Detecting extrasolar planets

Sam Dolan (Jan 2017)

"Shall all this immense space, which comprehends our Sun and our Planets, be only a little parcel of the Universe? Shall there be as many such spaces as there are Fixed Stars? This confounds me, troubles me, frights me."
– The Marquess in *Conversations on the Plurality of Worlds* (1686) [1].

1 Introduction

Does life exist elsewhere in our Universe? This fundamental question has inspired much speculation through the centuries, spurred by advances in technology. For example, Galileo's discovery of the moons of Jupiter (in 1610) led the writer de Fontenelle to imagine (in 1686) that Jovian astronomers could be watching Earth through their own telescopes [1]. Now, in the 21st century, our imagination is sparked by the discovery of new planets beyond our solar system.

An extrasolar planet (or exoplanet) is a planet in orbit around a star other than the Sun. Over 3500 exoplanets have been discovered since 1995. The rate of discovery is increasing. The Kepler space observatory alone, launched in 2009, has confirmed over 2300 discoveries. Improved techniques are leading to discoveries of less massive, rocky planets (like Earth) in addition to massive gas giants (like Jupiter).

To date, only a few (~ 21) planets have been found to lie on orbits within the 'habitable zone', in which the planetary surface may support liquid water if there is sufficient atmospheric pressure. This is seen as a necessary (but not sufficient) condition for the existence of life.

In July 2016, astronomers reported the discovery of a new exoplanet, named Proxima B, orbiting our nearest neighbour star [3]. Intriguingly, Proxima B is within the habitable zone of Proxima Centauri, a red dwarf star 4.6 light years away. Yet, as red dwarfs are comparatively dim, this implies that Proxima B is ~ 20 times closer to its parent star than the Earth is to the Sun. Life on Proxima B therefore seems unlikely, given that the planet is (most likely) tidally-locked to the star, with its magnetic field insufficient to protect it from flare events.

As many as 14 different methods have been used to detect extra-solar planets [4]. The most widely-used are transit photometry, and radial velocity measurement. With transit photometry, the light from a star is observed to dim and then brighten as a planet passes across the stellar disk. With the radial velocity method, emission lines in the spectrum of the star are found to periodically shift in frequency: a Doppler shift caused by the star's motion due to the gravitational attraction of the planet.

In this assignment, we consider the (fictitious but plausible) scenario that an astronomer has inferred the relative velocity of a star using 60 days of Doppler shift measurements. The periodic oscillation, seen in Fig. 1, is taken as evidence that the star has a companion, that is causing the star to orbit around the common centre of mass with a relative velocity of just a few metres per second. As the signal is not strictly sinusoidal, it follows that the motion is not strictly circular. In the following sections, we attempt to deduce the properties of the system by fitting various models to the data set. Data [5] and sample code [6] are available.

2 Model fitting

(a) A crude model. An astronomer has compiled a data file [5] of relative velocity measurements v_j (in metres per second) at times t_j (in days), with $j = 1 \ldots N$. There is some experimental error in the data, which appears to be random rather than systematic. In our first pass, we will fit the relative velocity data v_j with a simple sinusoidal model

$$f_0(t) = \beta_0 \sin(\omega t) + \beta_1 \cos(\omega t), \tag{1}$$

where $\omega = 2\pi/15$. The model is linear in its parameters, β_0 and β_1. As outlined in Lecture 7 [7], we may find the best-fit parameters that minimize the sum-of-squared-residuals by solving the *normal equations*

$$\mathbf{X}^T \mathbf{X} \boldsymbol{\beta} = \mathbf{X}^T \mathbf{v}. \tag{2}$$

Here $\mathbf{v} = [v_j]$ is the vector of relative velocity measurements, $\boldsymbol{\beta} = [\beta_0, \beta_1]$ is the vector of parameters, and $\mathbf{X}$ is an $N \times 2$ matrix with rows $[\sin(\omega t_j), \cos(\omega t_j)]$. As the normal equations are in the standard form $\mathbf{A}\mathbf{x} = \mathbf{b}$, it is straightfoward to find the best-fit parameters $\beta_0 \approx -0.706542$ and $\beta_1 \approx -0.129861$ using e.g. `np.linalg.solve()`.

Figure 1 (left) shows the data (blue dots) alongside the fitted model (green line). Although there is a reasonable agreement, it is clear that the sinusoidal model is too simple to capture all aspects of the data. For example, the oscillation in the data appears *asymmetric*, as it rises more slowly than it falls.

(b) A good fit? Figure 1 (right) shows the residuals after fitting the model, $r_j = v_j - f_0(t_j)$. If the fit were optimal, then the residuals would be dominated by experimental error, which we assume is randomly distributed. If this were the case, then there would be no visible correlation between successive errors. Instead, we observe in Fig. 1 that there is residual structure which has not been captured by the crude model (1).

One measure of the 'goodness of fit' of a model is the *root mean square deviation* (RMSD), defined by $\sqrt{\frac{1}{N}\sum_j r_j^2}$, where r_j are the residuals after fitting. The RMSD for this simple model is 0.2342. Superior models should exhibit a smaller RMSD. However, the RMSD cannot be reduced below the amplitude of the experimental error (~ 0.1).

(c) A better linear model. We may add a second harmonic to the simple model (1), to obtain

$$f_1(t) = \beta_0 \sin(\omega t) + \beta_1 \cos(\omega t) + \beta_2 \sin(2\omega t) + \beta_3 \cos(2\omega t). \tag{3}$$

Figure 2 shows this model fitted to the data. The RMSD is 0.1509, suggesting that the new model is a better fit to the data. Some high-frequency structure remains in the residuals, suggesting the model can be bettered.

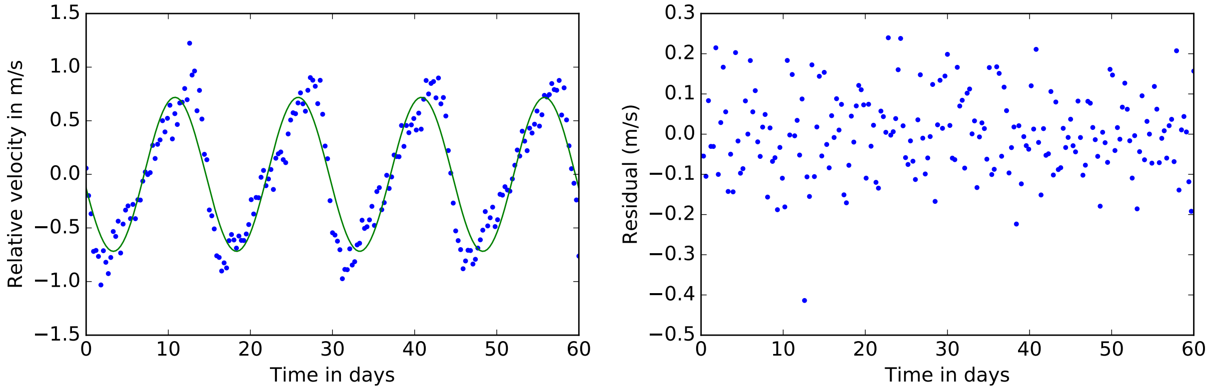

Figure 1: The left plot shows the astronomer's measurements of the relative velocity of the star, as a function of time [*blue dots*], and the best fit of the simple sinusoidal model $f_0(t)$, Eq. (1) [*green line*]. The right plot shows the residuals $r_j = v_j - f_0(t_j)$ after fitting.

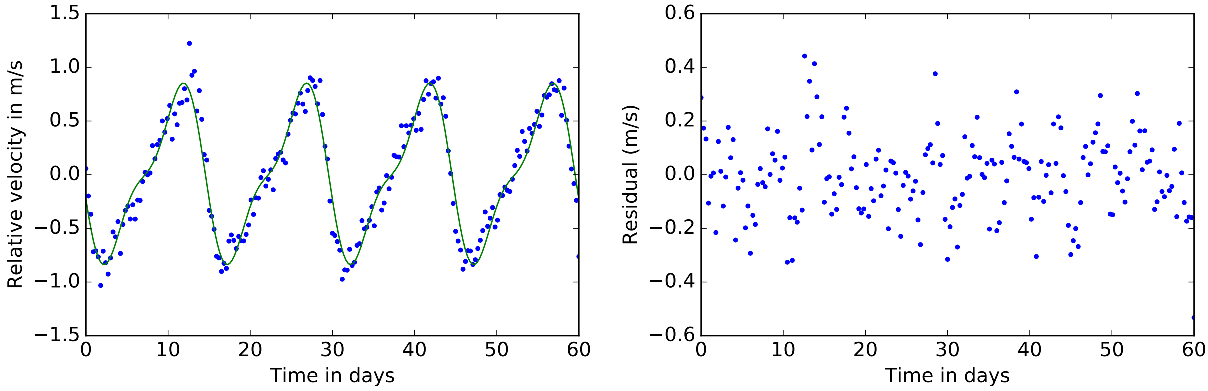

Figure 2: **Part 2(c)**. *Left:* the four-parameter linear model $f_1(t)$, Eq. (3), compared with data. *Right:* The residuals.

(d) A non-linear model. The `curve_fit()` function in the `scipy.optimize` package in Python can be used for fitting models which are not linear in their parameters. After some trial-and-error, I decided to fit a six-parameter model $f_2(t)$ which could recreate the asymmetry (in rise-and-fall) noted earlier, namely,

$$\beta_0 \sin\left(\beta_1 t + \beta_2 + \beta_3 \sin(\beta_4 t + \beta_5)\right) \tag{4}$$

The fit is shown in Fig. 3. The best-fit parameters were found to be
$\beta_i = [0.7977, 0.4230, 3.1963, 0.4189, 0.4225, 0.1002]$. Note that $\beta_1 = 0.4230$ is close to the frequency $\omega = \pi/15 \approx 0.4189$, used in part 2(b).

To fit this non-linear model using `curve_fit()` it was essential to choose sufficiently-good starting values for the parameters. Otherwise, the method typically failed, producing very bad fits.

The RMSD was 0.1520, suggesting that model (4) was not a significantly better fit than model (3), at least by this measure.

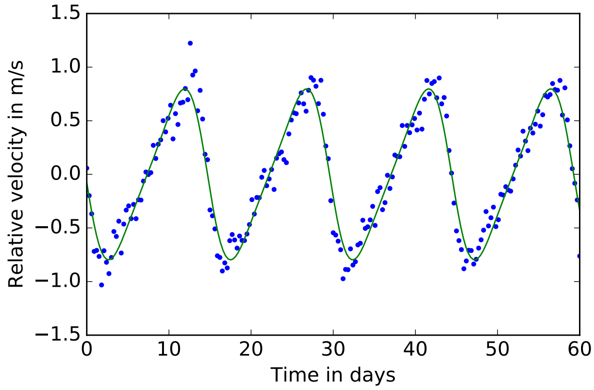

Figure 3: **Part 2(d)**. The non-linear model $f_2(t)$, defined in Eq. (4).

3 A physical model

A planet and a star interacting under gravity are governed by Newton's laws of motion ($\mathbf{F} = m\mathbf{a}$, $\mathbf{F}_{1\to 2} = -\mathbf{F}_{2\to 1}$) and Newton's law of universal gravitation, $\mathbf{F}_{1\to 2} = -Gm_1m_2\mathbf{r}_{1\to 2}/r^3$. Kepler's laws for a two-body system follow as a consequence of Newton's laws. Kepler's first law states that the bodies (planet and star) move on closed ellipses, each with a focus at the centre of mass. Kepler's second law states that "equal areas are swept out in equal times"; essentially, this a statement of the conservation of angular momentum. Kepler's third law relates the orbital period to the semi-major axis of the ellipse (see Sec. 3(d)).

It can be shown that Newton's laws lead to the following ordinary differential equations:

$$\begin{aligned} \ddot{r} &= -\frac{1}{r^2} + \frac{p}{r^3}, \qquad r(0) = \frac{p}{1+e}, \qquad \dot{r}(0) = 0, \\ \dot{\phi} &= \frac{\sqrt{p}}{r^2}, \qquad \phi(0) = \phi_0, \end{aligned}$$

Here $r(t)$ is the distance between the planet and star (in certain units), and $\phi(t)$ is the orbital angle (in radians). The parameters p and e are the semi-latus rectum and the eccentricity of the orbit. The parameter ϕ_0 is the initial orbital angle.

To write these equations in first-order form, we introduce a new variable $s(t)$ defined by $\dot{r} = s$. Thus we obtain a three-dimensional autonomous system of first-order ODEs,

$$\begin{aligned} \dot{r} &= s, \\ \dot{s} &= -\frac{1}{r^2} + \frac{p}{r^3}, \\ \dot{\phi} &= \frac{\sqrt{p}}{r^2}, \end{aligned} \tag{5}$$

with initial conditions $r(0) = p/(1+e)$, $s(0) = 0$ and $\phi(0) = \phi_0$.

(a) Numerical solution of ODEs with the midpoint method. System (5) is in the standard first-order form $\dot{\mathbf{X}} = \mathbf{F}(\mathbf{X}, t)$, where $\mathbf{X}(t) = [r(t), s(t), \phi(t)]$ is a 3-component vector, and $\mathbf{F}$ is determined by the right-hand side of (5). We may solve such equations

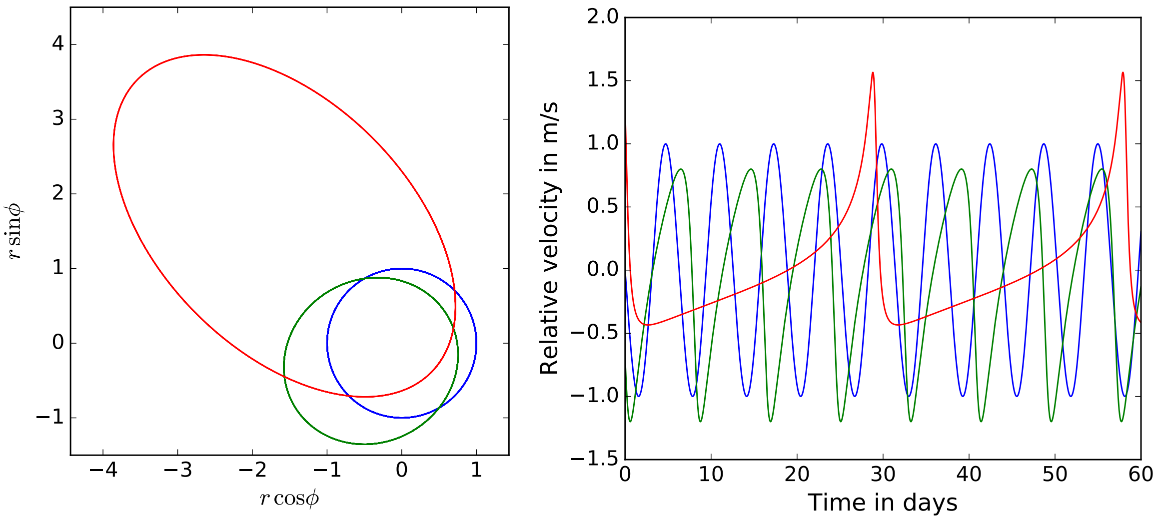

Figure 4: **Part 3(b)**. Left: Three exoplanet orbits around the centre of mass at $(0, 0)$. Right: The relative velocities of the star's motion (which will generate a Doppler shift in the star's emission lines).

numerically, by generating a sequence of estimates $\mathbf{X}_j = [r_j, s_j, \phi_j]$, where r_j is the estimate at time $t_j = jh$ with step size $h = \Delta t$, and $j \geq 0$ is an integer.

I used the midpoint method to obtain numerical solutions of Eq. (5). The midpoint method uses the iterative steps

$$\begin{aligned} \mathbf{X}_{j+1/2} &= \mathbf{X}_j + \frac{h}{2}\mathbf{F}(\mathbf{X}_j, t_j), \\ \mathbf{X}_{j+1} &= \mathbf{X}_j + h\,\mathbf{F}(\mathbf{X}_{j+1/2}, t_{j+1/2}). \end{aligned} \tag{6}$$

The midpoint method is second-order accurate, meaning that the global truncation error scales with h^2. I used the ratio test [7] to confirm that my implementation showed the correct convergence. In practice, I found that a step size of $h = 0.01$ was sufficient for adequately-accurate results.

(b) Orbits and signals. Three orbits are shown in Fig. 4 (left), for the parameters $(p, e, \phi_0) = (1, 0, 0)$ (blue), $(1, 0.4, \pi/6)$ (green) and $(1, 0.8, -\pi/4)$ (red). In accordance with Kepler's first law, the orbits are ellipses with one focus at the centre of mass $(0, 0)$. The orbits are *closed*, that is, that they do not precess[1]. The shape of the ellipse is determined by the eccentricity e. The $e = 0$ orbit (blue) is circular, whereas the $e = 0.4$ orbit is more elliptical. The body is moving fastest when it is closest to the focus (i.e. at the *periapsis*) and moving slowest when it is farthest from the focus (i.e. at the *apoapsis*).

The relative velocity of the star along the line of sight $v_\parallel$ is given by

$$v_\parallel = \dot{r}\cos\phi - r\dot{\phi}\sin\phi \tag{7}$$

[1]Closed orbits are a property of inverse-square-law forces, such as Newtonian gravity. Einstein's theory (1915) predicted that, in fact, two-body orbits are not closed, as they should slowly precess. His theory explained the anomalous precession of Mercury which had been observed since 1859.

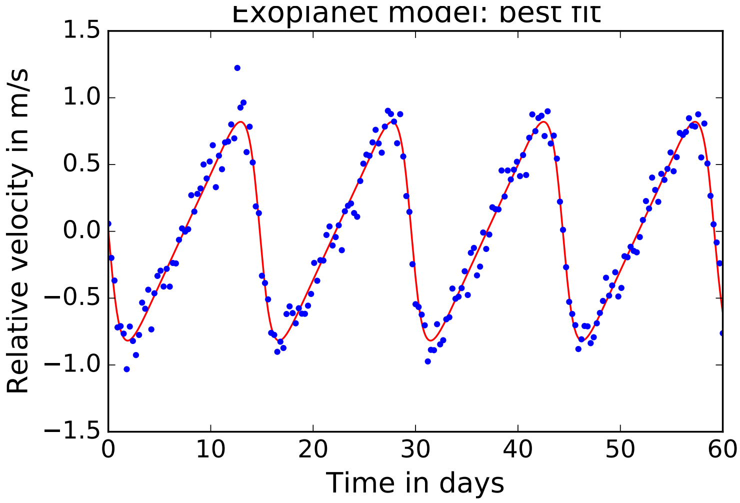

Figure 5: **Part 3(c)**. The best fit of the physical model to the data.

(Here, as instructed, we have assumed that the line of sight is in the plane of motion.) The relative velocities for the three orbits are shown in Fig. 4 (right). The most eccentric orbit ($e = 0.4$, red) has the longest period, and shows the greatest asymmetry (i.e. deviation from sinusoidal profile). Comparing with Fig. 3, it seems plausible that the astronomer's data was generated by an orbit with an eccentricity e greater than zero but less than 0.4. The next challenge is to estimate the orbital parameters p, e and ϕ_0 from the data.

(c) Fitting the physical model. We now have a physical model which, given three parameters p, e, ϕ_0, produces a candidate signal $v_{\parallel}(t)$ which can be compared with the data. In principle it is possible to use Python to find the best-fit values of the parameters. In practice, the implementation requires some care.

To use `curve_fit()`, one must supply a function which calculates the model given a value of t. This function is then called many times for each set of parameter values. However, this is inefficient in our case, because to calculate the model at time t, one must start from $t = 0$. But we do not want to start the numerical integration afresh each time the function is called. For this reason, I took a different approach.

I wrote a function `calcVres()` to take as input the parameter values, and give as output the array of residuals. The function calculates the difference between the model and the data over the whole domain, and calls the midpoint function just once. Then I used `scipy.optimize.leastsq()` to find the parameter values which minimized the sum-of-squared-residuals. See code file [6].

I found the best-fit parameters to be $p = 1.491$, $e = 0.397$ and $\phi_0 = -0.003$, after some trial-and-error with the initial parameter estimates. The best fit model is shown in Fig. 5. The RMSD was 0.1017, suggesting that this is a better fit than the heuristic model of part 2(d). The residuals were found to be randomly distributed, suggesting that the model captures all significant features of the data.

(d) Estimating the physical parameters. Kepler's third law may be written as

$$a_2^3 = \frac{G(m_1 + m_2)}{\Omega^2} \tag{8}$$

where a_2 is the semi-major axis of the planet's elliptical orbit, m_1 and m_2 are the masses of the star and planet, respectively, G is Newton's gravitational constant, and $\Omega = 2\pi/T$ with T the period of the orbit. As instructed, we shall assume that the star has the same mass as the Sun, $m_1 = M_\odot \approx 2 \times 10^{30}\,\text{kg}$. The star may be assumed to be much more massive than the planet ($m_1 \gg m_2$) so that we may neglect m_2, to a good approximation. The orbital angular frequency is $\Omega \approx \omega = 2\pi/15\,(\text{days})^{-1} = 4.848 \times 10^{-6}$ Hz. Using Eq. (8),

$$a_2 \approx 1.784 \times 10^{10}\,\text{m} = 0.119\,\text{au}$$

The *astronomical unit* (au) is defined to be the mean distance between the Sun and Earth, with $1\,\text{au} = 1.496 \times 10^{11}$ m. The Earth's orbit is approximately circular ($e \approx 0.0167$), whereas this extra-solar planet follows an elliptical orbit with $e \approx 0.397$ [Sec. 3(c)]. The distance of closest approach (periapsis) is $r_{\text{peri}} = a_2(1 - e) \approx 0.072$ au. For comparison, Mercury's periapsis is 0.307au. Thus, the exoplanet passes 4.3 times closer to its star than Mercury does to the Sun.

Both star and planet orbit on ellipses about a common centre of mass. Their eccentricities are equal, but their semi-major axes are different. The ratio a_1/a_2 (where 1 = star and 2 = planet) is equal to the ratio of masses m_2/m_1. To get a rough estimate of the planet's mass, we may assume that orbit is circular, and use the familiar formula $v = \Omega r$ (recall MAS112), setting v equal to the amplitude of the relative velocity ($v_{\parallel}^{\max} \approx 0.8\,\text{ms}^{-1}$) and r equal to $a_1 = m_2 a_2/m_1$. Rearranging,

$$m_2 \sim \frac{v m_1}{\Omega a_2} \approx 1.850 \times 10^{25}\,\text{kg}. \tag{9}$$

For comparison, the mass of the Earth is 5.972×10^{24} kg. Thus, our rough estimate suggests that the exoplanet has about 3 times the mass of the Earth.

To refine this estimate, we should take into account the effect of the elliptical nature of the orbit. One way to do this is to use the formula for the planet's velocity $v_{2,\text{peri}}$ at periapsis,

$$v_{2,\text{peri}} = \sqrt{\frac{G m_1 (1 + e)}{a_2 (1 - e)}} \tag{10}$$

The star's velocity at periapsis will be smaller by a factor of m_2/m_1, by similarity of ellipses. Hence

$$m_2 = v_{\parallel}^{\max} \sqrt{\frac{m_1 a_2}{G} \frac{1 - e}{1 + e}} \approx 1.215 \times 10^{25}\,\text{kg}. \tag{11}$$

With this improved argument, the planet is found to be about twice the mass of the Earth.

To conclude, by fitting a Newtonian-dynamics model to the astronomer's relative-velocity data [5], we have found evidence that there is an extra-solar planet in an eccentric orbit ($e \approx 0.4$) around a star. The planet is approximately twice the mass of the Earth, so it is likely to be a 'rocky' planet rather than a 'gas giant'. The planet is unlikely to support life as it passes 14 times closer to its star than the Earth does to the Sun.

References

[1] B. B. de Fontenelle, *Conversations on the Plurality of Worlds* (First published in 1686).

[2] W. J. Borucki et al., *Kepler planet-detection mission: introduction and first results*. Science **327**.5968, 977 (2010).

[3] G. Anglada-Escudé et al., *A terrestrial planet candidate in a temperate orbit around Proxima Centauri*. Nature **536**, 437 (2016).

[4] Methods of detecting exoplanets (Wikipedia, Jan 2017): `https://en.wikipedia.org/wiki/Methods_of_detecting_exoplanets#Radial_velocity`

[5] Data file: `http://sam-dolan.staff.shef.ac.uk/mas212/data/extrasolar_signal.txt`

[6] Example code: `http://sam-dolan.staff.shef.ac.uk/mas212/docs/assign3_code.ipynb`

[7] Course webpage: `http://sam-dolan.staff.shef.ac.uk/mas212/`