}}(\lambda, T_{\text{act}}, B). \quad (3.5)$$

To compute the RVs from active regions we first convolve the template spectrum  $S_{\text{quiet}}$  with the resolution of the spectrograph we want to simulate (in our case HARPS;  $R = 115000$ ). For the template we measure the line centroid and define this as reference wavelength. Then we measure the line centroid of the spectra including active regions  $S_{\text{tot}}$  and compute the RV shift relative to the reference.

## 3.4 Results

In this Section we will describe how different stellar parameters and line profiles affect RV curves of spotted stars. We investigate spectral types from F3 to K5. Unless explicitly stated we follow [Lagrange et al. \(2010\)](#) and adopt their temperature difference for dark spots of  $\Delta T = 550$  K relative to the quiet photosphere for all stars. For plagues we use the temperature law of [Meunier et al. \(2010\)](#) given in Eq. 3.1.

We remind the reader here that neither the spot temperature difference of 550 K nor the plague temperature law in Eq. 3.1 might be representative for active regions on stars other than the Sun. For the dark spots observational efforts have been made to measure their temperatures in other stars. We investigate the influence of spot temperature on RVs of other stars in Sect. 3.4.4. However, for plagues, no observational constraints exist. Therefore we use the temperature relation for the Sun (Eq. 3.1) throughout this paper.

In this Section we keep the spot and plague sizes constant at 1 % of the visible stellar disk which corresponds to a radius of about 6 degrees. The upper panel of Fig. 3.1 shows a schematic picture of our simulation set up with the dark spot moving across the stellar disk of the G2 star. The color code indicates the intensity of the surface elements. The dark spot starts at the anticenter coming into view after 90 degrees of rotation and leaving the stellar disk at 270 degrees.

### 3.4.1 Influence of the magnetic field strength on RV curves

#### RV curves derived from MHD line profiles

In order to derive meaningful results from the MHD line profiles we start by testing the influence of the mean magnetic field on the RV curves. We simulate the dark spot and a bright plague for our three F, G, K stars with their parameters given in Table 3.1. We use the MHD line profiles of both FeI lines and derive the RVs induced by the active regions. We compute four simulations per star with different field strength inside the active region:  $B = 0$  G,  $B = 20$  G,  $B = 100$  G and  $B = 500$  G. The results for both FeI lines are shown in Fig. 3.3 for the spot and in Fig. 3.4 for the plague. In both cases the results of the FeI line at 6165 do not substantially differ from the ones derived from the FeI line at 6173 (see solid gray and red dotted lines on the right panel of Figs. 3.3 and 3.4). Although the two FeI lines differ in their effective Landé factors ( $g_{\text{eff}, 6165} = 0.69$ ,  $g_{\text{eff}, 6173} = 2.5$ ) the fact that the results are very similar leads us to the conclusion that Zeeman broadening plays no important role in both lines at the magnetic field strengths considered here. This result is consistent with [Reiners et al. \(2013\)](#) where the RV amplitude caused by a purely magnetic spot (considering only the Zeeman effect and no temperature contrast) is predicted to be below 1 m/s in the optical for a magnetic field strength of 600 G.

In Figs. 3.3 and 3.4 we can see the evolution of the RV curves with increasing mean magnetic field inside the active region for the spot and plague respectively. In the case of  $B = 0$  G we only see the flux effect for all stars causing symmetric, sinusoidal RV variations when the active region is visible (and a flat RV curve when the active region

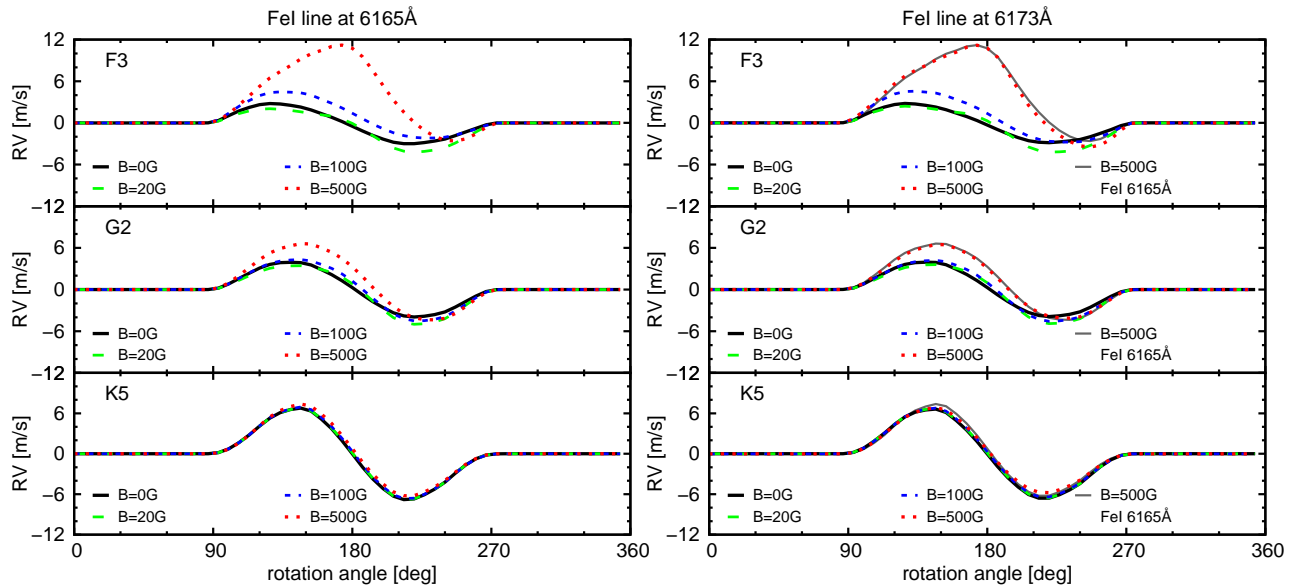


FIGURE 3.3: Influence of magnetic field strength on spot RVs. Line profiles with different field strengths were used inside the spot. Spot RVs calculated with  $B = 0$  G (black solid line), with  $B = 20$  G (green long dashed line), with  $B = 100$  G (blue short dashed line) and with  $B = 500$  G (red dotted line). The results for the FeI line at 6165 are shown in the left panel, and the results of the FeI line at 6173 can be seen in the right panel. The gray solid line in the right panel indicates the results of the FeI line at 6165 for comparison.

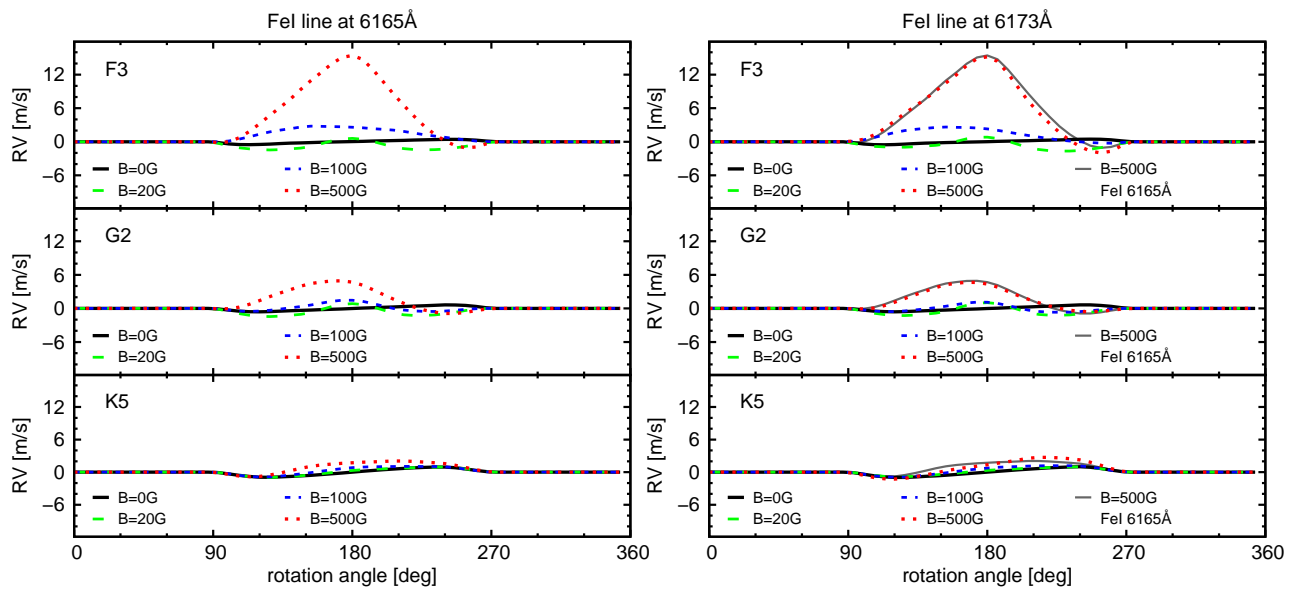


FIGURE 3.4: Influence of magnetic field strength on plage RVs. Line profiles with different field strengths were used for the plage. Plage RVs calculated with  $B = 0$  G (black solid line), with  $B = 20$  G (green long dashed line), with  $B = 100$  G (blue short dashed line) and with  $B = 500$  G (red dotted line). The results for the FeI line at 6165 are shown in the left panel, and the results of the FeI line at 6173 can be seen in the right panel. The gray solid line in the right panel indicates the results of the FeI line at 6165 for comparison.

TABLE 3.2: Measured active region velocities as a function of magnetic field strength  $B$ .

Star	Method	$B = 20$ G	$B = 100$ G	$B = 500$ G
F3	Gaussian	63 m/s	35 m/s	781 m/s
F3	Centroid	73 m/s	239 m/s	1556 m/s
G2	Gaussian	123 m/s	101 m/s	300 m/s
G2	Centroid	78 m/s	105 m/s	406 m/s
K5	Gaussian	30 m/s	64 m/s	76 m/s
K5	Centroid	20 m/s	50 m/s	97 m/s

rotates behind the star). When the magnetic field inside the active region increases we start to see growing asymmetries in the RV curves of both spot and plage. Up to an average magnetic field strength of 100 G these asymmetries are small and we mainly observe the flux effect in the RV curves. The only exception is the F3 star where a magnetic field strength of 100 G already has impact on the RV curves. When we increase the average magnetic field strength further to 500 G the RV curves of the F and G stars show a significant change. However, for the K5 star there is little change in the active region RV curves even at  $B = 500$  G. Do determine why different stellar types show different sensitivities to the magnetic field inside the active region we look into the convection of each star in more detail.

### Convection at different magnetic field strengths

We attribute the RV amplitude and symmetry changes seen with increasing magnetic field strength in Figs. 3.3 and 3.4 to the suppression of convection inside the active region (as explained in Sect. 3.2.2). To test this idea we measure the suppression of the convective blueshift as a function of the magnetic field strength. We use the local MHD line profiles at  $\mu = 1$  here. First we determine the line centers of the two FeI lines at  $B = 0$  G,  $B = 20$  G,  $B = 100$  G and  $B = 500$  G by fitting a Gaussian to the ten data points around the minimum of the line profile. The effective velocity difference between active region and quiet photosphere is then calculated from the difference of the line centers with magnetic field ( $B = 20$  G,  $B = 100$  G and  $B = 500$  G) relative to the line center determined for the  $B = 0$  G line:  $(\lambda_B - \lambda_{B=0})/\lambda_{B=0} \cdot c$ . This method will be denoted as the Gaussian method from now on.

A second way to measure the apparent active region velocity is to use the line centroids instead of a Gaussian fit. This method will be denoted as centroid method from now on.

The resulting apparent active region velocity are plotted in the left panel of Fig. 3.5. As the results differ slightly between the FeI line at 6165 and the FeI line at 6173 we shaded the area between the two measurements. Open symbols in Fig. 3.5 indicate the measurements of the Gaussian method and filled symbols show the results obtained by the centroid method. A summary of the mean velocity shifts plotted in Fig. 3.5 can also be found in Table 3.2 for each  $B$  step.

Before exploring the details of Fig. 3.5, we remind the reader here that the active region velocity measured depends on the depth of the line used and the lines wavelength (Gray, 2009; Meunier et al., 2017). The flux at the bottom of the FeI lines (at  $\mu = 1$ ),  $F$ , are  $F_{6165} = 0.9$  and  $F_{6173} = 0.83$  for the F3 star,  $F_{6165} = 0.6$  and  $F_{6173} = 0.34$  for the G2 star and  $F_{6165} = 0.47$  and  $F_{6173} = 0.17$  in the K5 star. Therefore the lines we used here to compute the apparent active region velocity are not representative for all

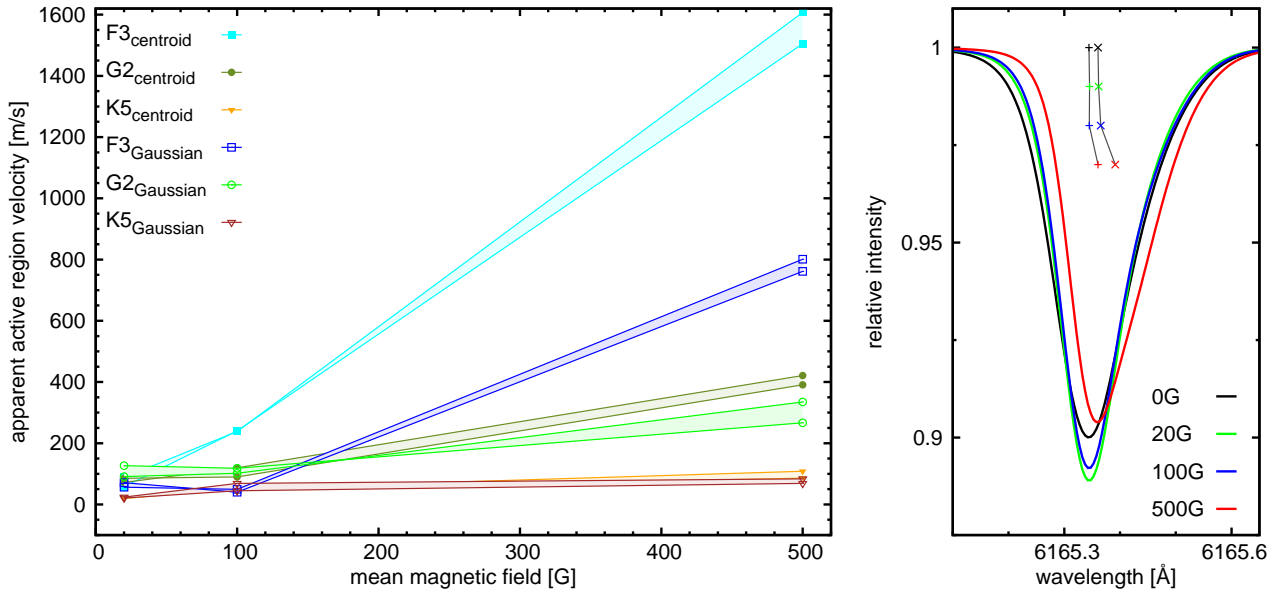


FIGURE 3.5: Left panel: variation of the active region velocity with the magnetic field strength  $B$  for different stellar types and measurement methods. Right Panel: MHD line profile of the FeI line at 6165 Å for different magnetic field strength (solid lines) and line center measurements with the Gaussian method (crosses) and the centroid method (x). The color code for the line profiles and the center measurements are the same: black for  $B = 0$  G, green for  $B = 20$  G, blue for  $B = 100$  G and red for  $B = 500$  G.

spectral lines. Nevertheless, the two lines used here give us a first impression how the active region velocities (derived from a consistent line list) change over spectral type and allow us to study the impact of this change on activity-induced RVs.

For the K5 and G2 stars the results obtained from the Gaussian method and the centroid method are similar (see table 3.2). For the K5 star there is only a small change in the apparent active region velocity when changing the magnetic field from  $B = 20$  G to  $B = 500$  G. Moreover, the apparent active region velocities are low. Therefore the absence of a change of the RV curves with increasing  $B$  shown in Figs. 3.3 and 3.4 can be explained by the small values of the apparent active region velocities.

For the G2 star the apparent active region velocities are also low and relatively stable for  $B = 20$  G and  $B = 100$  G. However, at  $B = 500$  G the suppression of convection becomes significant. Therefore, also in the G2 star the observed change in the RV curves of spot and plage coincide with the apparent active region velocities which increases for stronger magnetic fields.

For the F3 star the apparent active region velocities obtained with the Gaussian method differ significantly from the results of the centroid method (see left panel of Fig. 3.5 and Table 3.2). The reason for the large differences in apparent active region velocities among the two methods lies in the MHD line profiles themselves. In the right panel of Fig. 3.5 we plot the FeI line profile at 6165 Å for all simulated magnetic field strengths and indicate the line centers obtained by the Gaussian method with crosses and with the centroid method as x. The line centers obtained for  $B = 0$  G and  $B = 20$  G are similar when measured with both methods. However, with increasing magnetic field strength the centroid method returns highly redshifted values because the spectral line becomes significantly asymmetric. We argue that the apparent active

region velocity should be measured by the Gaussian method (as also used in [Beeck et al., 2013b](#)) because it measures the line core and thereby avoids a bias of the line centers towards redder wavelengths as a result of line asymmetries. Hence, the change in the RV curves of the F3 star in [Fig. 3.3](#) (spot) and [Fig. 3.4](#) (plage) only partially coincide with the inhibition of convection. The picture for the F3 star is more complex and line asymmetries as a function of magnetic field strength also play an important role here.

Because of the results obtained for the F3 star we decide to use the Gaussian method and its results of the apparent active region velocity in what follows in this work. Line asymmetries and their effect on RV curves of active regions are discussed later in more detail in [Sect. 3.4.3](#).

### RV amplitudes across different stellar types

It is interesting to note that the growth in asymmetry and RV amplitude with increasing magnetic field is seen more in hot stars than in cool ones. In the K star simulation the spot RV amplitude increases only from 6.8 m/s to 7.4 m/s (factor of 1.09) between the  $B = 0$  G and  $B = 500$  G run. For the G star the spot RV amplitude increases from 4 m/s to 6.6 m/s (factor of 1.65) if the magnetic field is increased from 0 G to 500 G and on the F star the RV amplitude grows from 2.8 m/s in the 0 G run to 11.2 m/s in the 500 G run (factor of 4).

For the plage simulation we observe a similar behavior. Changing the magnetic field inside the plage from 0 G to 500 G changes the RV amplitude in the K star simulation from 0.9 m/s to 2 m/s (factor of 2.2) in the G star simulation from 0.6 m/s to 5 m/s (factor of 8.3) and in the F star simulation from 0.5 m/s to 15.5 m/s (factor of 31).

We have to be careful when comparing the growth of the RV amplitude in [Figs. 3.3](#) and [3.4](#) between the F, G, and K stars. Because the same spot temperature contrast ( $\Delta T = 550$  K) and the same plage temperature contrast law ([Eq. 3.1](#)) are used for all stars, the intensity contrasts of the active regions depend on the stellar type. Thus, one possibility for the increase in amplitude towards hotter stars could be the changing active region intensity contrast.

The dark spot has a contrast of 0.74 on the F star, a contrast of 0.65 on the G star and a contrast of 0.45 on the K star. As explained in [Sect. 3.2.2](#), the suppression of convection is greater in RV curves of brighter active regions. Because the dark spot is relatively brighter on the F star compared to the G or K star, the suppression of convection is seen more in the spot RV curves of the hotter stars. However, for plages the effect should be exactly reversed because the plage is relatively brighter on the K star (contrast 1.35) than on the G star (contrast 1.19) or F star (contrast 1.13). According to the contrast we expect the plage signal to have the largest amplitude growth in the K star. Nevertheless, we observe the opposite and the plage RV amplitudes are also larger for hotter stars.

The decreasing sensitivity of the RV curves to magnetic fields in cooler stars is therefore not primarily caused by changing contrast. The spectral type dependence of the RV amplitude growth from  $B = 0$  G to  $B = 500$  G is better explained by the convection speeds. As shown in [Sect 3.4.1](#) the measured active region velocity increases towards earlier-type stars (76 m/s for the K5 star, 300 m/s for the G2 star and 781 m/s for the F3 star). Suppressing fast convective motions leads to a strong impact on the RV curves. Thus the increase in asymmetry and RV amplitude towards the  $B = 500$  G simulation is well explained by the different convective speeds of the stars and their suppression caused by the magnetic field.



A further consequence of the decreasing convective motion speeds towards late-type stars is that plages contribute less to the RV variation of an active star with decreasing effective temperature. As plages have low contrasts their signal comes mainly from the inhibition of convection within them. If convection is slow in the quiet photosphere plage signals become weaker. Nevertheless, we remind the reader that we simulate spots and plages covering only 1% of the visible stellar surface. On the Sun, plages cover larger areas than spots; therefore one has to be careful here when comparing absolute RV amplitudes of spots and plages. Plages can still produce large RV variations in cool stars because they are typically larger than simulated here.

### 3.4.2 Comparison to the literature

After finding that only simulations using the 500 G line profiles lead to a noticeable effect in the RV curves through the suppression of the convective blueshift we go ahead and attempt to validate our results by comparing them with published studies. The simulation of the G2 star is the perfect test ground for inter comparison as most other works in this field have been done with sun-like stars. Especially interesting is the comparison between our results and the ones obtained by [Dumusque et al. \(2014\)](#). Their simulations are based on observed spectra of the solar quiet photosphere and a solar spot while our results are completely based on MHD simulations. Although [Reiners et al. \(2016b\)](#) pointed out that the absolute wavelength scale of the data used in [Dumusque et al. \(2014\)](#) has some problems, a close match of the results can thus give us confidence to extend our method used to simulate activity RV curves to stars other than the Sun.

The setup of the simulations in this Section is the same as before using an equatorial active region with a size of 1 % of the visible stellar disk and the stellar parameters of [Table 3.1](#). For the active region we now only use the MHD line profiles with a mean magnetic field of 500 G because only then the suppression of the convective blueshift is significant. We use our active region setup and simulate the same configuration with the SOAP 2.0 code of [Dumusque et al. \(2014\)](#). We show our RV curves and the ones obtained with SOAP 2.0 for the dark spot and the bright plage in [Fig. 3.6](#).

For the dark spot our RV curve agrees well with the prediction of SOAP 2.0. We see in [Fig. 3.6](#) that the maximum deviation of the two curves occurs around a rotation angle of 240 degrees and is found to be 1.14 m/s or 17 % of the amplitude. If we compare our plage RV results of the G2 star to SOAP 2.0 we find the maximum difference between the two curves again around a rotation angle of 240 degrees. This time the maximum difference is 1.57 m/s or 31 % of the amplitude.

Both dark spot and plage RV curves agree well with the findings of SOAP 2.0. Nevertheless, we take a closer look at how the RV curves are derived in both cases and give a possible explanation for the differences. Equivalent width, FWHM, line bisectors, and convective blueshift of the MHD line profiles are known to be functions of  $\mu$  ([Beeck et al., 2013b, 2015b](#)). Changes in these parameters have impact on the RV signals computed from the line profiles. SOAP 2.0 uses observed high-resolution, spatially resolved spectra of a quiet part and a sunspot on the solar surface close to the disk center to simulate RV curves for spots and plages. As there are no observations available for different  $\mu$ , the model of SOAP 2.0 uses constant line profiles and therefore a constant convective blueshift (corresponding to the convective blueshift close to the Sun's disk center) to derive RV signals from spots and plages. Hence, the differences between our models and SOAP 2.0 are likely related to the fact that SOAP 2.0 uses constant line profiles while we use limb-angle dependent MHD line profiles.

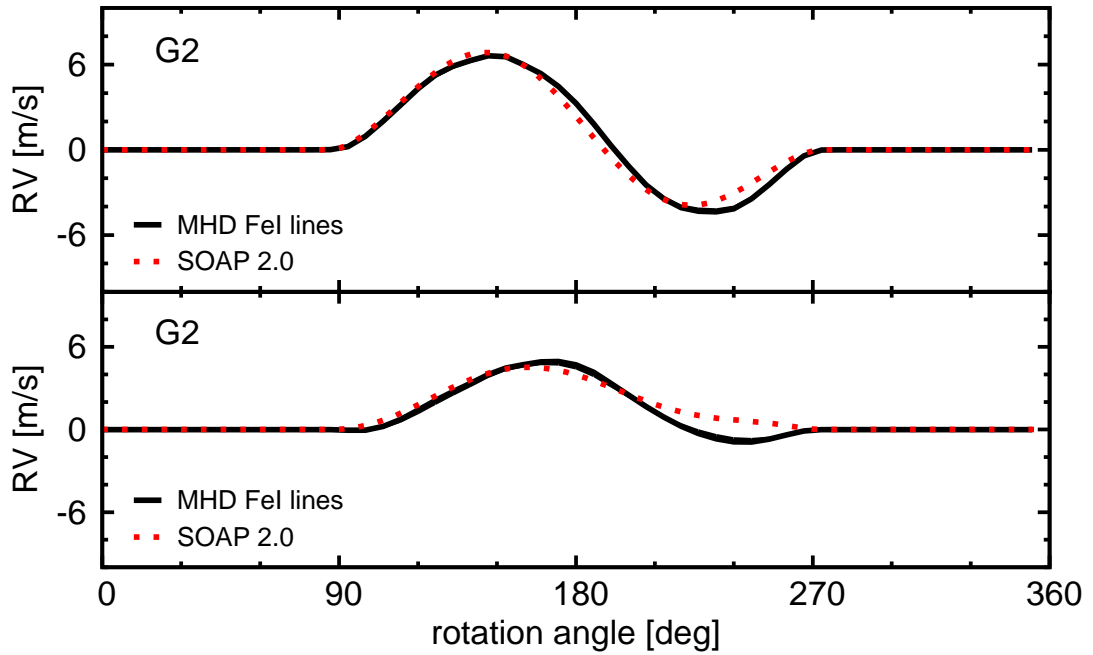


FIGURE 3.6: Comparison between the RV curves derived by our code using the MHD line profiles (black solid lines) and SOAP 2.0 (red dotted lines) for a dark spot (upper panel) and a bright plage (lower panel).

However, we note here that the MHD line profiles might not fully cover all changes in line shape and line depth with changing  $\mu$ . This was recently pointed out by [Reiners et al. \(2016a\)](#) who compared the MHD line profiles to observations made during a solar eclipse. Observations of line profile changes with  $\mu$  will be useful for future work in this area of research.

Despite the differences between our results and SOAP 2.0 the general results of both spot and plage on the G2 star agree well. The amplitudes match for both spot and plage, and the shape of the RV curves also agrees. In the following Sections we move on and explore the role of convective blueshift, line profiles, and spot contrast for activity RV curves of hotter and cooler main sequence stars.

### 3.4.3 Influence of limb-angle-dependent convective blueshift and line profiles on RV curves

In this Section we try to understand the role of changes in the convective blueshift and line profiles from center to limb in RV curves in more detail. For this purpose we try to recover the RV curves derived from the MHD line profiles with a simulation based on Voigt profiles with solar parameters. Also in this Section we explain how we can use this method to test whether or not convective blueshift variations and line profile changes across the disk effect the resulting RV curves.

#### The apparent active region velocity across the disk

As a first step we derive the apparent active region velocity across the disk for all spectral types. We use the local MHD line profiles for 0 G and 500 G available for 10  $\mu$  values between 0.1 and 1 and measure the line centers with the Gaussian method as described

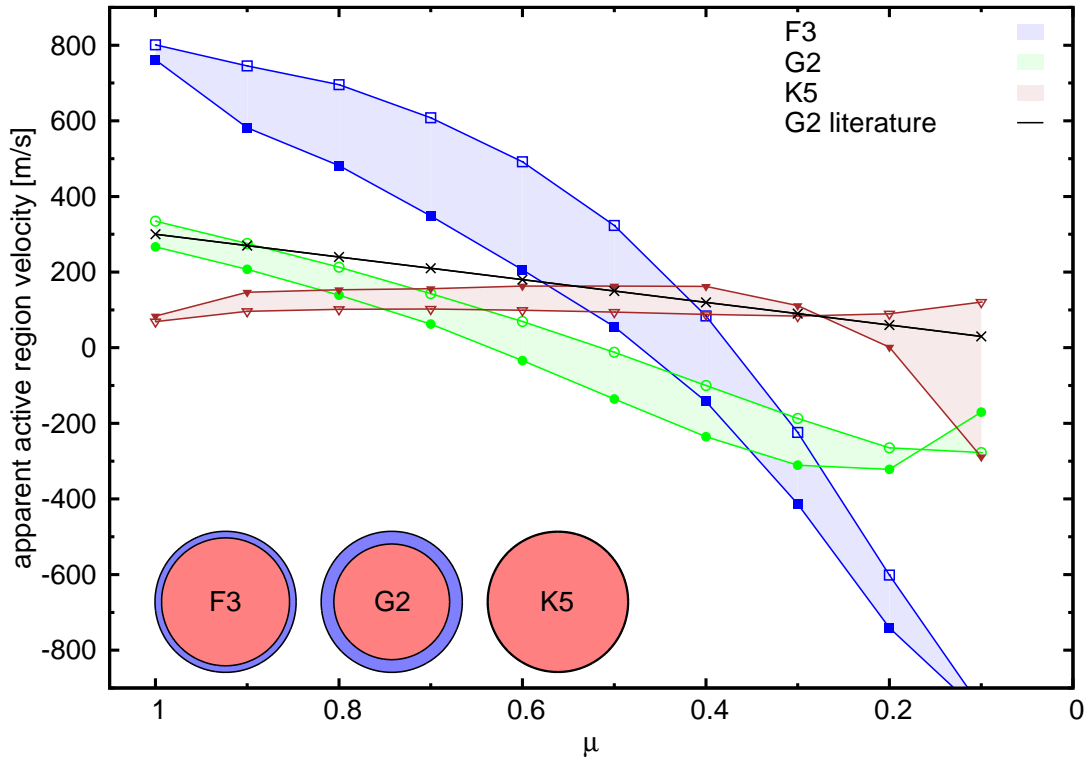


FIGURE 3.7: Variation of the active region velocity field between  $B = 0$  and  $B = 500$  G as a function of  $\mu$  for the F3 star (blue squares), G2 star (green circles) and the K5 star (red triangles). The results of the FeI line at 6165 is shown with open symbols and the result of the FeI line at 6173 is indicated as filled symbols. For comparison the constant active region velocity field taking into account only the projection effect with  $\mu$  (as done in previous works) is plotted as black line. Red-blue circles in the bottom left indicate where on the stellar disk active regions appear redshifted or blueshifted relative to the quiet photosphere.

in Sect. 3.4.1. The measurement is taken for both FeI lines at 6165 and 6173. The resulting velocity difference between active region and quiet photosphere is plotted in Fig. 3.7 for different  $\mu$  on the stellar disk.

It is important to note that our result is different in amplitude and span from the estimate of the convective blueshift in Beeck et al. (2013b). They measure the line core shift of the  $B = 0$  lines relative to the laboratory wavelength of the FeI lines. This results in an estimate of the absolute velocities of the convective motions. For our purpose however, it is important to measure the suppression of the convection by the magnetic field because this is what is seen in the RV curves of spots and plages.

From Fig. 3.7 we can study the evolution of the convective blueshift from center to limb. As already shown in Beeck et al. (2013b) the convective blueshift can turn into a redshift towards the limb of stars. Active region can therefore appear blueshifted when observed at low  $\mu$  values. Previous works studying the RV variation caused by active regions have often taken into account only the projection effect of the convective blueshift with  $\mu$ , shown as black line in Fig. 3.7 for the G2 star. Hence, the effect that active regions can also become blueshifted at the limb was not taken into account.

The area on the disk where active regions appear red or blueshifted changes with spectral type. As shown by the red-blue circles in the lower-left corner of Fig. 3.7 the

active regions on the K5 star appear redshifted almost on the entire disk. Only on the outermost 2% of the projected disk do active regions appear blueshifted. In the G2 star active regions appear blueshifted in the outer 33% of the disk and for the F3 star we measure this value to be 18%. Also the apparent convective motions across the stellar disk are changing with spectral type. The span is considerably high for the F3 and G2 stars, about 1800 m/s and 600 m/s, respectively, from center to limb. However, for the K5 star there is almost no variation.

### Using Voigt profiles with different convective blueshift parameters

With the measurement of the active region velocity field in hand, we try to recover the RV curves derived from the MHD line profiles with a model parametrized as simply as possible. We want to test if we can reproduce the RV curve from MHD line profiles by using a constant Voigt profile throughout the disk and different models for the effective velocity field of the active region. We calculate three sets of RV curves for each star. The first one uses a constant effective redshift for the active region of 781 m/s for the F3 star, 300 m/s for the G2 star and 76 m/s for the K5 star corresponding to the redshift seen at  $\mu = 1$ . This approach is close to the one used by [Dumusque et al. \(2014\)](#) with observed Sun spectra. In our second setup we take the projection effect of the convective blueshift into account (see black line in Fig. 3.7 for the G2 example). The third simulation uses a variable effective active region velocity field. To simulate the variable velocity field we place the Voigt profiles at the measured line center positions of the MHD profiles. For the quiet photosphere we place the Voigt profiles at the center wavelength of the  $B = 0$  G line profiles while for active regions we place the Voigt profiles at the center wavelength of the  $B = 500$  G MHD line profiles. In this way we reproduce exactly the active region velocity fields shown in Fig. 3.7 and any remaining difference between the spot and plage RV curves derived from MHD and Voigt profiles can only originate from the line shape. The four models for the spot are compared in Fig. 3.8 and the models for the plage can be seen in Fig. 3.9. The black solid lines are the MHD ( $B = 500$  G) simulations, the green dashed lines represent the models with constant active region velocities, the blue dashed lines represent the model taking into account the projection effect of the active region velocities and the red dashed lines use the variable active region velocity field from Fig. 3.7.

### Changes induced by convective blueshift models

In order to single out the influence of the convective blueshift alone on the RV curves of active regions we analyze the Voigt profile simulations. The only difference between the Voigt profile simulations shown in Figs. 3.8 and 3.9 is the active region velocity field across the disk. When the spot or plage is at disk center, at  $\mu = 1$ , which corresponds to a rotation angle of 180 degrees, all three simulations return the same result because the active region velocities are the same. However, when the active region moves towards the limb the difference in apparent active region velocity field impacts the RV curves.

For the spot as well as the plage we observe that the maximum differences among the Voigt profile simulations occur for all spectral types at rotation angles between 110 and 120 degrees and between 240 and 250 degrees (60 to 70 degrees from the disk center at  $0.34 < \mu < 0.5$ ). Although the difference in the apparent active region velocity among the simulations increases towards the limb (see Fig. 3.7 for the G2 example), the projected active region size decreases towards the limb. Therefore the maximum RV difference between the Voigt profile simulations is found at a distance to the disk

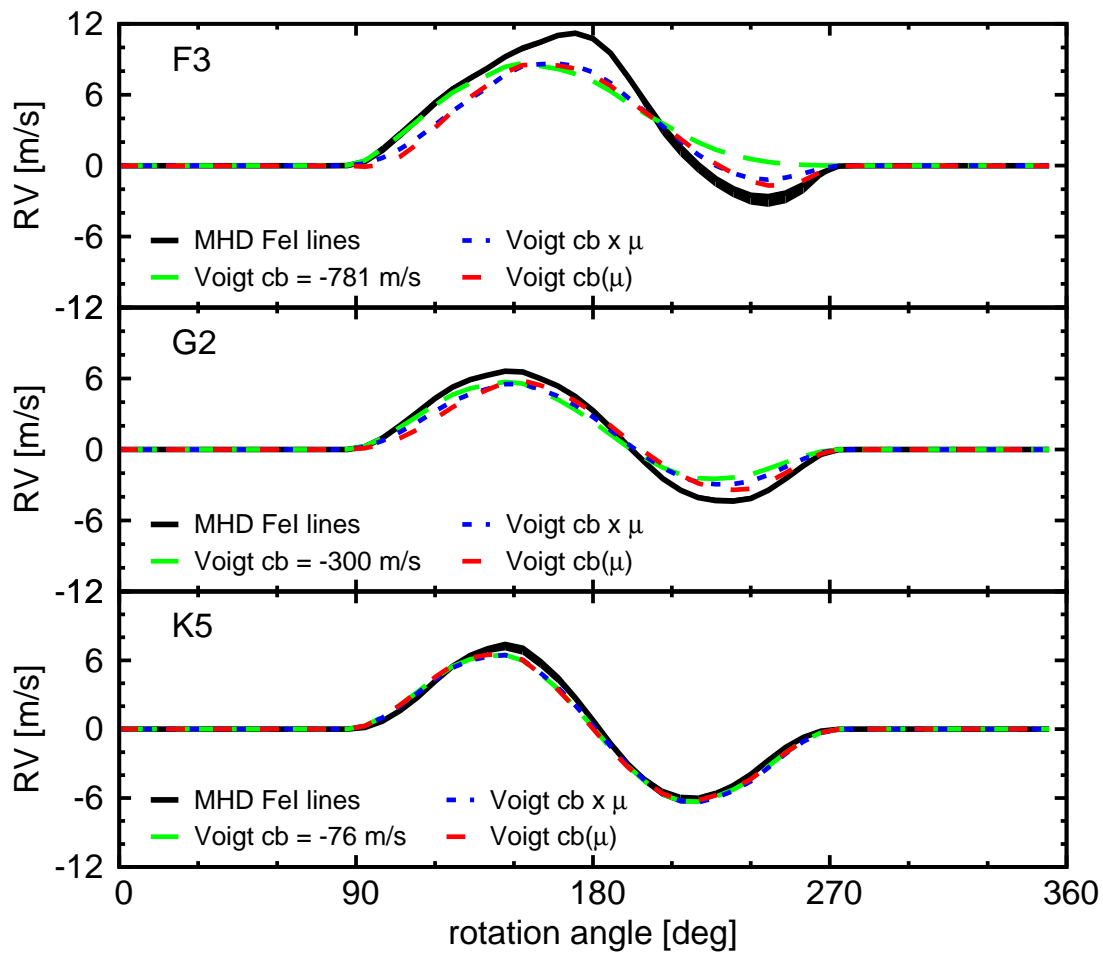


FIGURE 3.8: Spot RVs calculated for different stellar types with different line profiles. Calculations done with the 500 G MDH line profiles (black solid lines), Voigt profiles and a constant convective blueshift (green long dashed lines) and Voigt profiles including a varying convective blueshift (blue short dashed lines).

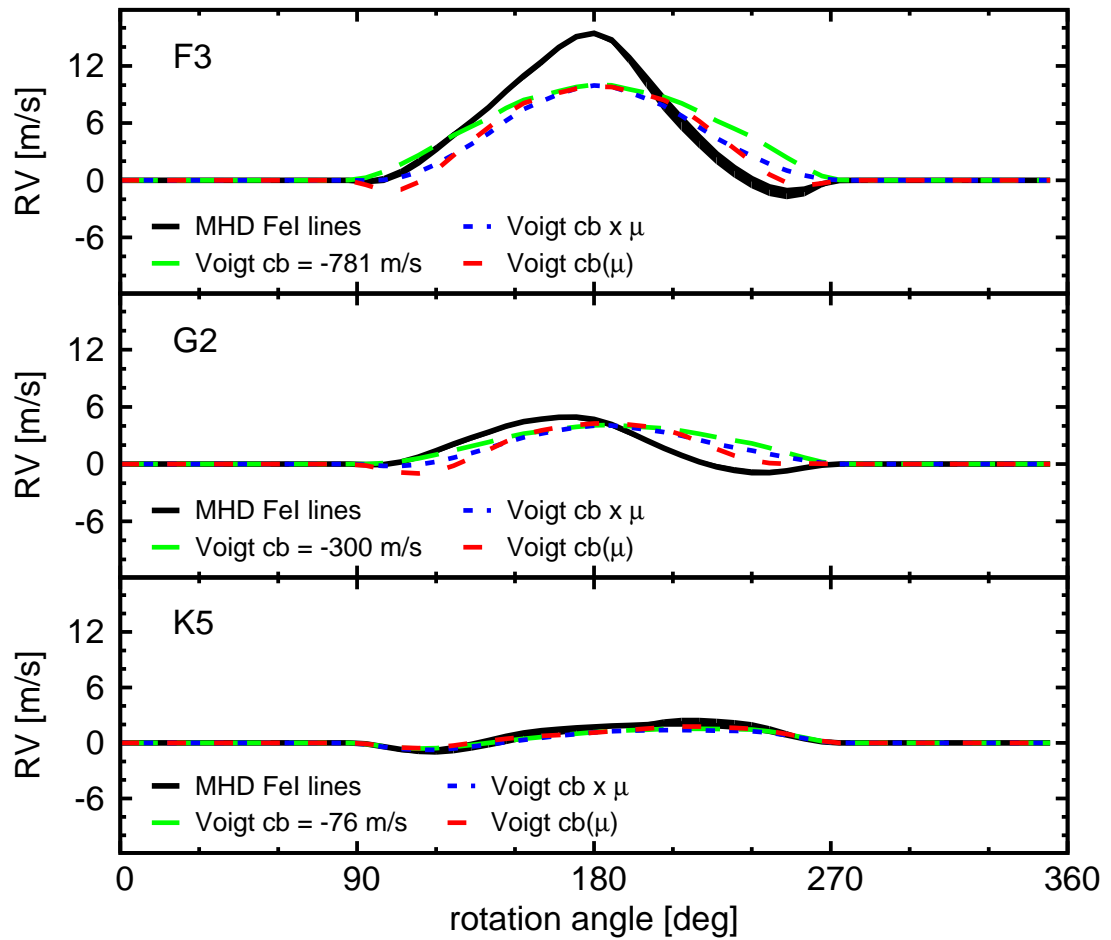


FIGURE 3.9: Plage RVs calculated for different stellar types with different line profiles. Calculations done with the 500 G MDH line profiles (black solid lines), Voigt profiles and a constant convective blueshift (green long dashed lines) and Voigt profiles including a varying convective blueshift (blue short dashed lines).



TABLE 3.3: Differences between Voigt profile models using different convective blueshift models.

Star	Models	diff spot [m/s]	diff plage [m/s]
F3	$cb = -781$ and $cb \cdot \mu$	1.6	2.0
F3	$cb = -781$ and $cb(\mu)$	2.0	2.8
F3	$cb \cdot \mu$ and $cb(\mu)$	0.6	1.3
G2	$cb = -300$ and $cb \cdot \mu$	0.6	0.8
G2	$cb = -300$ and $cb(\mu)$	1.2	1.8
G2	$cb \cdot \mu$ and $cb(\mu)$	0.6	1.0
K5	$cb = -76$ and $cb \cdot \mu$	0.1	0.2
K5	$cb = -76$ and $cb(\mu)$	0.1	0.2
K5	$cb \cdot \mu$ and $cb(\mu)$	0.1	0.4

center at which the active region velocities differ significantly while the active region itself does not appear too small to produce a significant RV signal. The maximum differences between the three Voigt profile simulations are summarized in Table 3.3.

In Table 3.3 we see that differences in the treatment of the convective blueshift can result in differences of the simulated RV curves in the order of a few m/s. From our results we conclude that for future RV models it is not only important to have accurate constraints for the absolute value of the convective blueshift but it is also necessary to take the variation of the convective blueshift from center to limb into account.

### Changes induced by line shape

Figures 3.8 and 3.9 show that the model using a variable active region velocity field (red broken line,  $cb(\mu)$  model) can follow the spot RV curves from the MHD line profiles best. This is not surprising because the  $cb(\mu)$  model follows the change in convective blueshift of the MHD line profiles from center to limb exactly. Nevertheless, we also observe large differences between the MHD and  $cb(\mu)$  models. For the F3 star, the maximum difference between the MHD and the  $cb(\mu)$  model is 3.1 m/s for the spot and 5.5 m/s for the plage. For the G2 star we find this difference to be 1.5 m/s for the spot and 2.3 m/s for the plage while the differences for the K5 star are 1 m/s for the spot as well as the plage. Spot RVs are generally less sensitive to line profile changes than plage RVs are. This is because spots are cooler than the surrounding photosphere and consequently darker. Hence the line profile inside the spot has less weight than the surrounding photosphere and the flux effect is dominant. For plages the situation is exactly the opposite and the RV curves become very sensitive to line profiles changes.

The differences between the RV curves derived from MHD and Voigt profiles suggest that the line profiles play an important role for the derived RV curves. Because the Voigt profile is a symmetric function, the remaining differences in the RV curves of spot and plage come from line shape and line depth (line bisectors, FWHM and equivalent width) changes with  $\mu$  in the MHD line profiles. The line profiles of both the quiet photosphere and the active region vary from center to limb as can be seen in Fig. 3.10 for the G2 star (Beeck et al., 2013b, 2015b). As reference we plot the Voigt profile used to derive the models in Figs. 3.8 and 3.9 as black (at  $\mu = 1$ ) and gray (at  $\mu = 0$ ) broken lines. To visualize the differences between the line profiles we plot the residuals of the MHD minus the Voigt profile in the bottom of Fig. 3.10.

It is interesting to note that the maximum difference between the MHD and the  $cb(\mu)$  model in the F3 star occurs at disk center (rotation angle 180 degrees) for both

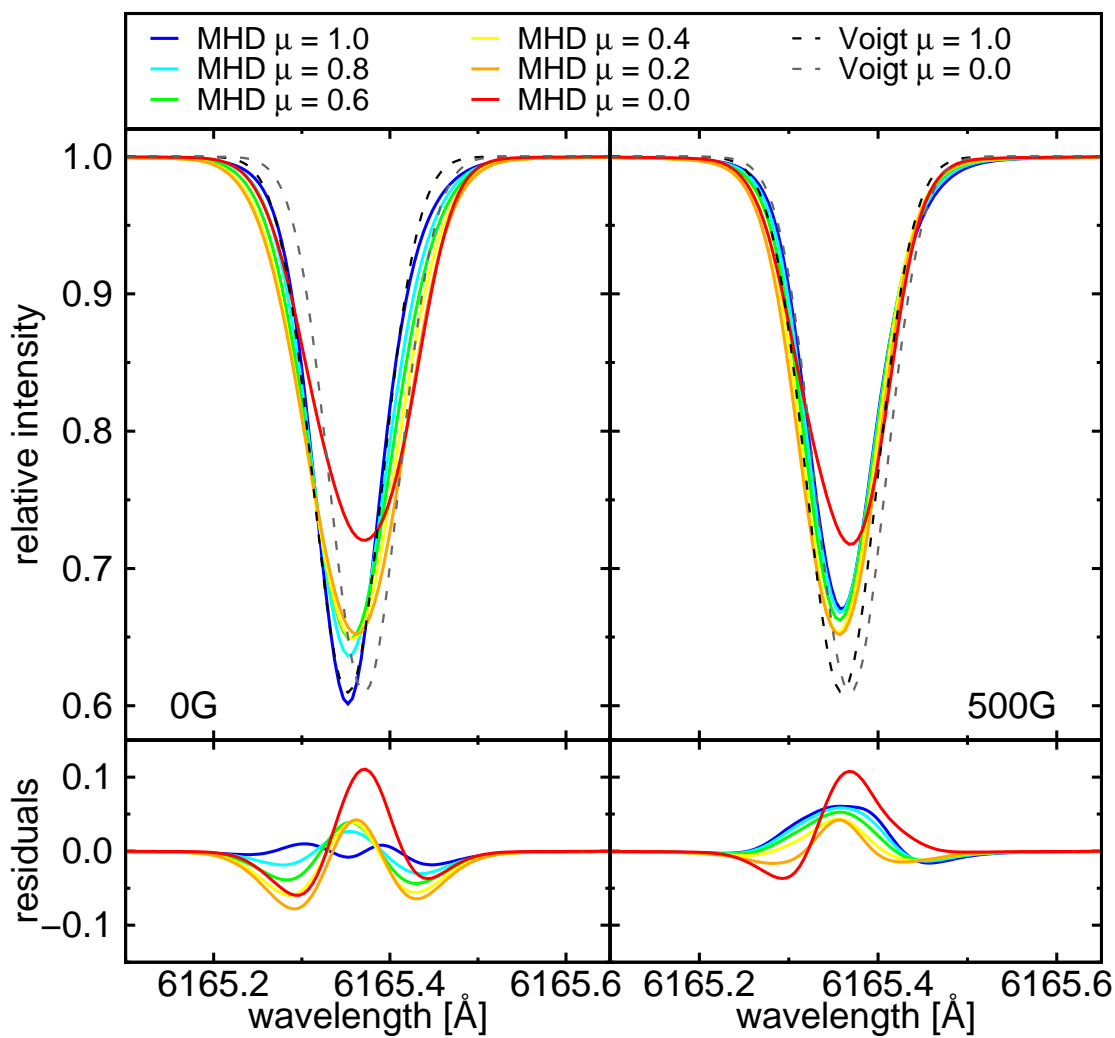


FIGURE 3.10: Variation of the FeI line at 6165 Å with  $\mu$  for  $B = 0$  G left, and  $B = 500$  G right panel. MHD line profiles at different  $\mu$  values are color coded from blue ( $\mu = 1$ ) to red ( $\mu = 0$ ). For comparison, Voigt profiles with solar parameters are overplotted as black ( $\mu = 1$ ) and gray ( $\mu = 0$ ) broken lines.

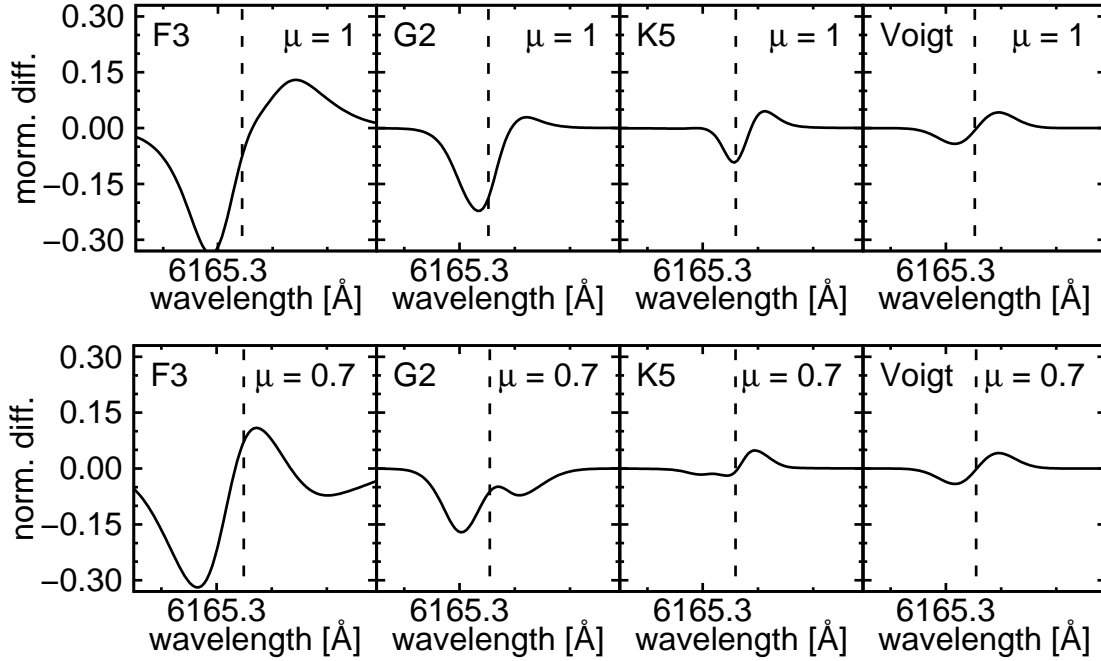


FIGURE 3.11: Difference between the  $B = 500$  G and  $B = 0$  G line profiles normalized by the line depth of the  $B = 0$  G line. Panel from left to right show the result for the F3, G2, K5 MHD lines, respectively. The rightmost panel depicts results from the  $cb(\mu)$  simulation using Voigt profiles with solar parameters. Upper panels: results at  $\mu = 1$ . Lower panel: results at  $\mu = 0.7$ .

spot and plage models (see Figs. 3.8 and 3.9). This is not the case for the G2 and K5 star where the differences are largest at rotation angles between 30 and 50 degrees away from the disk center. The fact that the maximum occurs at different rotation angles in the F3 star compared to the G2 and K5 star is related to the relative line profile changes of the MHD lines in different stars between the  $B = 0$  G (quiet photosphere) and  $B = 500$  G (active region) lines.

In Fig. 3.11 we plot the difference between the  $B = 500$  G and  $B = 0$  G line profiles normalized by the line depth of the  $B = 0$  G line. We show two rows of plots: one at  $\mu = 1$  and one at  $\mu = 0.7$  (where the maximum differences between Voigt and MHD simulations occur). As a reference we also depict the results obtained for the Voigt profiles of our  $cb(\mu)$  model in the rightmost panels. The vertical black dashed lines mark the line center of the  $B = 0$  G lines as measured by the Gaussian method. We note that the results shown in Fig. 3.11 are for a temperature contrast of  $\Delta T = 0$ .

Figure 3.11 can be interpreted as the perturbation introduced to the local line profiles by the magnetic field. The rightmost panels show the ideal cases in which the line profiles themselves do not change with  $B$ . The perturbation observed in the Voigt profile case originates from the two lines being shifted with respect to each other because of the suppression of convection. The MHD line profiles show different perturbation patterns that are asymmetric and not necessarily centered at line center. These asymmetric perturbation patterns are caused by the change in line depth and shape (asymmetries) when the magnetic field is turned on. Even in the disk center perturbations can be asymmetric resulting in the large differences between the  $cb(\mu)$  and MHD models seen in the F3 star (see Figs. 3.8 and 3.9). For this reason we cannot reproduce the RV curves of spots and plages using symmetric Voigt profiles.

### 3.4.4 Influence of spot temperatures in F, G, and K stars

Until now we have assumed that spots have the same temperature difference relative to the quiet photosphere in all stars. The spot temperature is a crucial parameter in our models as it directly influences the RV amplitude and the asymmetry of the RV curve. As explained in Sect. 3.2.1 a lower spot temperature leads to a larger spot contrast. For high spot contrasts the flux effect is dominant. For low spot contrasts however the suppression of convective motions in active regions becomes important and can be seen as asymmetries in the RV curves. The temperature difference between quiet photosphere and a spot of  $\Delta T = 550$  K seen in the Sun might not necessarily represent the situation in F and K stars. Therefore, we investigate the influence of spot temperatures on the RV curves of F to K stars in this Section. We aim to use realistic spot temperatures for individual spectral types and investigate what influence convective blueshift has on spot RVs if the spot temperature on other stars is taken into account.

#### Spot temperatures

There have been many efforts to measure the spot temperatures in other stars with a variety of methods. Berdyugina (2005) presented a summary of such measurements in their Table 5. We plot the measurements summarized in Berdyugina (2005) in Fig. 3.12 and fit a linear relation to the data excluding EK Dra as done in Berdyugina (2005). The relation found between effective temperature and temperature difference to the quiet photosphere of dark spots will be used in our simulations to produce more realistic spot RV curves for other stars.

The relation we find between effective temperature and temperature difference of the spot is  $\Delta T = 0.64 \cdot T_{eff} - 1842.64$ . The spot temperature measurements cover a range of stellar effective temperatures from  $T_{eff} = 3300$  K up to  $T_{eff} = 6000$  K which corresponds to spectral types of mid M to late F type stars. The F3 star is not covered by this temperature range. We extrapolate the spot temperature difference with our linear relation and find a temperature difference of  $\Delta T_{F3} = 2553$  K (contrast of 0.13). For the G2 star we find  $\Delta T_{G2} = 1840$  K (contrast of 0.15) and following our relation for the K5 star we find spot temperature differences  $\Delta T_{K5} = 897$  K (contrast of 0.24).

#### The symmetry parameter

To rate the change in the RV curves as a function of spot temperature we introduce here the symmetry parameter. We measure the maximum and minimum of the RV curve and name the two points  $A$  and  $B$  respectively. Our measure for the symmetry  $s$  is then simply the strength of the minimum  $B$  relative to the maximum  $A$ :

$$s = \frac{-B}{A}. \quad (3.6)$$

The closer the symmetry parameter  $s$  is to 1, the stronger the flux effect. The closer  $s$  approaches 0, the more influence the convective blueshift has. In the upper panel of Fig. 3.13 we show the RV curves for an equatorial spot of 1 % size with different temperatures for the F3 star. The blue dashed line represents the reference case of  $\Delta T = 550$  K. To show the RV curve evolution with decreasing spot temperature we also indicate the RVs for  $\Delta T = 1100$  K as green dashed line (arbitrary chosen intermediate case) and the RVs with the temperature of the spot derived from the model in Fig. 3.12,  $\Delta T_{F3} = 2553$  K, as black solid line. The points  $A$ ,  $B$  are indicated as dots for each curve.

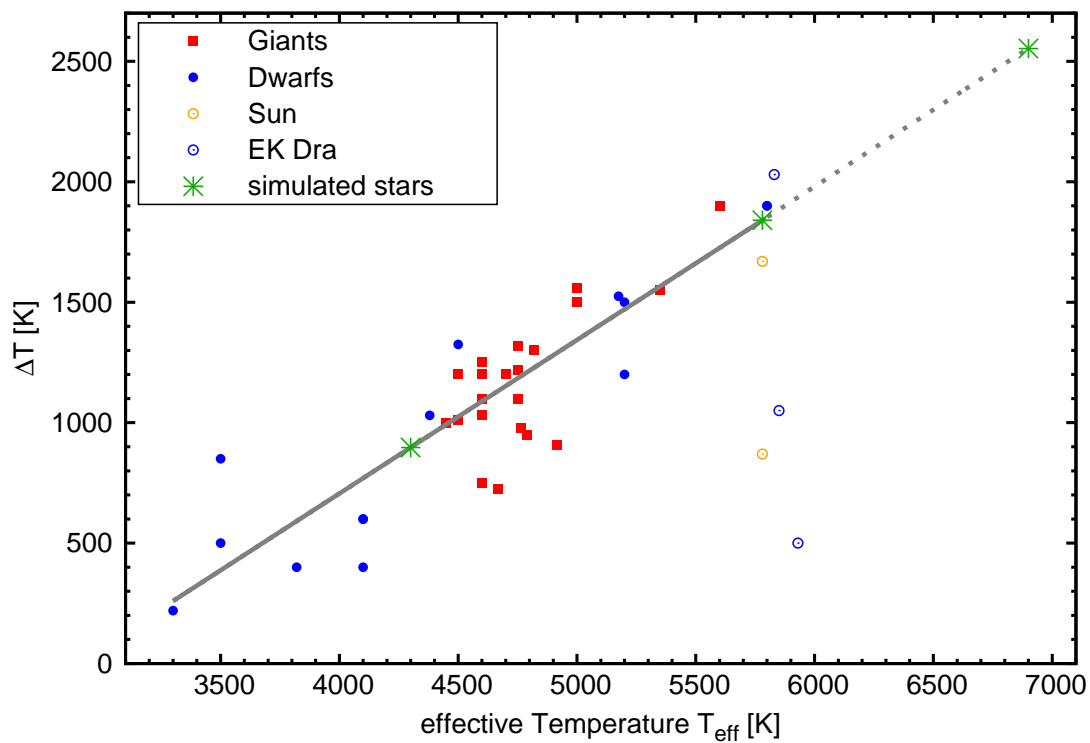


FIGURE 3.12: Measured spot temperatures in red giants (red filled squares) and main sequence dwarfs (blue filled circles). Measurements of the Sun's umbra and penumbra are shown as yellow open circles and EK Dra is shown as blue open circle. All values are taken from [Berdyugina \(2005\)](#). Our fit to the data is indicated as gray line and our simulated stars are indicated as green star symbols.

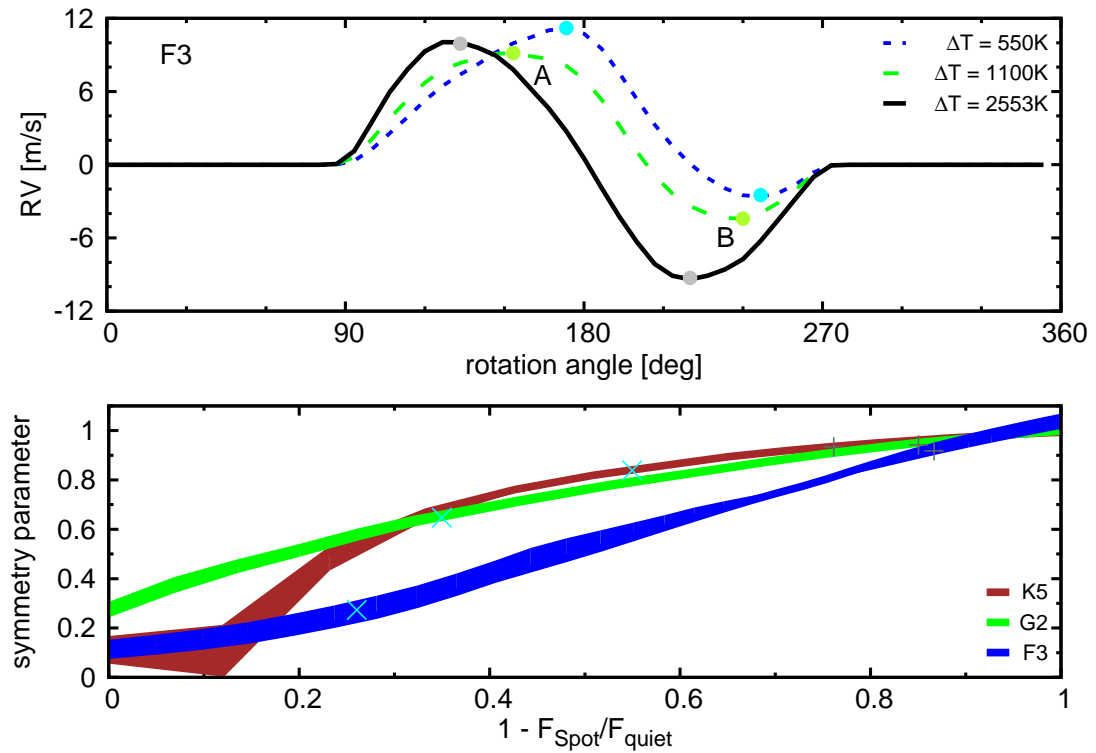


FIGURE 3.13: Top panel: Influence of spot temperature on spot RVs of the F3 star. Calculations done with  $\Delta T = 550\text{ K}$  (blue short dashed line),  $\Delta T = 1100\text{ K}$  (green long dashed), and  $\Delta T = 2553\text{ K}$  chosen according to the spectral type (black solid line). Lower panel: fractional symmetry parameter as a function of spot contrast for spectral types F3 (blue), G2 (green) and K5 (red). The gray crosses mark the contrasts of the spot temperatures derived from Fig. 3.12 for the F3, G2, and K5 stars, and the cyan 'x' symbols mark the contrasts associated to a spot temperature contrast of 550K used in the earlier sections.



From the upper panel of Fig. 3.13 we can see that there is an increase in symmetry for the spot RV curve with decreasing spot temperature (or increasing spot contrast). In the lower panel of Fig. 3.13 we plot the symmetry parameter as a function of the spot contrast. Higher values on the x-axis correspond to larger spot contrasts (or darker spots). The lines seen in the lower panel of Fig. 3.13 are calculated from both FeI lines of the MHD simulations and the area between the two results is shaded. For all three stars simulated in this work we see an increase in symmetry of the RV curves for cooler spots with higher contrasts. However, the symmetry in the spot RV curves is different in F, G, and K stars for the same spot contrast. As explained in Sect. 3.4.1 this is likely related to the different convective speeds for different spectral types. If the spot temperatures from our fit to Fig. 3.12 are used for the RV simulations, the symmetry parameter is above 90 % for all stars (see gray crosses in Fig. 3.13). In these cases the spot contrast is large and the convective blueshift plays only a minor role because the flux effect is dominant.

However, it is not clear whether or not the values of  $\Delta T$  used are typical for the F, G, and K stars. For the G2 star the temperature difference of  $\Delta T_{G2} = 1840$  K corresponds to the Umbra in the Sun. The average temperature difference of spots is lower if the Penumbra is also considered. It is debatable whether observations of spot temperatures on other stars are sensitive mainly to the darker parts of the spots (Berdyugina, 2005). If so, we underestimate the influence of the reduction of convection on the RV curves of spots in other stars.

When RV curves of active stars are modeled there is a degeneracy between active region temperature and size (decreasing the spot temperature and increasing the spot size both increase the RV amplitudes). The symmetry parameter could help to break this degeneracy because it quantifies the interplay between convection and spot temperature which gives the RV curves a characteristic shape. In practice, however, active stars have more than one spot on the surface and the convective blueshift is hard to constrain which will make it difficult to disentangle spot temperatures and sizes.

### 3.5 Summary and conclusion

This work presents simulations of activity-induced RV variations of cold, dark spots and hot, bright plagues on stars of spectral type F3, G2, and K5. In all of our simulations, spots and plagues have the same size, covering 1% of the visible stellar disk and are placed at the equator. Although plagues are observed to cover a larger surface area on the Sun, this setup enables us to investigate a variety of parameters and effects and allows us to compare our results obtained with different model setups.

We investigate the influence of convective blueshift, line shape, and spot temperature on activity-related RV variations on stars other than the Sun. Other approaches in this field use synthetic model spectra without magnetic fields or observed spectra of the Sun to simulate RV curves. In contrast to previous studies we use MHD line profiles of two FeI lines at 6165 and 6173 Å. These line profiles are available at ten different  $\mu$  and four different mean magnetic field strengths from  $B = 0$  G to  $B = 500$  G. In addition these line profiles are available for different stellar types which enables us to investigate the differences in the activity RV curves on different stars.

Magnetic fields are important when RV curves of spots and plagues are investigated. If the magnetic field is strong enough it hinders the supply of new material into the active region and suppresses convection. The result is an apparent redshift of the active region relative to the surrounding quiet disk which results in a bias of the RV curves

towards positive Doppler velocities. We have calculated four spot and plage models with magnetic fields of  $B = 0$  G,  $B = 20$  G,  $B = 100$  G and  $B = 500$  G inside the active regions. All models using low magnetic fields of up to 100 G show no significant signs of reduced convective motions in the active region and the flux effect dominates the RV curves. However, our simulations using a mean magnetic field of 500 G inside the active region produce significantly asymmetric RV curves because convection is suppressed.

For F and G stars the suppression of the convective blueshift is significant and easily seen in RV curves of both spots and plages. The amplitudes of plage RV signals, especially, increase with increasing magnetic field strength in F and G stars. Increasing the magnetic field from 0 G to 500 G in the plage increases the RV amplitude from 0.5 m/s to 15.5 m/s in the F star and from 0.6 m/s to 5 m/s in the G star. For K stars, however, the magnetic field does not influence the shape of the RV curves significantly. As the convective blueshift of spectral lines is very weak, the suppression of convection has only a weak influence on the RV curves in the K5 star. This leads to a symmetric spot RV signal and a negligible plage signal.

Convective blueshift is not constant across the disk and we investigate the influence a variable convective blueshift has on the RV curves of spots and plages. We find that using symmetric Voigt profiles in combination with a constant convective blueshift does not well reproduce the results from the MHD line profiles. Differences are as large as 3.5 m/s for the spot on the F star and 5.5 m/s for the plage. Including a variable convective blueshift for the Voigt profile simulation improves the match with the RV curves derived from the MHD line profiles. However, especially for the plage simulations of the F and G stars, the maximum differences remain large. As a general trend we find that differences between the Voigt and MHD simulations are larger for hotter stars.

As we placed the Voigt profiles at the wavelength measured for the MHD line centers, there is a possibility that the remaining differences between the simulations originate from line profile differences. Voigt profiles are symmetric while MHD profiles have a red tail due to fast downstreams which create line asymmetries. As the projected velocity fields change across the disk, the MHD line profiles change with limb angle and we suspect these line profile changes to be the reason for the differences seen in simulations computed with Voigt and MHD line profiles. However, the differences between simulations with MHD and Voigt profiles are larger for the plages than for the dark spots. We conclude that this is a result of the active region brightness. For brighter active regions, line profile changes result in larger changes of the RV signal because the line profiles of bright active regions contribute more to the disk integrated spectrum of the star than to the line profiles of dark active regions.

We also tested the influence of spot temperatures on the RV curves. The literature suggests that spots in early-type stars are cooler as compared to the quiet photosphere than in late-type stars. We simulate RV curves with the MHD line profiles and vary the spot temperature (and contrast). We define the symmetry parameter that allows us to measure the influence of the convective blueshift on spot RVs as a function of spot contrast. If spot temperatures according to observational values are used in our simulations we find that the flux effect is dominating the spot RV curve in all stars. Thus we can conclude that for high spot contrasts convective blueshift effects are negligible in RV signals, while for bright plages the situation is reversed and convection effects dominate the RV signatures.



## Chapter 4

# Discussion and conclusion

More than two decades have passed since the first exoplanet was discovered in 1995 (Mayor & Queloz, 1995). Since the first exoplanet detection, major improvements have been made to all detection methods in order to find smaller planets in wider orbits. In the beginning all methods were only sensitive to massive, jovian-like planets in close orbits that are entitled ‘hot Jupiters’. Nowadays the field has advanced to a point where it is possible to detect small, rocky and moderately temperate planets (e.g., Anglada-Escudé et al., 2016, 2014, 2013; Wright et al., 2016; Torres et al., 2015). However, a true Earth analogue, with a similar mass, orbit and host star is still beyond reach with current technology (Fischer et al., 2016). To advance further and find planets similar to Earth all detection methods need to undergo improvements. For the RV method, this requires finding ways to overcome the current the m/s precision level and advance towards the cm/s level. Another major challenge with this ultra precise instruments operating at the cm/s level will then be astrophysical noise that hinders the detection of small planets. This thesis has focused on improving the wavelength solution and understanding astrophysical RV jitter originating from stellar spots and plages.

To advance beyond the current m/s precision with the next generation of instruments, better calibration is one key factor for success. RV measurements of stars can only be done with the same precision as the calibration. State-of-the-art instruments achieve m/s precision with HCLs but these standards are insufficient to advance to the cm/s level in the future (Fischer et al., 2016). HCLs emit strong gas lines that saturate the detector, have a limited amount of lines and spectral regions without features. To overcome the limitations of HCLs, this thesis has investigated FPIs for the use in wavelength calibration.

FPIs produce equidistantly spaced spectral lines with almost homogeneous amplitudes. The spacing of the lines depends on the distance between the two mirrors of the device and thus FPIs can be designed according to the spectrograph resolution. This adjustment flexibility makes FPIs very appealing calibrators. However, FPIs have a disadvantage as calibrators: the wavelengths of the FPI lines are poorly constrained because the cavity width is not known and moreover also not constant but wavelength dependent (Wildi et al., 2010). Calibrating a spectrograph without knowing the reference is not possible and hence FPIs were not considered as standards until now. This thesis proposed a solution for this shortcoming of FPIs. FPIs can be referenced to absolute standards like HCLs to calibrate the cavity width with sufficient accuracy to use the interference peaks in wavelength calibration.

The HARPS and CARMENES projects both test FPIs. This thesis characterized the FPIs of these instruments with the HCL referencing method. When the FPIs are used in wavelength calibration small systematics on the order of a few 10 m/s were discovered in both instruments. These systematics can not be seen with HCLs and are only noticeable when enough, well known calibration lines are available. This clearly

demonstrates that it is possible to advance below the calibration limit of HCLs when FPIs are used. HARPS also has a LFC installed. A direct comparison between the FPI and the LFC was performed and only minor differences in the wavelength solution were observed. This shows the abilities of FPIs in wavelength calibration and demonstrates that FPIs are high precision standards.

At the present day there is only one RV survey that uses FPIs on a daily bases. The CARMENES project is pioneering the FPI wavelength solution approach that is proposed in this work. CARMENES uses two FPIs to calibrate a visual and near-infrared spectrographs. The CARMENES calibration strategy is an important test case to demonstrate that FPIs will not only be useful in the future but have already become valuable calibration devices for today's spectrographs. For some upcoming RV instruments FPIs are discussed (e.g., SPIROU [Artigau et al., 2011](#) or HPF [Halverson et al., 2014](#)). The FPI approach used for HARPS and CARMENES is simple to use and easy to implement also for instruments in the near future. Calibration lessons learned from HARPS and CARMENES can potentially already push the detection limits of current and near future instruments below the m/s level. Until the next generation of instruments are build for cm/s precision, FPIs might already be established as standard calibration sources for high precision RV work.

When the next generation of RV instruments will approach the cm/s precision level, new challenges will emerge. While currently RV measurement errors are dominated by instrumental systematics, planet detection with the next generation of instruments will mainly be limited by astrophysical noise ([Fischer et al., 2016](#)). Solar type stars produce a number of phenomena that cause what is known as RV jitter. These RV signals are related to magnetic fields. Activity features on the stellar surface, like spots or plages, can hamper the detection of small planets in the system. Contrary to the instrumental effects, astrophysical noise cannot be suppressed by improving any hardware parts. Astrophysical noise needs to be understood and modeled to reveal small, hidden planetary signals in the RV curves.

Modeling RV jitter with an accuracy of 10 cm/s is challenging and requires knowledge about the underlying physical phenomena of the activity signal. Different methods to simulate RV curves originating from active regions exist (e.g., [Saar & Donahue, 1997](#); [Hatzes, 2002](#); [Meunier et al., 2010](#); [Dumusque et al., 2014](#)). Data driven models use observational data of the Sun ([Dumusque et al., 2014](#)). Spatially resolved spectra taken from active and inactive regions on the solar surface are used to set up simulations and compute the RV jitter. This approach has the advantage that it automatically includes the physical processes on the surface that lead to RV jitter. However, this method also has downsides. Measurement errors in the observed spectra (noise or uncertainties in the wavelength scale; [Reiners et al., 2016b](#)) influence the derived RV curves. While data driven models might be accurate for the Sun, they might not represent other stars because their physical properties differ from the solar case. Moreover, in data driven models the physical processes in active regions that lead to the RV jitter are not investigated.

Other approaches use stellar model atmospheres like the ones of [Kurucz \(1993\)](#) or [Husser et al. \(2013\)](#) to construct simulations of RV jitter (e.g., [Desort et al., 2007](#); [Herrero et al., 2016](#)). These model atmospheres provide a physical understanding of the emerging spectra in the quiet photosphere but do not cover the physical properties of active regions. Temperature changes and inhibition of convection driven by the strong magnetic fields inside active regions are not simulated in the model atmospheres. These processes are crucial to understand RV jitter but the temperature and the reduction

of convective velocities in active regions need to be artificially introduced in the simulations. The temperature and inhibition of convection can be constrained by solar observations but this again limits the applicability of the RV jitter models to the solar case.

This thesis uses a different approach to improve the current understanding of activity-related RV signals. The study presented in this work uses 3D magneto-hydrodynamic models of F, G and K stars from [Beeck et al. \(2013b, 2015b\)](#). Line profiles of FeI lines computed with different magnetic field strength were used to model the RV signatures of spots and plages that rotate across the star. The magnetic fields in the magneto-hydrodynamic simulations hinder convection and thus, the RV models presented in this thesis are not bound to the free parameter of convective blueshift. However, the magnetic field strength available in the magneto-hydrodynamic simulations are not sufficient to force temperature changes. Therefore, the temperature of spots is taken from observations of other stars.

Comparing the RV signatures derived with the magneto-hydrodynamic line profiles to data driven models yields only small differences. Hence, the RV jitter models presented in this thesis agree well with what is expected for the Sun. Since the magneto-hydrodynamic simulations of [Beeck et al. \(2013b, 2015b\)](#) have been also performed for stars other than the Sun, it is possible to explore RV jitter for F and K type stars as well. The properties of convection as well as the temperatures of active regions differ among stars of different spectral types ([Beeck et al., 2013a,b, 2015a,b](#); [Berdyugina, 2005](#)). The models presented in this thesis take into account the different properties of stars and thus give some insights into the RV signals from spots and plages for a broader range of stars. It turns out that RV jitter is produced mainly by dark spots in late type stars while in early type stars bright plages significantly contribute to RV jitter. Because the field of exoplanet science is currently moving towards later spectral types it is important to understand the origin of activity related signals also for these cool stars.

Although the models of RV jitter in this thesis has shown that activity signals differ among stars of different spectral types, the results presented are just a first step. Magneto-hydrodynamic simulations are computationally expensive and only single line profiles can be computed. RV observations use a wide spectral range to achieve high precision. Thus, RV jitter simulations must be performed over the same wavelength range as the measurements ([Desort et al., 2007](#)). Magneto-hydrodynamic simulations of a wider spectral range would therefore be desirable because different spectral lines respond differently to the presence of magnetic fields. Stronger magnetic fields are also required in the magneto-hydrodynamic simulations. This would enable temperature changes in the magneto-hydrodynamic simulations. Then the RV jitter simulations would not rely on observational constraints of spot temperatures which is desirable because it is not clear whether the measured spot temperatures are representative.

Although the models of RV jitter are becoming more detailed, a comparison of the models to high precision RV measurements has yet to be performed. Currently, the first projects to make RV observations of the Sun as a star are underway. [Dumusque et al. \(2015\)](#) presents a solar tracker for HARPS-N which will provide state-of-the-art high precision RV data. RV observations of the Sun are the first step to understand RV jitter in detail. The Sun is currently the only star for which high-resolution images of the stellar surface are possible. RV jitter simulations can use spot maps derived from these high-resolution images. The resulting model RV curves can then be compared and calibrated to the observations of the Sun. Such a comparison would provide valuable information about the current state of the physical understanding of RV jitter.



Models can be improved and extended to other stars which in turn will help to reduce astrophysical noise in RV curves when hunting for true Earth analogues.

# Acknowledgements

- Thank you Ansgar for your support, advice and guidance throughout my years in Göttingen. The friendly atmosphere you created in the working group has led to interesting discussions and made the work enjoyable. I always felt welcome and understood in your office.
- Thank you Sandra and Uwe. The discussions during the GRK meetings were always useful for me. Your guidance and support over the years helped in finishing this thesis.
- Thank you Mathias for the countless discussions and your help whenever I needed it. Working with you was enjoyable and has certainly improved my understanding in the research field. Thank you also for your feedback on this thesis which has substantially improved the content.
- Thank you Philipp for explaining me the principles of the laser frequency comb and for correcting my thesis so carefully. The comments you gave me improved this work a lot. I enjoyed working with you.
- Thanks to everybody in the working group and beyond for all the useful discussions in the social room or my office. Many ideas in this work were motivated by these small chats during the coffee breaks.
- Great thanks also to my parents and family who have always supported me. I was always happy to visit and the time I could spend with you recharged my batteries.
- Cheers to all of my friends with whom I have spent so many great moments. You guys really made life enjoyable ;-)
- I acknowledge support from the Deutsche Forschungsgemeinschaft under DFG GrK 1351. Part of this work was supported by the ERC Starting Grant Wavelength Standards.



# Curriculum vitae

Name : Florian Franziskus Bauer  
Date of Birth : 10<sup>th</sup> August 1989  
Place of Birth : Oberwart (Austria)  
Nationality : Austria

## Education

---

- 09/2013 – 12/2016 | PhD studies at Institute for Astrophysics  
Georg-August-University, Göttingen (Germany)  
Thesis title: 'Radial velocities in low mass stars: improving the wavelength solution of astronomical spectrographs and understanding stellar noise'
- 10/2010 – 07/2013 | Master of science in Astronomy  
University of Vienna (Austria)  
Thesis title: 'A search for transiting planets orbiting Kepler K-giants'
- 10/2007 – 08/2010 | Bachelor of science (Bakk. rer. nat) in Astronomy  
University of Vienna (Austria)  
Thesis title: 'Planetentstehung in Doppelsternsystemen'

## Publications

---

1. [F. F. Bauer](#), M. Zechmeister and A. Reiners, 'Calibrating echelle spectrographs with Fabry-Pérot etalons', 2015A&A...581A.117B
2. A. Reiners, U. Lemke, [F. F. Bauer](#), B. Beeck, P.Huke, 'Radial velocity observations of the 2015 Mar 20 eclipse - A benchmark Rossiter-McLaughlin curve with zero free parameters', 2016A&A...595A..26R

## Conference Contributions

---

- IAU General Assembly, Honolulu, 3<sup>rd</sup> - 14<sup>th</sup> August 2015  
Poster: 'Wavelength calibration with Fabry-Pérot etalons - yes we can'
- Extreme precision radial velocities, Yale, 5<sup>th</sup> - 8<sup>th</sup> July 2015  
Poster: 'Calibrating echelle spectrographs using Fabry-Pérot interferometers'
- Towards other Earth II, Porto, 15<sup>th</sup> - 19<sup>th</sup> September 2014  
Poster: 'Improving the spectrograph wavelength solution with Fabry-Pérot interferometers'



# Bibliography

- Anderson, J. & King, I. R. 2000, *PASP*, 112, 1360
- Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, *Nature*, 536, 437
- Anglada-Escudé, G., Arriagada, P., Tuomi, M., et al. 2014, *MNRAS*, 443, L89
- Anglada-Escudé, G., Tuomi, M., Arriagada, P., et al. 2015, *ArXiv e-prints*
- Anglada-Escudé, G., Tuomi, M., Gerlach, E., et al. 2013, *A&A*, 556, A126
- Artigau, É., Donati, J.-F., & Delfosse, X. 2011, in *Astronomical Society of the Pacific Conference Series*, Vol. 448, 16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. C. Johns-Krull, M. K. Browning, & A. A. West, 771
- Baglin, A. & COROT Team. 1998, in *IAU Symposium*, Vol. 185, *New Eyes to See Inside the Sun and Stars*, ed. F.-L. Deubner, J. Christensen-Dalsgaard, & D. Kurtz, 301
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, *A&AS*, 119, 373
- Basri, G., Borucki, W. J., & Koch, D. 2005, *New A Rev.*, 49, 478
- Bauer, F. F., Reiners, A., Beeck, B., & Jeffers, S. V. 2017, *A&A*, in press
- Bauer, F. F., Zechmeister, M., & Reiners, A. 2015, *A&A*, 581, A117
- Beeck, B., Cameron, R. H., Reiners, A., & Schüssler, M. 2013a, *A&A*, 558, A48
- Beeck, B., Cameron, R. H., Reiners, A., & Schüssler, M. 2013b, *A&A*, 558, A49
- Beeck, B., Schüssler, M., Cameron, R. H., & Reiners, A. 2015a, *A&A*, 581, A42
- Beeck, B., Schüssler, M., Cameron, R. H., & Reiners, A. 2015b, *A&A*, 581, A43
- Benedict, G. F., Nelan, E., McArthur, B., et al. 1993, *PASP*, 105, 487
- Berdugina, S. V. 2005, *Living Reviews in Solar Physics*, 2
- Beuzit, J.-L., Feldt, M., Dohlen, K., et al. 2008, in *Proc. SPIE*, Vol. 7014, *Ground-based and Airborne Instrumentation for Astronomy II*, 701418
- Boisse, I., Bonfils, X., & Santos, N. C. 2012, *A&A*, 545, A109
- Borgniet, S., Meunier, N., & Lagrange, A.-M. 2015, *A&A*, 581, A133
- Born, M. & Wolf, E. 1999, *Principles of Optics*
- Bouchy, F., Pepe, F., & Queloz, D. 2001, *A&A*, 374, 733
- Butler, R. P., Bedding, T. R., Kjeldsen, H., et al. 2004, *ApJ*, 600, L75
- Butler, R. P., Marcy, G. W., Williams, E., et al. 1996, *PASP*, 108, 500



- Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, *ApJ*, 529, L45
- Cochran, W. D., Smith, H. J., & Smith, W. H. 1982, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 331, *Instrumentation in Astronomy IV*, 315
- Crosswhite, H. M., Dieke, G. H., & Legagneur, C. S. 1955, *J. Opt. Soc. Am.*, 45, 270
- de Boor, C. 2001, *A Practical Guide to Splines*, *Applied Mathematical Sciences* (Springer New York)
- Degraffenreid, W. & Sansonetti, C. J. 2002, *Journal of the Optical Society of America B Optical Physics*, 19, 1711
- Desort, M., Lagrange, A.-M., Galland, F., Udry, S., & Mayor, M. 2007, *A&A*, 473, 983
- Dierckx, P. 1995, *Curve and Surface Fitting with Splines*, *Monographs on numerical analysis* (Clarendon Press)
- D'Odorico, V. & CODEX/ESPRESSO Team. 2007, *Mem. Soc. Astron. Italiana*, 78, 712
- Dumusque, X., Boisse, I., & Santos, N. C. 2014, *ApJ*, 796, 132
- Dumusque, X., Glenday, A., Phillips, D. F., et al. 2015, *ApJ*, 814, L21
- Dumusque, X., Pepe, F., Lovis, C., et al. 2012, *Nature*, 491, 207
- Dumusque, X., Udry, S., Lovis, C., Santos, N. C., & Monteiro, M. J. P. F. G. 2011, *A&A*, 525, A140
- Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, *PASP*, 128, 066001
- Follert, R., Dorn, R. J., Oliva, E., et al. 2014, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 9147, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 19
- Foukal, P. 2008, *Solar Astrophysics* (Wiley)
- Gray, D. F. 2009, *ApJ*, 697, 1032
- Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, *AJ*, 132, 161
- Griffiths, P. R. & de Haseth, J. A. 2007, *Fourier Transform Infrared Spectrometry* (Wiley)
- Halverson, S., Mahadevan, S., Ramsey, L., et al. 2014, in *Proc. SPIE*, Vol. 9147, *Ground-based and Airborne Instrumentation for Astronomy V*, 91477Z
- Hatzes, A. P. 2002, *Astronomische Nachrichten*, 323, 392
- Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, *ApJ*, 529, L41
- Herrero, E., Ribas, I., Jordi, C., et al. 2016, *A&A*, 586, A131
- Huke, P., Tal-Or, L., Sarmiento, L. F., & Reiners, A. 2016, *Hollow-cathode lamps as optical frequency standards: the influence of optical imaging on the line-strength ratios*
- Husser, T.-O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, 553, A6

- Isobe, T., Feigelson, E. D., Akritas, M. G., & Babu, G. J. 1990, *ApJ*, 364, 104
- Jeffers, S. V., Barnes, J. R., Jones, H. R. A., et al. 2014, *MNRAS*, 438, 2717
- Kalas, P., Graham, J. R., Chiang, E., et al. 2008, *Science*, 322, 1345
- Kerber, F., Nave, G., & Sansonetti, C. J. 2008, *ApJS*, 178, 374
- Kotani, T., Tamura, M., Suto, H., et al. 2014, in *Proc. SPIE*, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 914714
- Küker, M. & Rüdiger, G. 2011, *Astronomische Nachrichten*, 332, 933
- Kurucz, R. 1993, ATLAS9 Stellar Atmosphere Programs and 2 km/s grid. Kurucz CD-ROM No. 13. Cambridge, Mass.: Smithsonian Astrophysical Observatory, 1993., 13
- Lagrange, A.-M., Desort, M., & Meunier, N. 2010, *A&A*, 512, A38
- Lanza, A. F., Bonomo, A. S., Moutou, C., et al. 2010, *A&A*, 520, A53
- Latham, D. W., Stefanik, R. P., Mazeh, T., Mayor, M., & Burki, G. 1989, *Nature*, 339, 38
- Latham, D. W., Stefanik, R. P., Mazeh, T., Torres, G., & Carney, B. W. 1998, in *Astronomical Society of the Pacific Conference Series*, Vol. 134, Brown Dwarfs and Extrasolar Planets, ed. R. Rebolo, E. L. Martin, & M. R. Zapatero Osorio, 178
- Lovis, C. & Pepe, F. 2007, *A&A*, 468, 1115
- Ludwig, H.-G., Caffau, E., Steffen, M., et al. 2009, *Mem. Soc. Astron. Italiana*, 80, 711
- Mahadevan, S., Ramsey, L., Wright, J., et al. 2010, in *Proc. SPIE*, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 77356X
- Marcy, G. W. & Butler, R. P. 1992, *PASP*, 104, 270
- Mayor, M., Pepe, F., Queloz, D., et al. 2003, *The Messenger*, 114, 20
- Mayor, M. & Queloz, D. 1995, *Nature*, 378, 355
- McMillan, R. S., Moore, T. L., Perry, M. L., & Smith, P. H. 1994, *Ap&SS*, 212, 271
- McMillan, R. S., Perry, M. L., Smith, P. H., & Merline, W. J. 1988, in *Astronomical Society of the Pacific Conference Series*, Vol. 3, Fiber Optics in Astronomy, ed. S. C. Barden, 237–246
- Meunier, N., Desort, M., & Lagrange, A.-M. 2010, *A&A*, 512, A39
- Meunier, N., Lagrange, A.-M., Borgniet, S., & Rieutord, M. 2015, *A&A*, 583, A118
- Meunier, N., Lagrange, A.-M., Mbemba Kabuiku, L., et al. 2017, *A&A*, 597, A52
- Molaro, P., Esposito, M., Monai, S., et al. 2013, *A&A*, 560, A61
- Neely, A. W. & Janesick, J. R. 1993, *PASP*, 105, 1330
- Ortiz, A., Solanki, S. K., Domingo, V., Fligge, M., & Sanahuja, B. 2002, *A&A*, 388, 1036
- Palmer, B. A. & Engleman, R. 1983, *Atlas of the Thorium spectrum*

- Pasquini, L., Avila, G., Dekker, H., et al. 2008, in Proc. SPIE, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, 70141I
- Perryman, M. 2011, *The Exoplanet Handbook*
- Piskunov, N. E. & Valenti, J. A. 2002, *A&A*, 385, 1095
- Press, W. 2007, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University Press)
- Probst, R. 2015, PhD thesis
- Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2011, in IAU Symposium, Vol. 276, IAU Symposium, ed. A. Sozzetti, M. G. Lattanzi, & A. P. Boss, 545–546
- Radick, R. R., Mihalas, D., Lockwood, G. W., et al. 1983, *PASP*, 95, 621
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Experimental Astronomy*, 38, 249
- Reader, J., Kramida, A., & Ralchenko, Y. 2012, in American Astronomical Society Meeting Abstracts, Vol. 219, American Astronomical Society Meeting Abstracts #219, 443.01
- Redman, S. L., Lawler, J. E., Nave, G., Ramsey, L. W., & Mahadevan, S. 2011, *ApJS*, 195, 24
- Redman, S. L., Nave, G., & Sansonetti, C. J. 2014, *ApJS*, 211, 4
- Reiners, A. 2014, in IAU Symposium, Vol. 302, Magnetic Fields throughout Stellar Evolution, ed. P. Petit, M. Jardine, & H. C. Spruit, 156–163
- Reiners, A., Banyal, R. K., & Ulbrich, R. G. 2014, *A&A*, 569, A77
- Reiners, A., Bean, J. L., Huber, K. F., et al. 2010, *ApJ*, 710, 432
- Reiners, A., Lemke, U., Bauer, F., Beeck, B., & Huke, P. 2016a, ArXiv e-prints
- Reiners, A., Mrotzek, N., Lemke, U., Hinrichs, J., & Reinsch, K. 2016b, *A&A*, 587, A65
- Reiners, A., Shulyak, D., Anglada-Escudé, G., et al. 2013, *A&A*, 552, A103
- Robertson, P., Mahadevan, S., Endl, M., & Roy, A. 2014, *Science*, 345, 440
- Robertson, P., Roy, A., & Mahadevan, S. 2015, *ApJ*, 805, L22
- Saar, S. H. & Donahue, R. A. 1997, *ApJ*, 485, 319
- Schäfer, S. & Reiners, A. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 94
- Schroeder, D. J., ed. 2000, *Astronomical optics*
- Schwab, C., Gurevich, Y., Stuermer, J., et al. 2015, in AAS/Division for Extreme Solar Systems Abstracts, Vol. 3, AAS/Division for Extreme Solar Systems Abstracts, 106.03
- Seager, S. & Deming, D. 2010, *ARA&A*, 48, 631

- Torres, G., Kipping, D. M., Fressin, F., et al. 2015, *ApJ*, 800, 99
- Udry, S., Bonfils, X., Delfosse, X., et al. 2007, *A&A*, 469, L43
- Verschueren, W. 1995, in *European Southern Observatory Conference and Workshop Proceedings, Vol. 53, Calibrating and Understanding HST and ESO Instruments*, ed. P. Benvenuti, 77
- Vogt, S. S., Butler, R. P., Rivera, E. J., et al. 2010, *ApJ*, 723, 954
- Wilken, T., Lovis, C., Manescau, A., et al. 2010, *MNRAS*, 405, L16
- Wildi, F., Pepe, F., Chazelas, B., Lo Curto, G., & Lovis, C. 2010, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, 4
- Wilken, T. 2010, PhD thesis, lmu
- Wilken, T., Curto, G. L., Probst, R. A., et al. 2012, *Nature*, 485, 611
- Wilken, T., Lovis, C., Manescau, A., et al. 2010, *MNRAS*, 405, L16
- Wright, D. J., Wittenmyer, R. A., Tinney, C. G., Bentley, J. S., & Zhao, J. 2016, *ApJ*, 817, L20