

# Resolving stellar atmospheres

## I. The H $\alpha$ line and comparisons to microlensing observations

C. Thurl<sup>1</sup>, P. D. Sackett<sup>1</sup>, and P. H. Hauschildt<sup>2</sup>

<sup>1</sup> Research School of Astronomy and Astrophysics, Australian National University, Mt. Stromlo Observatory, Cotter Road, Weston ACT 2611, Australia  
e-mail: cthurl@mso.anu.edu.au

<sup>2</sup> Sternwarte Hamburg, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany

Received 23 December 2005 / Accepted 2 March 2006

### ABSTRACT

**Context.** We present work on H $\alpha$  spectral line characteristics in PHOENIX stellar model atmospheres and their comparison to microlensing observations.

**Aims.** We examine in detail the H $\alpha$  equivalent width (*EW*) and the line shape characteristics for effective temperatures of 4500 K <  $T_{\text{eff}}$  < 5600 K where H $\alpha$  is a strong spectral feature.

**Methods.** We find that H $\alpha$  *EW* in models calculated under the assumption of local thermodynamic equilibrium (LTE) is up to 15% smaller than in models without this assumption, non-LTE models (NLTE) and that line shapes vary significantly for the two model types. A comparison with available high quality microlensing data, capable of tracing H $\alpha$  absorption across the face of one G5III giant, shows that the LTE model that fits the *EW* best is about 100 K hotter than and the best-fitting NLTE model has a similar  $T_{\text{eff}}$  as predicted by the spectral type analysis of the observed star but agree within the uncertainties of the observationally derived temperature.

**Results.** Neither LTE nor NLTE models fit the line shape well. We suspect unmodelled chromospheric emission. Line shape diagnostics suggest lower gravities than derived for the star and are unacceptable low in the case of the LTE models. We show that *EW* alone is insufficient for comparison to stellar model atmospheres, but combined with a new shape parameter we define is promising. In stellar parameter ranges where the H $\alpha$  line is strong, a NLTE approach of modeling stellar atmospheres is not only beneficial but mandatory.

**Key words.** stars: atmospheres – gravitational lensing

## 1. Introduction

Most of our basic knowledge about the interior physics of stars, even of the Sun, is based on our understanding of stellar atmospheres. The modeling of such atmospheres allows us to derive the star's fundamental characteristics such as temperature, gravity, chemical composition and age. However many outstanding questions still need to be answered about stellar atmospheres. The “solar model problem”, i.e., the disagreement between model predictions for the neon-oxygen abundance ratio and measurements by helioseismology, is hotly debated (Drake & Testa 2005; Asplund et al. 2003). Derivation of elemental abundances is based on the accuracy of stellar model atmospheres and as such make and shake our knowledge not only about stellar, but also about galactic and cosmic evolution (Asplund 2003).

The main difficulty in constraining stellar atmosphere models with observations is that most techniques only measure the disk-integrated flux or disk-integrated spectra of stars. The biggest advance in one-dimensional (1D) atmospheres modeling has been to introduce consistent treatment of Non-Local Thermodynamic Equilibrium (NLTE) processes (Anderson 1989; Hauschildt & Baron 1999). However observational tests of NLTE effects have been difficult to establish, because studying atmospheres of stars more distant than our Sun has proven difficult and tedious.

Various methods have been tried to resolve stars, i.e. measuring their sizes and resolving surface structure. Direct observations have been restricted to the Sun (e.g. Blackwell et al. 1995) or, when using the Hubble Space Telescope (HST), to nearby supergiants (e.g. Gilliland & Dupree 1996). Doppler imaging (e.g. Oláh et al. 2003) produces observed center-to-limb variations of the integrated flux that are in good agreement with values derived from current LTE model atmospheres. Whereas Interferometry (e.g. Aufdenberg et al. 2005) show less limb darkening than predicted by stellar atmospheres models. Due to the uneven surface of the moon, surface measurements of nearby giants via lunar occultations (Richichi & Lisi 1990) have not yet yielded sufficient precision to distinguish between various model atmospheres. O'Donoghue et al. (2003) and Popper & Guinan (1998) have used eclipsing binaries to study surface structure. In such systems both stars are often in a similar evolutionary state, which complicates tremendously the decomposition of the single star spectra. Furthermore, if one star is evolved and has (almost) filled its Roche Lobe so that mass transfer occurs, then neither star is a good probe for single-star atmospheres. However all of these methods are restricted to nearby stars.

Microlensing was proposed to offer an elegant solution to the problem of observing distant stellar atmospheres (Schneider et al. 1992). During a microlensing event a differential magnification pattern across the face of the source star enables the observer to resolve its surface (see Sect. 5 for more

**Table 1.** Observed microlensing events with characteristics and results. In the first column are listed microlensing events discussed in greater detail by the authors in the second column. The last column lists the main findings for the events, i.e.  $H\alpha$  equivalent width variations and limb darkening coefficients were measured in the respective events.

Event Name	Publication	results
MACHO Alert 95-30	Alcock et al. (1997)	$H\alpha$ <i>EW</i>
MACHO 97-BLG-28	Albrow et al. (1999a)	limb darkening
MACHO 97-BLG-41	Albrow et al. (2000)	limb darkening
OGLE-1999-BUL-23	Albrow et al. (2001a)	limb darkening
MACHO 98-SMC-1	Albrow et al. (1999b)	limb darkening
EROS BLG-2000-5	Albrow et al. (2001b)	limb darkening
	Castro et al. (2001)	$H\alpha$ <i>EW</i>
	Fields et al. (2003)	
MOA 2002 -BLG-23	Abe et al. (2003)	limb darkening
OGLE-2002-BUL-069	Cassan et al. (2004)	$H\alpha$ <i>EW</i>

detail). Several authors have studied the effect theoretically. Valls-Gabaud (1998) presented predictions of spectroscopic signatures of microlensing on spectral lines like  $H\beta$  and  $\text{CO } 2.3 \mu\text{m}$  and proposed microlensing as a unique test for stellar atmosphere models. Heyrovský & Sasselov (2000) studied how microlensing can be used to detect stellar spots on the surfaces of stars. Heyrovský et al. (2000) demonstrated the effects of microlensing on synthesized optical spectra of red giant model atmospheres. Heyrovský (2003) discussed the difficulty of microlensing light curve inversion methods. Bryce et al. (2002) and Hendry et al. (2002) examined the detectability of star spots via microlensing and their possible effects on multicolor microlensing light curves. Bryce et al. (2003) examined the change in line profiles due to bulk motion in circumstellar envelopes during microlensing fold caustic crossing events.

An early attempt by Lennon et al. (1996, 1997) used the magnification effect of the microlensing event 96-BLG-3 for spectral studies of stellar atmospheres. They further suggested an effect of center to limb variations on the  $H\alpha$  line profile during such an event. A number of well-studied microlensing events have resulted in measurements that resolve the source star; we discuss one of these in more detail in Sect. 5.1. The names and major characteristics of these events are listed in Table 1, with the respective publications. The major findings were promising comparisons of observationally-deduced limb darkening coefficients with model atmospheres and variations of equivalent width (*EW*) of the  $H\alpha$  line. One of these events, EROS BLG-2000-5, is a cusp-crossing event, which shows that despite the increased complexity in modeling not only fold caustic crossings can be used for these purposes. For the most recent event, OGLE-2002-BUL-69 (Cassan et al. 2004; Kubas et al. 2005), a well-sampled light curve and high resolution spectra are available at crucial times during the event, making it a good target with which to compare high resolution model atmospheres. Four events of similar quality have been observed in 2004 and another four in 2005<sup>1</sup> and many more are to be expected in the years to come, making microlensing a viable future for testing stellar model atmospheres.

This is the first in series of papers aimed at developing methods to test stellar atmosphere models by confronting them with observations that resolve stellar surfaces photometrically or spectroscopically. Based on the most current theoretical models, we describe what observations can be made in the future in order

to distinguish between various stellar models and learn the most from comparisons to observations.

We lead into this work by describing, in Sect. 2, our methods and the PHOENIX stellar models. In Sect. 3 we explain the reasons for and analysis techniques of studying the  $H\alpha$  line. In Sect. 4, we introduce the spectral parameter range studied and show that spectral differences between LTE and NLTE cannot be ignored within the  $H\alpha$  line for stars with  $4500 \text{ K} < T_{\text{eff}} < 5600 \text{ K}$ . How microlensing can be used as a tool to resolve the surfaces of stars is described in Sect. 5, where we also present results of our comparison of model atmospheric data to microlensing  $H\alpha$  *EW* line data for the event OGLE-BULGE-2002-069.

In Sect. 6, we illustrate how to most appropriately choose the best-fitting model out of a set of models for a given parameter or set of parameters. In Sect. 7, we introduce a new line shape diagnostic and repeat the analysis similar to the *EW* analysis. Our conclusions are summarized in Sect. 8.

## 2. Methodology

### 2.1. Stellar models

We study the state-of-art atmospheric models produced by the multi-purpose code PHOENIX (Hauschildt et al. 1999a,b). These models are static and assume one-dimensional, spherically-symmetric radiative transfer. We have chosen PHOENIX because its models can be calculated for all stellar parameters from dwarfs to giants, for both static and variable stars. A major advantage of PHOENIX models is that the code can include self-consistently NLTE effects. LTE assumes that, within a confined region, absorption and emission of photons occurs at the same rate and the underlying region’s temperature structure follows a black body. NLTE models relax this assumption so that atomic physics may produce situations in which local areas need not be in thermal equilibrium. Spectra calculated from the PHOENIX models can be calculated for almost any resolution ranging from the UV to radio wavelengths.

In order to compare stellar model atmospheres to the observations of real stars we must “transform” the models into observables by simulating the effects of external phenomena on the spectra. Microlensing, for example see Sect. 5, would be such a transformation, as it occurs externally to the source star, but does affect the light collected by the observer as a function of time. Convolution of model data with this transformation produces synthetic data, which we can then analyze in the same way as we would analyze observational data. We conduct a comparison between PHOENIX LTE and NLTE synthetic spectra as well as a comparison to high resolution observational spectral data from a microlensing event.

## 3. $H\alpha$ and its equivalent width

In this work we focus on the effects that can be studied with observations of the  $H\alpha$  line. The  $H\alpha$  line is a strong line for stars with  $4500 < T_{\text{eff}} < 5600 \text{ K}$  and is therefore a frequent observational target. Different parts of the  $H\alpha$  line are formed in different regions of a stellar atmosphere, some which are more affected by NLTE than others (Gray 1992). This makes  $H\alpha$  a good candidate to study possible differences between LTE and NLTE modeling by testing these differences against observations. Another incentive to study  $H\alpha$  is that published data are available for the resolved source star during some microlensing events (Albrow et al. 2001b; Cassan et al. 2004). In this section, we concentrate on the *EW* of the line as a figure of merit with which to compare

<sup>1</sup> <http://planet.iap.fr>

LTE with NLTE effects, and models with observations. We adopt the usual definition of  $EW$ :

$$EW = \int_{\text{line}} 1 - \frac{F_{\text{line}}(\lambda)}{F_{\text{cont}}(\lambda)} d\lambda \quad (1)$$

where  $F_{\text{line}}$  is the radiant flux as a function of wavelength  $\lambda$ , and  $F_{\text{cont}}$  the value of the flux from the continuous spectrum outside the spectral line. The analytical potential of  $EW$  as a tool for studying atmospheric models depends on one's ability to measure its value consistently for both model and observational spectra, which requires a careful choice of the continuum and the rejection of spectral blends. In order to reduce systematically-introduced effects to the analysis, we apply the same algorithm to all our model and observational data.

### 3.1. Extracting equivalent width from model spectra and data

For both model and observational data, the  $EW$  is determined as follows. We first apply a standard vacuum-to-air conversion to all model data. We then remove blended spectral lines and from this set of model/data points we determine the continuum around the spectral line that we wish to analyze and normalize the spectra. Note that by “continuum”, we mean the “pseudo continuum” determined by fitting, as there is no way of determining the *real continuum* for spectral line observations. After smoothing the data, we obtain the  $EW$  by spline interpolation followed by direct integration.

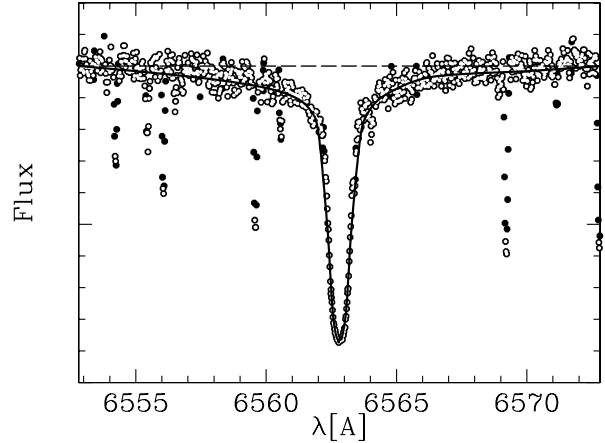
Our clipping algorithm removes lines blended with  $H\alpha$  from the spectral data, which would otherwise contribute to the  $H\alpha$   $EW$  during direct integration. We discard all points from the initial spectral data (whether model or fully-reduced observational data) that differ in flux from that of either of their neighbor points by more than a threshold,  $\mathcal{T}$ , which is determined by the signal-to-noise of the data and its spectral resolution,  $\Delta\lambda$  via,

$$\mathcal{T} = C\sigma^2/\Delta\lambda, \quad (2)$$

where  $\sigma$  is the rms scatter of all flux points about their mean across the wavelength range considered.  $C$  is a constant used to adjust the magnitude of the threshold to ensure optimal clipping. We chose  $C$  by testing our clipping routine on different modeled and observed spectra, and found that a different value of  $C$  must be used within the core of the line in order to not reject spectral points corresponding to the line itself. After clipping, the continuum is then determined from the remaining points.

Figure 1 illustrates how we obtained the  $H\alpha$   $EW$  in both synthetic and observed data. We assume the continuum to be linear over the narrow spectral range considered. We compute the mean flux value of the outermost 1 Å of the spectra on either side of the  $H\alpha$  line, and fit a continuum line through these values. Although the synthetic data are virtually noiseless, we determine the continuum in this way so that the same method can be applied to the noisier observational data. We have tested and confirmed this method by determining the  $EW$  for integration ranges from 1 Å to 10 Å. The slight increase in  $EW$  was linear from 9 Å onwards, indicating that the line signature lies within  $\pm 9$  Å and it is safe to determine the continuum outside of this range.

When smoothing, we divide the spectral data into three separate sections: (a)  $6564.8 \pm 1$  Å to treat the core of the line, (b) the region excluding the core out to  $6564.8 \pm 3$  Å, to cover the wings, and (c) the transition to the continuum of the spectral range of  $6564.8 \pm 10$  Å, excluding the core and the wing regions. Within each of these spectral regions we adapt the sizes of bins over



**Fig. 1.** Illustration of our  $EW$  integration algorithm: black solid points are the model output data around the  $H\alpha$  line with added noise, distributed at random following a Gaussian with a root mean squared (rms) scatter  $\sigma$  of 2.7%. Marked in white are those points chosen by the algorithm to be used for the integration. These are then binned, spline interpolated (black solid line), and integrated up to the continuum (long dashed line) to give the  $EW$  of the line.

which we average so as to sample the continuum less densely than the wings, and the wings less densely than the core. The averaged flux values in each bin are then spline interpolated to give the  $EW$  of the line via direct integration.

## 4. The stellar parameter grid: LTE vs. NLTE

Using this algorithm, we obtained disk-integrated  $H\alpha$   $EW$  values over a grid of 168 PHOENIX LTE and NLTE models. The stellar parameters ranged over a grid of  $2.0 < \log(g) < 5.0$  in steps of 0.5 and  $4500 \text{ K} < T_{\text{eff}} < 5600 \text{ K}$  in steps of 100 K. A metallicity of  $[\text{Fe}/\text{H}] = -0.5$  was assumed throughout.

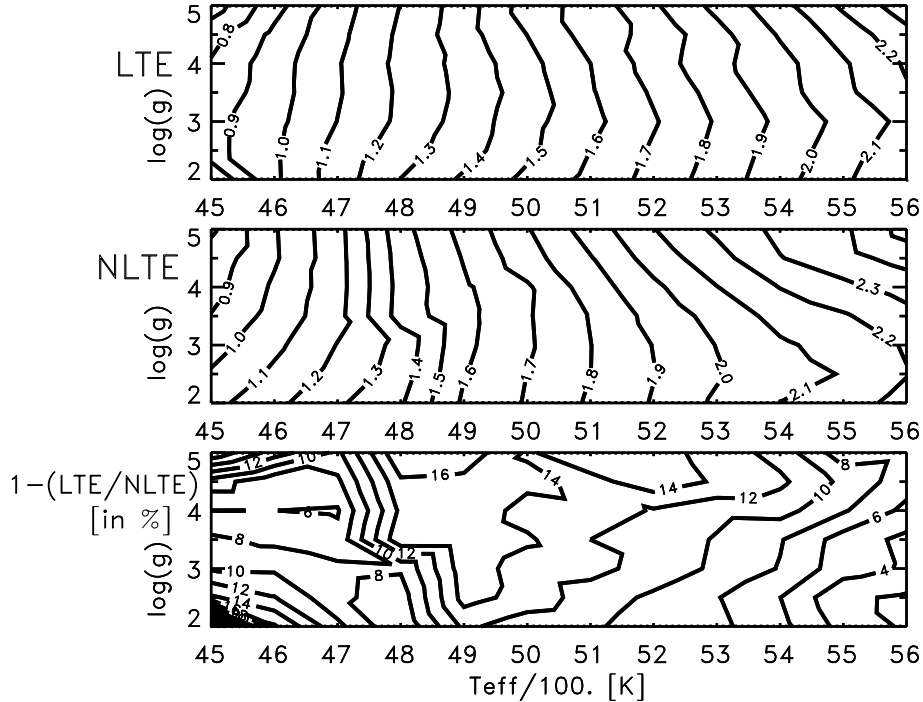
In Fig. 2, we indicate contour lines of equal  $H\alpha$   $EW$  values, the results for both LTE and NLTE across the atmosphere model grid.

One can see immediately that  $H\alpha$   $EW$  is very sensitive to changes in  $T_{\text{eff}}$ , but much less so for changes in gravity. NLTE models with higher  $T_{\text{eff}}$  show a slightly stronger dependence on gravity than do the respective LTE models. Previous studies (e.g. Gray 1992) have shown that NLTE calculations produce a deeper core for the  $H\alpha$  line than in the LTE regime, which we confirm. The bottom panel shows the fractional difference in  $H\alpha$   $EW$  between the two model types.

Over the middle range of the grid, the fractional difference between LTE and NLTE is about 15%. This shows clearly that, within our stellar parameter grid, the difference between NLTE and LTE is significant, and that large systematic errors may compromise any conclusions drawn for the  $H\alpha$  line in this regime if LTE is assumed for its formation. This work also indicates that over the stellar range  $T_{\text{eff}} > 5400 \text{ K}$  and  $\log(g) < 3.0$  a LTE approach may be reasonable for a study of the  $H\alpha$  line, depending on the precision required.

## 5. Microlensing as a test: resolving the surface of stars

So far we have only compared models of stellar atmospheres. Can these differences be observed and, if so, how can this help us to understand better the underlying physics of line formation



**Fig. 2.** *Upper panel:* contours of equal  $H\alpha$  EW (integrated over the face of the stellar disk) over a range of stellar parameters in  $T_{\text{eff}}$  and  $\log(g)$  for LTE atmosphere models. *Middle panel:* same, but for the NLTE atmospheric models. For the stellar parameter range shown, NLTE models have higher EW compared to LTE models at the same  $T_{\text{eff}}$  and  $\log(g)$ , as expected for the NLTE-affected  $H\alpha$  line. *Lower panel:* contours of fractional difference  $1 - (EW_{\text{LTE}}/EW_{\text{NLTE}})$ , given in percent.

and stellar modeling? In particular, do we have ways to distinguish between NLTE and LTE effects at the limb and the center of stars? We might expect to observe larger differences at the limb, where less light originates from deeper layers of the star, as NLTE processes are more important in the outer, less dense regions of the atmosphere. Resolving the stellar surface then becomes crucial. A fairly new tool of resolving the surface of stars farther away than the Sun is microlensing (Bryce et al. 2002). When a stellar object, the lens, crosses the line of sight to a source star and the lens is a binary, an asymmetric magnification pattern, which peaks at the so-called caustic, moves across the face of a source star (see Fig. 3). During such a caustic crossing there is a strongly differential magnification across the face of the microlensed star, which allows its surface to be resolved.

Figure 3 illustrates the microlensing effect on the total integrated light of the star (upper right panel) and on the  $H\alpha$  EW (lower right panel) during a caustic exit. Light from the parts of the star directly coincide with the position of the caustic are most highly magnified, and thus contribute a larger fraction to the integrated light observed. Figure 4 illustrates the different shapes of the  $H\alpha$  spectral line across the face of the star for different values of  $\mu = \cos\theta$ , where  $\theta$  is opening angle to the line-of-sight.  $H\alpha$  is a strong absorption line in the center of the star ( $\mu = 1$ ) and goes into emission near the limb. This emission explains the dip in the  $H\alpha$  EW during the caustic crossing while the limb is highly magnified.

Due to the relative motion of source and lens, the magnification pattern moves across the source star and produces a light curve of integrated stellar light, with a typical shape for microlensing events. The magnification pattern may be rather complex, but in most cases one can make the assumption that the caustic crossing is a *fold caustic crossing*, i.e., that the caustic has no curvature across the face of the star. For what follows, we further assume that the source star is spherically symmetric,

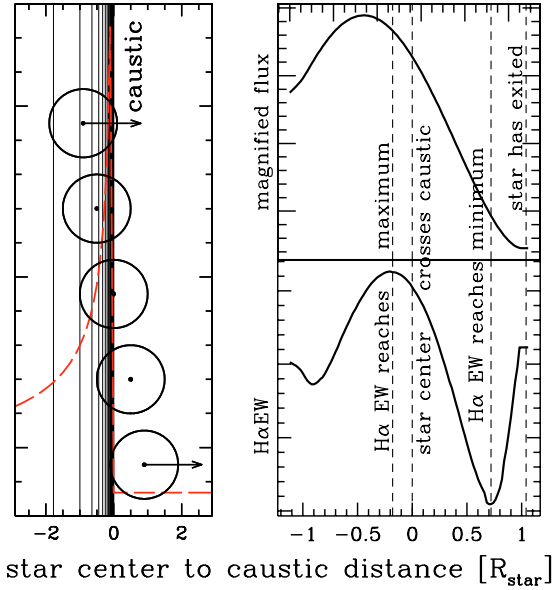
and that the relative motion of source star and caustic is constant and rectilinear. For fold caustics, a five-parameter fit to the light curve data is sufficient to extract the magnification pattern of the event and thus infer the contribution each part of the source star makes to the integrated light. The five parameters describe the time and duration of the event, the maximum magnification, how long it takes to reach maximum magnification and the underlying linear magnification after the caustic crossing (Albrow et al. 1999b). Entry into a caustic is unpredictable, but caustic exits can often be predicted so that proper observations can be put into place. Outside the caustic, typically there is little or no significant differential magnification across the stellar disk and the spectrum is similar to that of the unmagnified star. During the exit, integrated light is substantially influenced by the highly magnified light from the limb, only where differential magnification is important.

### 5.1. Variations of the $H\alpha$ line EW during a fold caustic crossing: a worked example

Because the  $H\alpha$  line is in emission at the limb, the integrated EW of the  $H\alpha$  line drops during a caustic exit compared to the non-differentially magnified case. This occurs first when the leading limb crosses the caustic and experiences maximum magnification. The effect is stronger when the trailing limb crosses the caustic, as now most of the stellar disk has crossed into the low magnification region outside of the caustic and contributes a smaller fraction to the total integrated light than when most of the star still lies in the caustic interior.

During the microlensing event microlensing event OGLE-BULGE-2002-069 the PLANET collaboration<sup>2</sup> was able to obtain high resolution VLT UVES data during the caustic exit

<sup>2</sup> <http://planet.iap.fr>



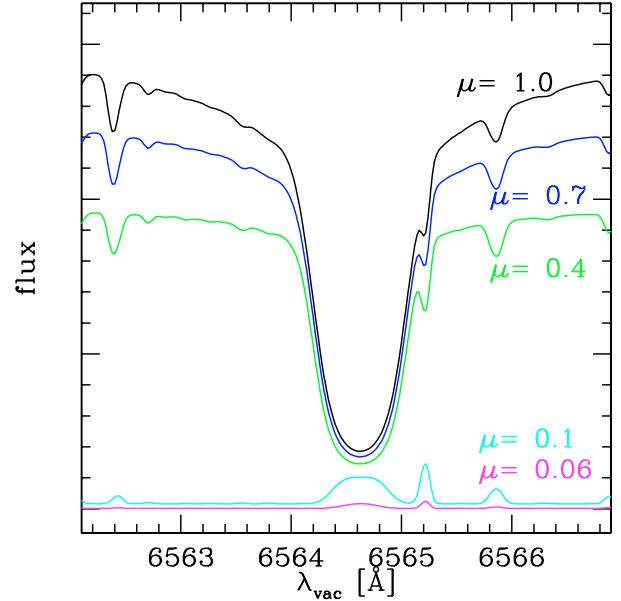
**Fig. 3.** The effect of a microlensing caustic crossing on the integrated light and  $EW$  of the  $H\alpha$  line. *Left panel:* as a star crosses a caustic, it becomes differentially magnified due to the asymmetric magnification pattern (red long dashes). Only the exit is shown. The vertical thin lines show contours of equal magnification leading up to the fold caustic line. Magnification drops immediately outside the caustic. *Right panels* show synthetic data for which a stellar model atmosphere has been convolved with a microlensing magnification pattern. The *top panel* shows how the total flux received from the star changes due to the magnification of the caustic. The *bottom panel* shows the respective changes in  $H\alpha$   $EW$  for this model. Note the weaker/stronger dips in  $EW$  that occur when the leading/trailing limb crosses the caustic, contributing emission (negative  $EW$ ) to integrated light.

(Cassan et al. 2004), and the collaboration kindly provided us with the fully-reduced spectra around the  $H\alpha$  line.

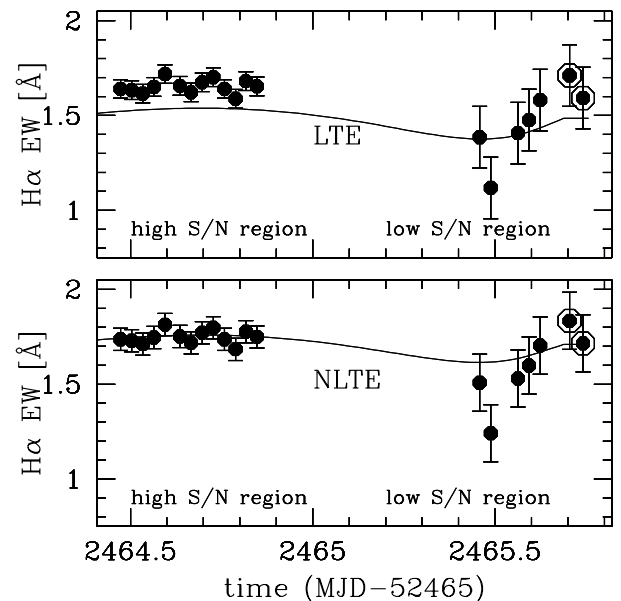
We adopted the parameterization of the fold caustic exit of the event from Cassan et al. (2004). For each time at which a UVES spectrum was available we produced an equivalent synthetic microlensed spectrum for both LTE and NLTE PHOENIX model atmospheres, for each atmospheric model on our stellar parameter grid.

We measured the  $H\alpha$   $EW$  of the synthetic microlensed data via the algorithm described in Sect. 3, and applied the same algorithm to the observational data. Due to the virtual non-existence of noise in the synthetic data, we adjusted some values of the fitting parameters for the observational data: the clipping parameter  $C$  was increased, and we have added an extra  $8 \text{ \AA}$  outwardly to the range over which we determine the continuum to achieve a higher stability in the continuum. Also, since the observed width of the line is not reproduced by the models, as we describe later, the “core region” was defined as  $H\alpha \pm 0.9 \text{ \AA}$  in the data instead of  $H\alpha \pm 0.75 \text{ \AA}$  in the models.

Figure 5 shows our results, where the  $H\alpha$   $EW$  is plotted for each time at which observations have been taken. (The last two points were taken long after the caustic exit to ensure two baseline points not affected by the differential magnification.) We plot the  $H\alpha$   $EW$  measurements for only one set of stellar parameters, namely:  $T_{\text{eff}} = 5000 \text{ K}$ ,  $\log(g) = 2.5$  and  $[\text{Fe}/\text{H}] = -0.5$ . These are the parameters from our grid of stellar models that are closest to the stellar parameters given by Cassan et al. (2004)  $T_{\text{eff}} = 5000 \text{ K}$ ,  $\log(g) = 2.5$  and  $[\text{Fe}/\text{H}] = -0.6$ . The difference in metallicity of 0.1 dex is well within observational errors of



**Fig. 4.**  $H\alpha$  spectral lines from stellar model atmospheres for different values of  $\mu = \cos \theta$ , where  $\theta$  is opening angle to the line-of-sight. At the limb  $\mu = 0$  and  $\mu = 1$  in the center of the stellar disk.  $H\alpha$  is a strong absorption feature for most of the stellar disk, but near the limb  $H\alpha$  goes into emission.



**Fig. 5.** Changes in the  $H\alpha$   $EW$  for OGLE-BULGE-2002-069 during the course of its caustic exit. The solid lines represent the synthetic data for LTE (*upper panel*) and NLTE models (*lower panel*) for  $T_{\text{eff}} = 5000 \text{ K}$ ,  $\log(g) = 2.5$  and  $[\text{Fe}/\text{H}] = -0.5$ , taken from our model grid, closest to those published by Cassan et al. (2004). The filled circles show our derivation of  $H\alpha$   $EW$  from the spectra taken by Cassan et al. (2004). We corrected the observations for the systematic uncertainties in the  $EW$  determination for LTE and NLTE models computed from Monte-Carlo simulations, and applied their statistical uncertainties as errorbars. We did this separately for the high and low  $S/N$  regions. The double-ringed datapoints show observations taken 36 days after caustic exit.

the metallicity determination. Further, in previous tests, we did not detect any measurable changes in  $H\alpha$   $EW$  from models with  $[\text{Fe}/\text{H}] = -0.5$  to models with solar metallicity, so we are confident in our assumption that there are no effects from the 0.1 dex difference in metallicity.

**Table 2.** Parameters of Gaussian fits to the Monte-Carlo simulation results for determining systematic and statistical errors. We have separately determined these for LTE and NLTE models at high and low noise levels. Both models have the same stellar parameters. The systematic error,  $\mu_{\Delta EW}$  is the amount that must be subtracted, on average, from the  $EW$  determined from noisy “data” compared to that determined from noiseless models. The scatter in  $EW$  determination about its known mean is  $\sigma_{EW}$ .

Simulation	$\mu_{\Delta EW}[\text{\AA}]$	$\sigma_{EW}[\text{\AA}]$	$\sigma_{EW}[\%]$
LTE (high $S/N$ )	0.03379	0.0492	3.28
LTE (low $S/N$ )	0.03175	0.1626	10.8
NLTE (high $S/N$ )	-0.0623	0.0586	3.45
NLTE (low $S/N$ )	-0.0912	0.1509	8.88

We quantify the uncertainty in our determination of the  $H\alpha$   $EW$ , using a Monte-Carlo approach simulating 999 noisy spectra by adding random noise to the virtually noiseless synthetic data and rerunning the algorithm to determine the  $EW$ . The simulations have been restricted to a stellar model of  $T_{\text{eff}} = 5000$  K and  $\log(g) = 2.5$ . We assume no variation of the fractional error distribution across the whole stellar parameter grid. The random noise we add follows a Gaussian distribution with an rms scatter  $\sigma$  equivalent to a signal-to-noise ( $S/N$ ) of 130 or 50, which Cassan et al. (2004) quote for their spectra at a resolution of  $\lambda/\Delta\lambda = 30\,000$ . These values translate into a signal-to-noise per spectral data point, spaced at  $0.0174$  Å, of  $(S/N)_{\text{pp}} = 36.5$ , or 2.7% noise, at those times where the center of the star is most highly magnified. At the caustic exit and post-exit time we obtain a  $(S/N)_{\text{pp}} = 14$ , or 7% noise. When we add this noise to the model spectra at the appropriate times, the model and observed spectra indeed appear to have a similar noise level. Thereafter, we shall use *high* and *low*  $S/N$  or *noise* to refer to the 2.7% and 7% noise levels respectively.

From the 999 simulations per noise level and model type, we determined a histogram of  $\Delta EW = EW_{\text{model, noisy}} - EW_{\text{model, noise-less}}$  from noisy to noise-less data, fitted a Gaussian to these. The fitting parameters are listed in Table 2.

The scatter in the observational  $EW$  points is similar to the  $\sigma_{EW}$  determined from the model spectra. Large scatter appears to be mainly a continuum fitting effect. We fit a continuum to the spectra normalized by Cassan et al. (2004) in the same way as we do it for the models. Since the observational data is noisy, the continuum is more uncertain compared to noiseless model data. Even slight differences in the choice of the continuum will result in a larger scatter of  $EW$  values. Since the data have already been normalized by Cassan et al. (2004), for the observations we fit a continuum line with a fixed slope of zero. Our measurement of the continuum is slightly lower than that chosen by Cassan et al. (2004), but is consistent with the way we pick the continuum for the models.

The  $EW$  of the data determined in this way is shown in Fig. 5 as dots; in the same figure, model predictions are also indicated. The overall trend of the observational data is reproduced by the models, however the offsets between them is obvious.

The LTE models do not match the magnitude of the  $EW$  of the  $H\alpha$  line of OGLE-BULGE-2002-069 after the caustic exit (double-ringed data points), where no differential magnification occurs and also at times when the center of the source star crosses the caustic. This could indicate that either the stellar atmosphere models are insufficient and/or we have chosen the wrong stellar parameters from the grid.

## 6. Giving the models a fair chance

Before condemning the model atmospheres prematurely, we need to be sure that we have chosen models that best fit the  $H\alpha$   $EW$  data within uncertainties in the determination of  $T_{\text{eff}}$  and  $\log(g)$ . We quantify the discrepancy between models and observations by defining  $\chi^2$ , the goodness-of-fit figure of merit.

$$\chi^2 = \sum_{t_{\text{obs}}} [EW_{\text{obs}}(t_{\text{obs}}) - EW_{\text{model}}(t_{\text{obs}})]^2 / (\sigma_{EW}^2(t_{\text{obs}})), \quad (3)$$

summing over all points in time  $t_{\text{obs}}$  at which an observed spectrum has been taken. We assume that any systematic errors introduced by our fitting algorithm and statistical uncertainties in fitting model spectra are the same across the model grid. Thus, for  $\sigma_{EW}$  we use the values derived in Sect. 5.1, namely, around 10% and 3.5% for the high and low noise data, respectively.

We calculate  $\chi^2$  for each model in our model stellar grid, and show in Fig. 6, contours of equal  $\Delta\chi^2$ ,

$$\Delta\chi^2 = \chi^2 - \chi_{\text{min}}^2, \quad (4)$$

where  $\chi_{\text{min}}^2$  is the minimal value of  $\chi^2$  across the model grid, i.e. the goodness-of-fit for the best fitting model. Models along a contour line fit the observed measurements equally well. In Fig. 6, we superimpose on the model grid the stellar parameters determined by Cassan et al. (2004) for OGLE-BULGE-2002-069. We follow an analysis of Minniti et al. (2002), used by Cassan et al. (2004), indicating uncertainties in determining stellar parameters of  $\pm 200$  K in  $T_{\text{eff}}$ ,  $\pm 0.5$  in  $\log(g)$ , and  $\pm 0.3$  in  $[\text{Fe}/\text{H}]$ .

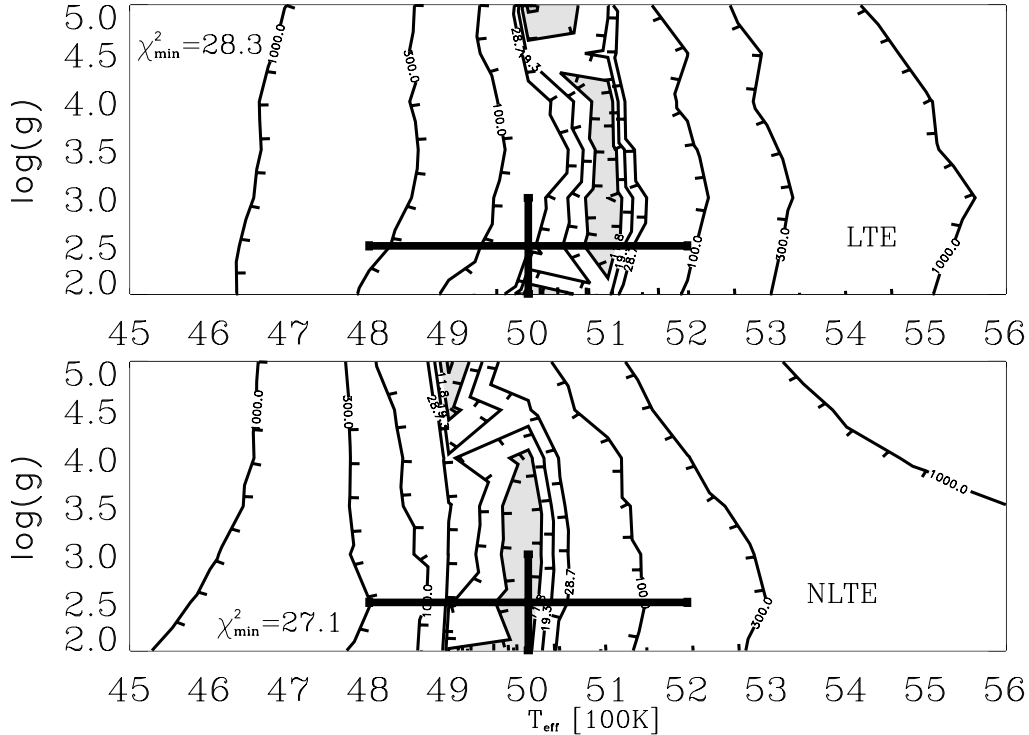
Within the uncertainties of the stellar model, LTE models with  $\approx 100$  K higher  $T_{\text{eff}}$  than and NLTE models at the same values quoted by Cassan et al. (2004) provide the smallest  $\chi^2$ . Gravity is not well constrained. The overall goodness of fit is similarly poor for both LTE and NLTE models at a values of  $\approx 28$  for 20 observations, which could be caused by random noise, with only about 7% confidence. Only the best fitting models should be examined to study further differences to the observations. Note that the best-fitting LTE and NLTE models have *different temperatures*.

## 7. The importance of line shape

Equivalent width collapses all the information about a spectral line into one value that only considers the area of the line under the continuum, but is independent of the line shape. This means that, despite a good agreement between the  $EW$  of a model spectral line, and of the observed spectral line any variations in the line shape will be ignored if this is the only figure of merit used to test models.

Examples of the  $H\alpha$  spectral line (top panel) and their fit from the algorithm (bottom) panel are shown in Fig. 7 for one observation. The line from corresponding differentially magnified LTE and NLTE models with parameters as chosen from Cassan et al. (2004) are shown in the left panels, right panels give the best fitting stellar models as determined in Fig. 6. Although the changes appear small, allowing stellar parameters  $T_{\text{eff}}$  and  $\log(g)$  to vary within the observationally allowed uncertainties, can alter the  $EW$  by 10–15%, as can be seen in Fig. 2.

That the models do not generally fit the width of the  $H\alpha$  line (see Fig. 7) suggests that there are other physical processes occurring which are yet to be explained by stellar models. One explanation may be chromospheric influences, which Cassan et al. (2004) claim to have detected in their data. Since



**Fig. 6.** *Upper panel:* contours of equal  $\Delta\chi^2 = \chi^2 - \chi_{\min}^2$ , where  $\Delta\chi_{\min}^2$  is minimal when best agreement between LTE models and observations of the  $H\alpha$  EW of OGLE-BULGE-2002-069 are achieved. *Lower panel:* same for NLTE.  $\Delta\chi_{\min}^2$  is different for LTE and NLTE models and its value is shown in the left corner of the figure panels. The grey area indicates all models that lie within  $3\sigma$  of the best fitting model for two degrees of freedom, namely  $T_{\text{eff}}$  and  $\log(g)$ . Further contour lines are the  $4\sigma$ ,  $5\sigma$  and higher boundaries. The tick marks on the contour lines indicate the direction of decrease in  $\Delta\chi^2$ . From their spectral analysis, Cassan et al. (2004) determined the stellar parameters to be  $T_{\text{eff}} = 5000$  K,  $\log(g) = 2.0$ ,  $[\text{Fe}/\text{H}] = -0.6$ . We mark this choice of parameters with the large cross with uncertainties of  $\Delta T_{\text{eff}} = \pm 200$  K and  $\log(g) = \pm 0.5$  as given in Minniti et al. (2002) for this type analysis.

$H\alpha$  is a chromospheric line but the standard model atmospheres do not include chromospheric calculations, more research is needed to explain this phenomenon.

This comparison of individual spectra clearly shows where discrepancies in  $EW$  occur, primarily in the wings of the line, but for LTE models also in the core of the line. However, the differences in shape in the core and the wings balance each other out when only total  $EW$  is considered. The discrepancies due to the different line shapes are washed out. Therefore we have devised a different diagnostic to compare model spectra with observational data by defining the line shape diagnostic,  $\psi^2$ ,

$$\psi^2(t_{\text{obs}}) = \int_{\text{line}} [\mathcal{F}_{\text{obs},o}(\lambda, t_{\text{obs}}) - \mathcal{F}_{\text{model},o}(\lambda, t_{\text{obs}})]^2 d\lambda, \quad (5)$$

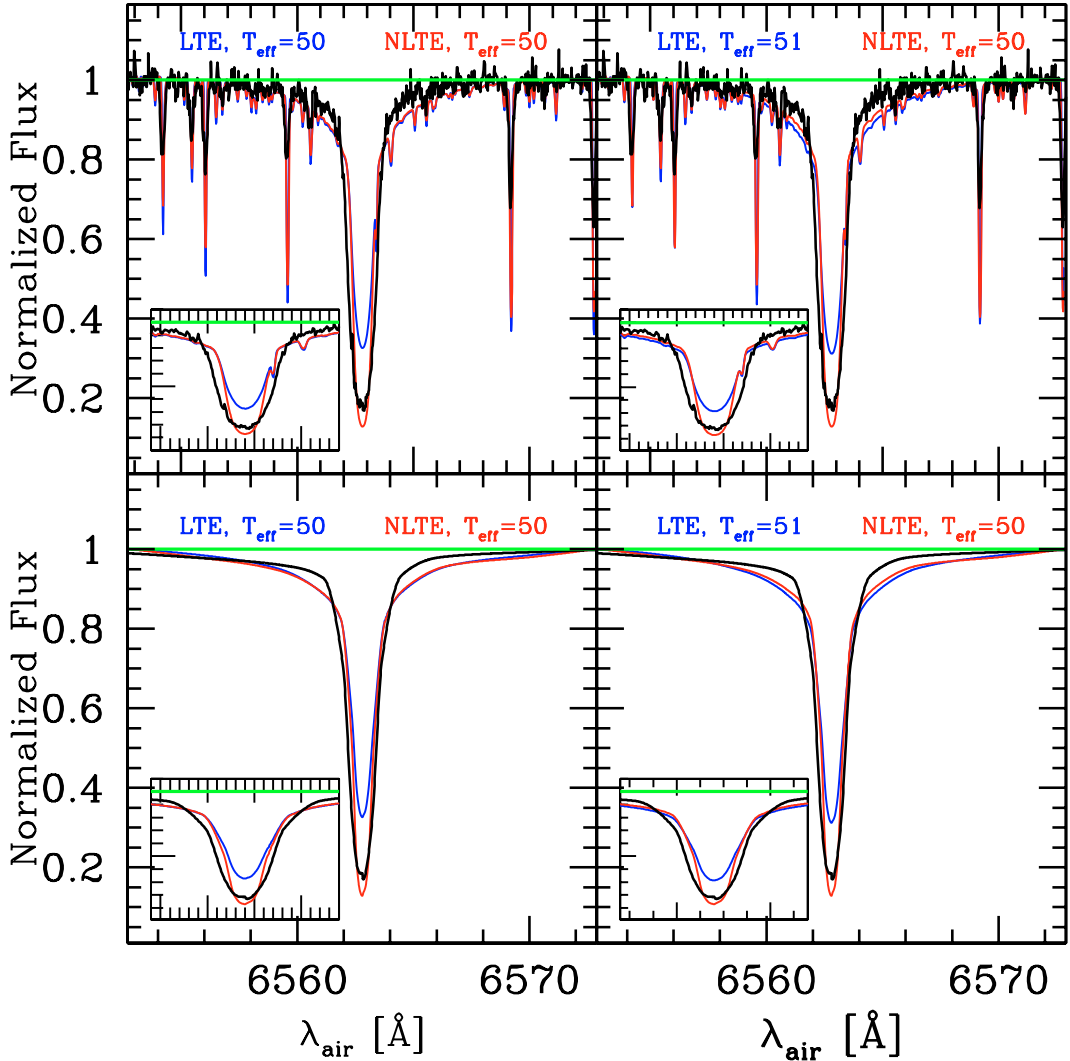
where we integrate the square of the flux difference that results when we subtract the model spectrum from the observational spectrum, after both have been normalized, clipped and spline-interpolated. Note, that the units of this new diagnostic are  $\text{\AA}$ , since the flux has already been normalized. Similarly, we can obtain  $\psi^2$  for line shape comparisons between LTE and NLTE models. The result of which is shown in Fig. 8 for non-microlensed model data. In order to compare LTE and NLTE models to observational data taken during a microlensing event, and to find the best-fit model based on the measured shape parameter, we define a new shape figure of merit

$$\Psi \equiv \sum_{t_{\text{obs}}} \psi^2(t_{\text{obs}}) \quad (6)$$

which is sum of equally weighted  $\psi^2$  over each observation during the event. The units of  $\Psi$  are still  $\text{\AA}$ .

Defining the shape parameter  $\Psi$  allows us to collapse the available time series data to one quantity per stellar model, while conserving the information of the line shape comparison. In effect,  $\Psi$  is sum of  $\chi^2$  comparison of all the spectra themselves, across a wavelength range defined by the line, equally weighted. We determine uncertainties in  $\Psi$  by fitting a Gaussian to the distribution of  $\Psi$  values resulting from 999 Monte-Carlo simulations. In each simulation we sum up 13 low noise and 7 high noise  $\psi^2$  values randomly drawn from the distribution of  $\psi^2$  values, which we have determined via 999 Monte-Carlo simulations in a similar fashion as for  $EW$  in Sect. 5.1 for low and high noise, LTE and NLTE models. We fit the resulting distributions of  $\psi^2$  with a combination of a Gaussian and a straight line as the distribution is asymmetric because  $\psi^2$  is positive definite. We fit the  $\Psi$  distribution with a Gaussian convolved with a straight line. From this we deduce 3, 2 and  $1\sigma$  levels of confidence which we mark in Fig. 9 as dark grey, light grey and the enclosed white areas.

In Fig. 9, we show contour plots of equal  $\Psi - \Psi_{\min}$ , where  $\Psi_{\min}$  is the lowest value of  $\Psi$  across the model grid. For LTE models this is  $\Psi_{\min} = 0.71$ , for NLTE models  $\Psi_{\min} = 0.47$ . The shape parameter  $\psi^2$  is much more dependent on  $\log(g)$  than is  $EW$ , which was strongly dependent only on  $T_{\text{eff}}$ . The dark grey shaded area indicates the 99.73% level of confidence around the best fitting model. The enclosed light grey and white areas show the 95.4% and 68.3% confidence levels, respectively. The best-fitting LTE model has a  $\log(g)$  of less than 2.0, substantially less than what the observationally-determined stellar parameters suggest. As with fits to  $EW$ , LTE fits to the shape parameter,  $\Psi$ , again suggest a higher  $T_{\text{eff}}$  by about 200 K, i.e. 100 K higher than

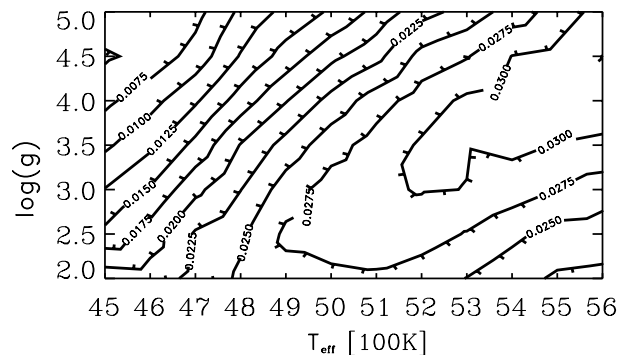


**Fig. 7.** Differences in the spectra between observations of the H $\alpha$  line for OGLE-BULGE-2002-069 (in black) and models (LTE in blue, NLTE in red) at the same magnification. The left panel shows comparison to the models with the stellar parameters closest to the published values, the right panel shows the spectra from models as chosen from Fig. 6. Bottom panels present the clipped and spline interpolated lines after the observations and models have passed through the *EW* algorithm from Sect. 3.1. Each inset shows a narrower region around the core of the line. For all diagrams  $\log(g) = 2.5$  and  $T_{\text{eff}}$  is given per 100 K.

the *EW*-best-fitting LTE model suggested. Although the best-fitting NLTE model for the shape parameter leans towards lower gravities, these are in agreement with the observationally determined stellar parameters. The overall fit to  $\Psi$ , however, is fairly poor: the distribution of  $\Psi$  at random shows that 99.9% of values drawn randomly from the distribution of  $\psi^2$  for NLTE models suggests a value below 0.07. For LTE models this reduces to a value less than 0.01.

## 8. Conclusions

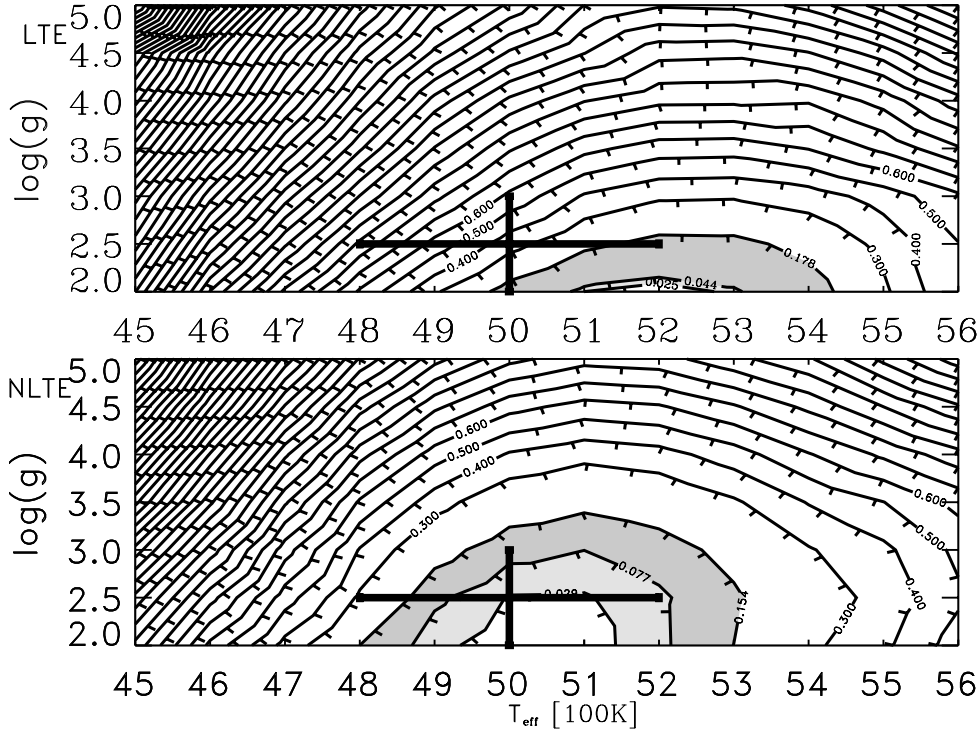
As a part of a larger program to compare model atmospheres to observational data, we have developed an algorithm to consistently and automatically determine and compare spectral line characteristics. Here, we have focused on equivalent width (*EW*) of the H $\alpha$  line and also introduced a line shape parameter  $\psi^2$ , which emphasizes flux differences between two spectra. We have studied a stellar parameter range of  $4500 < T_{\text{eff}} < 5300$  and  $2.0 < \log(g) < 5.0$  at  $[\text{Fe}/\text{H}] = -0.5$ , a range typical of microlensed source stars in the Galactic Bulge.



**Fig. 8.** Contours of equal  $\psi^2$  for a NLTE-LTE model comparison. The tick marks on the contour lines indicate the direction of decrease. Measured by this quantity, line shapes for NLTE and LTE models differ least for lower  $T_{\text{eff}}$  and higher  $\log(g)$  within the parameter grid. Larger differences are found for higher  $T_{\text{eff}}$  and higher  $\log(g)$ .

We have determined the uncertainty of this algorithm by performing Monte-Carlo simulations in which noise is introduced to a model spectrum and the extraction procedure repeated. The





**Fig. 9.** Contours of equal  $\Psi - \Psi_{\min}$  for comparison of the observational data with LTE models (*upper panel*) and NLTE models (*lower panel*). The contours clearly constrain not only the  $T_{\text{eff}}$  but also the  $\log(g)$  of the best fitting model parameters. This means that the best fitting LTE model has not only a higher  $T_{\text{eff}}$  and a lower  $\log(g)$ , whereas the best fitting NLTE model has stellar parameter closest to the one suggested by Cassan et al. (2004), indicated by the black cross.

simulations have been restricted to a stellar model of  $T_{\text{eff}} = 5000$  K and  $\log(g) = 2.5$ . We assume no variation of the fractional error distribution across the whole stellar parameter grid. The *EW* of the  $H\alpha$  spectral line as predicted by PHOENIX atmospheric models is substantially different for LTE and NLTE model. The wrong choice of model type may result in uncertainties in *EW* larger than 8% for most models with  $4500 < T_{\text{eff}} < 5300$  and  $2.0 < \log(g) < 5.0$ . The shape parameter  $\psi^2$  in an LTE-NLTE comparison is statistically significant and largest for high  $T_{\text{eff}}$  and high  $\log(g)$ .

Choosing the right model to compare to the observations is crucial when one wants to understand stellar atmosphere modeling. As an example, we study  $H\alpha$  spectral data from the microlensing event OGLE-BULGE-2002-069. First, we determine the best-fitting model to the  $H\alpha$  *EW* data over the range of our stellar grid for both LTE and NLTE models over the course of the caustic exit, during which the source star is resolved. We find that in order to fit the observed data an LTE model requires a  $T_{\text{eff}}$  about 100 K larger than that proposed by Cassan et al. (2004) based on spectral type analysis for this source star. The best-fitting NLTE model, on the other hand, requires a  $T_{\text{eff}}$  as proposed by Cassan et al. (2004). The absolute  $\chi^2$  is lower for NLTE PHOENIX models than for LTE models, suggesting a slight preference for the NLTE model.

Second, we determine the best-fitting model to the  $H\alpha$   $\psi^2$  which we define, and find that the best-fitting LTE model has a  $T_{\text{eff}}$ , that is 200 K higher than the observationally determined  $T_{\text{eff}}$  and a significantly lower  $\log(g)$ . The best-fitting NLTE model agrees with the observationally determined values for  $T_{\text{eff}}$  and  $\log(g)$  within the uncertainties, although it tends to underestimate both  $T_{\text{eff}}$  and  $\log(g)$  for this star. The NLTE models produced a better figure of merit than the LTE models for both diagnostics.

We caution, however, that declaring either LTE or NLTE models a failure based on analysis of  $H\alpha$  only is premature in the case of OGLE-BULGE-2002-069. Chromospheric effects may alter the characteristics of the  $H\alpha$  line significantly, and the atmosphere models used here do not include chromospheres. Indeed, Cassan et al. (2004) claim to have detected an anomaly in the  $H\alpha$  line at some times during the caustic crossing that may be due to a chromosphere.

We would like to point out that with our method we are already placing strong constraints on  $T_{\text{eff}}$  and  $\log(g)$  by using only two parameters measured for one particular line. Future studies of multiple lines and multiple diagnostics will provide even stronger constraints. In subsequent work, we will determine the most effective choices of spectral lines and their diagnostics for testing model atmospheres and include chromospheric calculations in the models for direct comparison to observations of OGLE-BULGE-2002-069.

*Acknowledgements.* We would like to thank John Tonry for kindly sharing with us his Marquardt-fitting and minimization routine. We would also like to thank the members of PLANET for providing us with the valuable UVES spectra for OGLE-BULGE-2002-069 and fruitful discussions of the matter. This work was supported by ANU Supercomputing facility (APAC). We would like to thank the National Facility help desk, and in particular Ben Evans, for support and help with the APAC system. We are indebted to Martin Asplund for helpful discussions in preparation of this paper.

## References

- Abe, F., Bennett, D. P., Bond, I. A., et al. 2003, *A&A*, 411, L493
- Afonso, C., Alard, C., Albert, J. N., et al. 2000, *ApJ*, 532, 340
- Afonso, C., Albert, J. N., Andersen, J., et al. 2001, *A&A*, 378, 1014
- Albrow, M. D., Beaulieu, J.-P., Caldwell, J. A. R., et al. 1999a, *ApJ*, 522, 1011 (PLANET collaboration)

- Albrow, M. D., Beaulieu, J.-P., Caldwell, J. A. R., et al. 1999b, *ApJ*, 522, 1022 (PLANET collaboration)
- Albrow, M. D., Beaulieu, J.-P., Caldwell, J. A. R., et al. 2000, *ApJ*, 534, 894 (PLANET collaboration)
- Albrow, M. D., An, J., Beaulieu, J.-P., et al. 2001a, *ApJ*, 549, 759 (PLANET collaboration)
- Albrow, M., An, J., Beaulieu, J.-P., et al. 2001b, *ApJ*, 550, L173 (PLANET collaboration)
- Alcock, C., Allen, W. H., Allsman, R. A., et al. 1997, *ApJ*, 491, 436
- An, J. H., Albrow, M. D., Beaulieu, J.-P., et al. 2002, *ApJ*, 572, 521 (PLANET collaboration)
- Anderson, L. S. 1989, *ApJ*, 339, 558
- Asplund, M., Grevesse, N., Guedel, M., & Sauval, A.J. 2005, *Nature*, 436, 525
- Asplund, M. 2003, *ASP Conf. Ser.*, 304, 275
- Aufdenberg, J. P., Ludwig, H.-G., & Kervella, P. 2005, *ApJ*, 633, 424
- Blackwell, D. E., Lynas-Gray, A. E., & Smith, G. 1995, *A&A*, 296, 217
- Bryce, H. M., Hendry, M. A., & Valls-Gabaud, D. 2002, *A&A*, 388, L1
- Bryce, H. M., Ignace, R., & Hendry, M. A. 2003, *A&A*, 401, 339
- Cassan, A., Beaulieu, J. P., Brillant, S., et al. 2004, *A&A*, 419, L1
- Castro, S., Pogge, R. W., Rich, R. M., DePoy, D. L., & Gould, A. 2001, *ApJ*, 548, L197
- Drake, J. J., & Testa, P. 2005, *Nature*, 436, 525
- Fields, D. L., Albrow, M. D., An, J., et al. 2003, *ApJ*, 596, 1305 (PLANET collaboration)
- Gilliland, R. L., & Dupree, A. K. 1996, *ApJ*, 463, L29
- Gray, D. F. 1992, *Cambridge Astrophysics Series* (Cambridge: Cambridge University Press), 2nd ed.
- Hauschildt, P. H., & Baron, E. 1999, *J. Comput. Appl. Math.*, 102, 41
- Hauschildt, P. H., Allard, F., Ferguson, J., Baron, E., & Alexander, D. R. 1999a, *ApJ*, 525, 871
- Hauschildt, P. H., Allard, F., & Baron, E. 1999b, *ApJ*, 512, 377
- Hendry, M. A., Bryce, H. M., & Valls-Gabaud, D. 2002, *MNRAS*, 335, 539
- Heyrovský, D. 2003, *ApJ*, 594, 464
- Heyrovský, D., & Sasselov, D. 2000, *ApJ*, 529, 69
- Heyrovský, D., Sasselov, D., & Loeb, A. 2000, *ApJ*, 543, 406
- Kubas, D., Cassan, A., Beaulieu, J. P., et al. 2005, *A&A*, 435, 941
- Lennon, D. J., Mao, S., Fuhrmann, K., & Gehren, T. 1996, *ApJ*, 471, L23
- Lennon, D. J., Mao, S., Reetz, J., et al. 1997, *The Messenger*, 90, 30
- Minniti, D., Barbuy, B., Hill, V., et al. 2002, *A&A*, 389, 419
- O'Donoghue, D., Koen, C., Kilkeny, D., et al. 2003, *MNRAS*, 345, 506
- Oláh, K., Jurcsik, J., & Strassmeier, K. G. 2003, *A&A*, 410, 685
- Popper, D. M., & Guinan, E. F. 1998, *PASP*, 110, 572
- Richichi, A., & Lisi, F. 1990, *A&A*, 230, 355
- Trujillo, I., Aguerri, J. A. L., Cepa, J., & Gutiérrez, C. M. 2001, *MNRAS*, 328, 977
- Schneider, P., Ehlers, J., & Falco, E. E. 1992, *Gravitational Lenses* (Berlin, Heidelberg, New York: Springer-Verlag), XIV, 560
- Valls-Gabaud, D. 1998, *MNRAS*, 294, 747