Navigation 14

To fulfil its mission, a spacecraft needs to know where it is and how fast it is moving. This is the goal of navigation. Here we look first at how navigation is done for spacecraft in the solar system, using internal measurements as well as artificial signals sent from the Earth acting as navigation beacons. The latter has some relation to the concept of global navigation satellite systems such as GPS, which we also take a brief look at. At interstellar distances, signals from the Earth would be too weak and take too long to arrive, so interstellar spacecraft will have to navigate autonomously. We will see how this can be done using measurements of the positions or radial velocities of stars, by exploiting aberration and parallax, or by timing pulses from neutron stars.

14.1 The need for navigation

An interstellar spacecraft must arrive close enough to its target star to perform the intended science. Suppose that for a flyby mission we want to pass within 1 au of the star. For Proxima Cen at a distance of 1.3 pc, this means that the spacecraft's velocity vector, when launched ballistically from the solar system, needs to be accurate to better than $0.8''$. If our spacecraft is travelling at $0.085\,c$ (50 years to Proxima Cen), this means the velocity transverse to the line of sight to the star has to be less than $95\,\mathrm{m\,s^{-1}}$. This could be a real challenge for rocketless propulsion.

It is actually more complicated than this because stars are moving relative to the Sun. Proxima Cen, for example, has an angular velocity, or proper motion, of about $3.8''\mathrm{yr}^{-1}$, meaning that in 50 years it will have moved $193''$. So we have to target the spacecraft this much ahead of the current star's position. However, the current uncertainty in this proper motion is $0.074\,\mathrm{mas\,yr^{-1}}$, which accumulates to $3.7''$ in 50 years ($1\,\mathrm{mas} = 0.001''$), much more than our targetting requirement. On top of this comes the fact that Proxima Cen is approaching the Sun at $22\,\mathrm{km\,s^{-1}}$, meaning that in 50 years it will be about 0.001 pc closer. Some targets are also in binary stars systems with relatively short periods (e.g. Alpha Cen A and B with a period of 80 years), and this too has to be taken into account. The upshot is that at the current level of accuracy, we cannot target a spacecraft to better than a few au at the target star if we just send it off ballistically with no means of changing its course.

Although the position and motion of the target system would be characterized better by the time an interstellar mission starts, there remain many factors that can alter the trajectory of the spacecraft, by altering the magnitude or direction of its velocity. One factor is the means of propulsion. A rocket may not perform quite as expected, or the laser power arriving at a sail may be slightly different from that predicted, or small-scale instabilities

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

may cause the sail to veer slightly off course. If the Δv provided by the propulsion system differs from expectations by just 0.1%, then after 50 years of travel the spacecraft will arrive at the expected interception point 18 days too late or too early. As Proxima Cen is moving at $32\,\mathrm{km\,s^{-1}}$ relative to the Sun, it would then be 0.34 au away from its expected position. Interstellar magnetic fields acting on the spacecraft, the cumulative effect of particle impacts (drag), and of course the gravity of the target star will change the spacecraft's course and velocity. Finally, during the long travel time, we on Earth are likely to learn a lot more about the target system, such as the presence and orbits of planets, so we might want to steer the spacecraft to a different flyby or rendezvous point within the target system.

All this means that we need to have some control over the velocity of the spacecraft at multiple points during the mission. Aside from making velocity changes, knowledge of the spacecraft's position and velocity are essential for interpreting its science results (e.g. the variation of particle flux as a function of stellar distance), for determining when to switch on instruments, and for knowing in what direction the Earth is for transmitting data.

Navigation is the process of determining the three-dimensional (3D) position and 3D velocity of the spacecraft in some reference frame. This frame is typically the International Celestial Reference Frame (ICRF), an inertial reference frame centred on the solar system barycentre (SSB). Generally speaking, navigation also needs some measure of time. As relativistic time dilation is likely to be significant, this is typically the time in the ICRF (at the SSB, properly known as the barycentric coordinate time, TCB), but it could be the time in the spacecraft's rest frame. The spacecraft will also need a means of determining its orientation in the reference frame, so as to point instruments, antennae, or engines in the right direction.

14.2 Inertial navigation

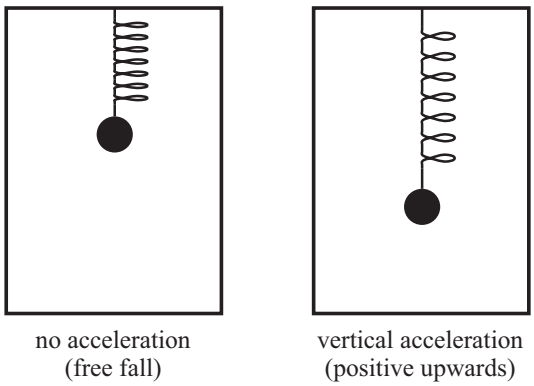
The idea of inertial navigation is to determine the change in position, velocity, and orientation of a spacecraft without using external references.

When a spacecraft moves away from a reference point with known constant velocity in a known direction, determining the future position is a simple matter of measuring the time of travel with an onboard clock. This is known as *dead reckoning*. The error in the determined position comes directly from the error in the initial velocity and orientation and in the measured time. If moving at a relativistic speed, this calculation needs to accommodate the time dilation (or length contraction) to determine the distance travelled in the reference system's rest frame.

14.2.1 Accelerometers

When a spacecraft is accelerating, this can be measured without external reference using an accelerometer. This is essentially a mass on a spring that can move freely within a rigid case and works via the principle of inertia (figure 14.1). In a non-rotating rocket in free space that is not firing its engines, the mass and spring hang freely and there is no acceleration. When the rocket fires its engines, the spring will compress or expand by a length Δx. Knowing the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015



The principle of an accelerometer. Fig. 14.1

spring constant k and the mass m, we compute the acceleration as $(\Delta x)k/m$ from Hooke's law. By orienting three such accelerometers perpendicular to one another, the acceleration in the three spatial directions can be measured. Integrating the acceleration over time we can determine the velocity over time, and by integrating again we establish the position as a function of time.

Accelerometers are insensitive to uniform gravity because gravity acts on all parts of the device equally. Thus an accelerometer in free fall, for example in a Keplerian orbit about a point mass, will read zero acceleration ($\Delta x = 0$). Likewise, the accelerometer on a skydiver who has just jumped out of a plane will read zero (neglecting air drag). Note that in both cases there is a net force acting on the accelerometer. The orbiting accelerometer is accelerating in the sense that its velocity relative to the central mass is changing with time. But this does not register on the accelerometer. In contrast, an accelerometer at rest on the surface of the Earth reads $9.8\,\mathrm{m\,s^{-2}}$ in the upward direction, because there is now a reaction force from the ground, which is transmitted through all the rigid parts of the accelerometer case. In this situation we tend to think of gravity as pulling the mass down to stretch the spring, but of course gravity acts on every other part of the accelerometer too. But in contrast to the rigid parts, there is no reaction force from the ground on the mass, which is free to move and stretch the spring. The same happens in the situation with the accelerometer in the accelerating rocket, and the equivalence of these two is in fact one statement of the equivalence principle. It is the reaction force that makes us feel the effects of gravity when we are standing on the surface of the Earth. Whether we are skydiving in a gravitational field or are in a non-accelerating rocket in deep space, there is no reaction force and so we feel no gravity.

In general relativity, gravity is not considered as a force, but is rather a result of the curvature of spacetime. If there are no (other) forces acting, then objects are unaccelerated and move along so-called geodesics, an example of which is a Keplerian orbit. The apparent force of gravity only becomes measurable when something prevents motion along a geodesic, such as the reaction force of the surface of the Earth when we stand on it. Because they work on the principle of inertia, accelerometers measure the proper acceleration rather than the coordinate acceleration. As discussed in section 7.4, coordinate acceleration is a relative

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

(frame-dependent) concept, like velocity (see section 7.2). But proper acceleration, in the sense of being the acceleration relative to a geodesic, is in some sense absolute.

14.2.2 Gyroscopes and orientation

Knowing and controlling the orientation, or attitude, of a spacecraft is important for many reasons, such as pointing instruments, antennae, and engines in the right direction. It is also needed for dead reckoning, so that the integration of the acceleration and/or velocity is done in a known direction.

Gyroscopes are used in some aircraft to determine orientation. A mechanical gyroscope is a spinning wheel mounted such that its spin axis is free to move in all three dimensions without any torque acting on it (figure 14.2). The angular momentum of the wheel is therefore conserved, so even if the vehicle in which the wheel is mounted rotates, the axis of the wheel remains pointing in a constant direction. By measuring the change in orientation of the vehicle relative to the axis, we can determine the change in the vehicle's orientation.

Mechanical gyroscopes inevitably spin down over time due to friction, so need to be spun up from time to time. An alternative to a mechanical gyroscope is the fibre optic gyroscope, shown schematically on the right of figure 14.2. This comprises a piece of fibre optic wrapped into a circle in a plane. Light from a laser beam is injected into the coil at a point, split into two using a half-silvered mirror, and each part sent in opposite directions through the fibre optic. If the coil is not rotating about the axis perpendicular to the plane, the distance travelled by each beam is the same: once around the circle. Hence the beams arrive back at the injection point still in phase. If the coil rotates clockwise by some angle during the travel, then because the photons are not dragged along with the medium, the beam travelling in the clockwise direction must travel further before it arrives back at the injection point. The anticlockwise beam, in contrast, travels less than the full circle. The beams will now be out of phase when they arrive back at the injection point. The phase difference, measured by interferometry, is a measure of the rotation angle in a given time period. By monitoring the change in the angle continuously over time, we can not only track the orientation but also determine the angular velocity and acceleration.

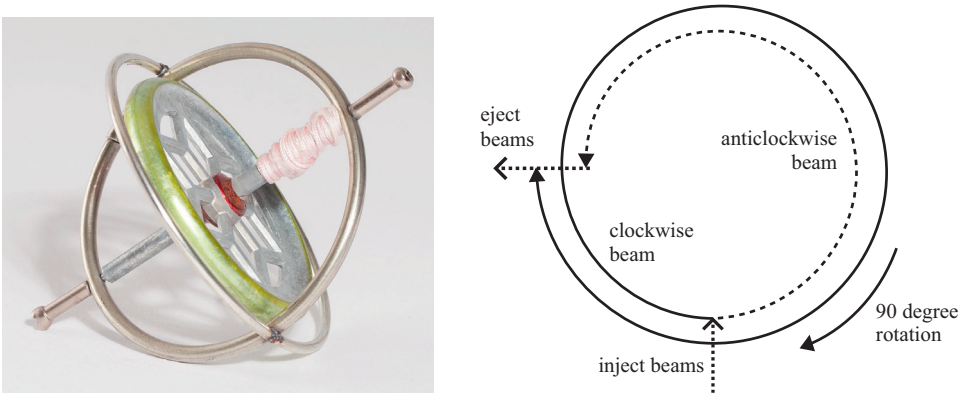


Fig. 14.2 A mechanical gyroscope (left; licensed under CC0-1.0, from Wikimedia Commons) and the principle of a fibre optic gyroscope (right).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

This principle underlying the fibre optic gyroscope is called the Sagnac effect and works by exploiting the constancy of the speed of light in a given medium. In order to build up a more easily measurable phase shift, a fibre optic gyroscope in practice comprises a long (kilometres) loop of fibre optic coiled up many times to increase the difference between the path lengths. Despite this, fibre optic gyroscopes are typically compact (centimetres in size) and, by having no moving parts, are more robust than mechanical gyroscopes.

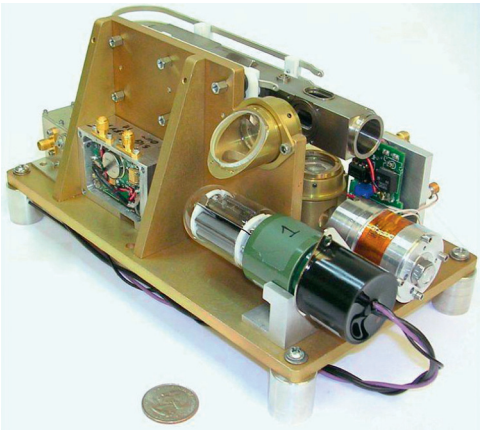
Often a set of three mutually perpendicular accelerometers and three gyroscopes are combined together into a single unit called an *inertial measurement unit* (IMU). IMUs are used extensively for navigation in cars, ships, aircraft, and missiles, often as a complement to satellite navigation (see section 14.6), in particular when satellite signals are not available, such as in tunnels or from fear of interference. Many smartphones and digital wearables for fitness also have IMUs.

14.2.3 Clocks

Onboard clocks are vital for many purposes. They are an essential part of inertial navigation for integrating the linear accelerations measured by an accelerometer. As we will see below, accurate clocks are also required for pulsar navigation, and in satellite navigation systems.

Wrist watches that use a quartz oscillator to maintain time are accurate to about 10 s per year, or one part in 3×10^6. A watch or phone connected to the internet appears to be more accurate than this because it can be regularly synchronized to a reference clock.

The most accurate manufactured clocks are atomic clocks, which exploit the universality of the energy levels in atoms. Electrons transitioning between energy levels release photons of very precise frequency. The constancy of this is even used in the SI definition of the second. Atomic clocks use the hyperfine transitions between energy levels that are split due to electromagnetic interactions between the nucleus and the electrons. These transitions are in the microwave region. In the case of the definition of the second, it is a 9 GHz transition in caesium-133.



A NASA deep space atomic clock based on the hyperfine transition of mercury. Credit: NASA. Fig. 14.3

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

An atomic clock commonly used in spacecraft is based on a transition in rubidium-87, which results in an accuracy of 1 part in 10^{12}, or $30\,\mu s\,yr^{-1}$. Another technology is hydrogen maser clocks that use the 1.4 GHz (21 cm) transition in hydrogen. These are used in the Galileo satellites for Earth-based navigation, where they have an accuracy of 0.45 ns in 12 hours, or 1 part in 10^{14}. A more recent development is based on mercury–iron with an accuracy of 1 part in 10^{15}, or $0.03\,\mu s\,yr^{-1}$. Perhaps the best accuracy achieved so far in an experiment – not yet in a deployed clock – is based on strontium that has achieved an accuracy of 1 part in 10^{18}.

14.2.4 Use in space travel

The main disadvantage of dead reckoning for navigation is the steady accumulation of errors with time. Commercial accelerometers for navigation have accuracies of order $\delta a = 10^{-4}\,m\,s^{-2}$. Over a period of time t at constant acceleration, the errors in velocity and distance from dead reckoning are δat and $(1/2)\delta at^2$, respectively. These are $0.4\,m\,s^{-1}$ and 650 m, respectively, after one hour, and $3.2\,km\,s^{-1}$ and 0.3 au after one year. These are too high for many applications.

Nonetheless, if we can get periodic external determinations of position and velocity, then inertial navigation can be reset and used to provide updates to the position and velocity over the shorter timescales between updates. Inertial navigation was used on the Apollo missions to take humans to the Moon, for example, in particular for the periods when the spacecraft had no contact with the Earth. (The astronauts were also provided with sextants to enable a manual determination of the orientation of the spacecraft.) IMUs are frequently used even in modern spacecraft, for example in the New Horizons spacecraft that encountered Pluto in 2015. Inertial methods are particularly useful during impulsive manoeuvres, when the ranging techniques that we'll discuss in the next section would take too long to provide a new determination.

14.3 Navigation in the solar system by ranging

Navigation of spacecraft in the solar system is routinely assisted by tracking them from the Earth. By sending a signal to a spacecraft, which it then returns, the spacecraft's distance from the transmitter can be computed from the time delay. This is called *two-way ranging*. The radial velocity is determined from the Doppler effect. One-way ranging – in which the time delay over just one leg of the journey is measured – can also be used, but this suffers from any discrepancies between the clocks on the spacecraft and ground station, so is generally less accurate. It may be necessary to use this for distances more than several light hours, however, because the ground station may lose line-of-sight contact to the spacecraft as the Earth rotates. Alternatively a second ground station could be used, to give three-way ranging.

For sending ranging signals, ESA and NASA generally use X-band radio at 7–8 GHz. Sometimes the higher-frequency Ka-band (32–34 GHz) is used as this is dispersed less by the ionosphere and solar wind (recall the plasma frequency discussed in section 12.3.1), but it is more attenuated by dust and water in the troposphere. Using a 35 m radio telescope with a minute of exposure time to collect many signals, the accuracy of these ranging systems is

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

about 1 m in distance and 0.1 mm s^{-1} in radial velocity. We need to know the radial velocity v to get an accurate range, because if the spacecraft is at a distance d, the light travel time is $t = 2d/c$, during which time the spacecraft moves $vt = 2dv/c$. With $v = 1\,\mathrm{km\,s^{-1}}$ and $d = 1$ au, for example, this is a displacement of 1000 km.

The accuracy of the angular position of the spacecraft determined from the radio telescope during ranging is relatively poor. The spacecraft's transverse position can be determined more accurately by measuring the radial distance and velocity as a function of time, and then fitting a model for the motion of the spacecraft. This model will typically be a Keplerian orbit (if the spacecraft's engines are not firing), although deviations from this due to the non-sphericity of the Earth or planetary perturbations, for example, may also have to be fitted. During critical parts of a mission, such as orbit insertion or target encounter, this modelling may not be accurate, especially if the position and velocity are changing rapidly.

Angular coordinates can be determined better for distant spacecraft using *differential one-way ranging* (DOR). Two ground stations A and B with known positions receive a signal from the spacecraft (figure 14.4, left). The difference in the arrival times Δt gives the direction towards the satellite in the plane containing the spacecraft and the two ground stations through simple trigonometry:

$$\cos\theta = \frac{c\Delta t}{b}\,. \tag{14.1}$$

This assumes that the wavefronts arriving at the two ground stations are parallel, which is a good approximation if the distance to the spacecraft is much larger than the separation between the ground stations.[1] The second angular coordinate (out of the plane) can generally be found by waiting for the baseline to rotate as the Earth rotates, or by using a third ground station to define a second plane. This method is called one-way ranging because it does not

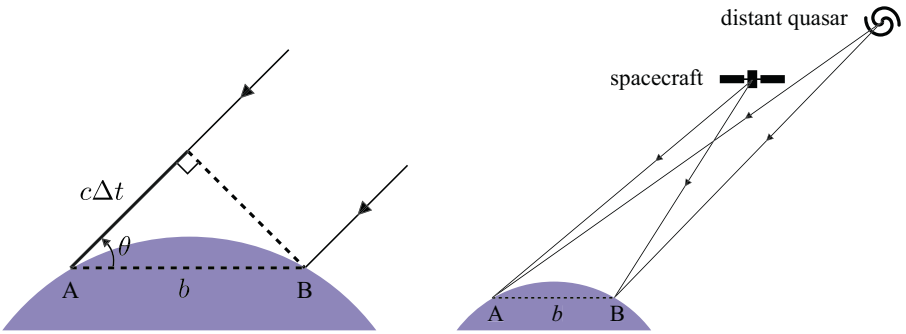


Fig. 14.4 An illustration of DOR (left) and delta-DOR (right) for finding the angular coordinates of a spacecraft. The right diagram is not to scale. The distances to the spacecraft and quasar are much larger, such that the incoming wavefronts from the spacecraft to A and B are virtually parallel, and likewise for the quasar, such that the geometry in the left panel still holds to a good approximation. The angular separation between the spacecraft and quasar is of the order of a few degrees.

[1] If we had no idea of the distance or direction to the sources, then a measured difference in the arrival times would only allow us to place the spacecraft on a hyperbola with the semi-major axis equal to half the ground station separation. This is the basis of *hyperbolic navigation*, which has been used historically for terrestrial navigation over small enough areas such that the Earth's curvature can be neglected. A third 'ground station' in the plane constrains the position to the interception of two hyperbolae, which is just two points, additional information then normally being used to choose between them.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

rely on the signal being sent back, and it is differential because it only uses the difference in the path lengths. As the baseline is very large, often transcontinental, this is also called very long baseline inferferometry (VLBI), which is commonly used in astronomy to get high-resolution radio maps rather than for ranging (see section 11.2.2).

The accuracy of DOR is limited by a number of factors. When the spacecraft is at a finite distance, the incoming wavefronts are not parallel and the time delay does not strictly follow from the simple geometry in the left panel of figure 14.4. The error in the angular position from making this assumption is very small, however (see exercises). When the ground stations are separated by 1 Earth radius, and the spacecraft is 0.1 au away (corresponding to a parallax between the ground stations of $88''$), then when the spacecraft is positioned at $\theta = 45°$ from the baseline (which maximizes the error), the angle θ inferred from equation 14.1 is wrong by only 2.3 mas. With the spacecraft at 0.5 au, this drops to 0.1 mas.

More significant factors limiting the accuracy of DOR include clock timing errors and instrument noise, as well as phase delays introduced by the fact that the waves take different paths through the solar plasma and the Earth's atmosphere to the two ground stations. These problems can be reduced by repeating the measurement on a source with well-known angular coordinates that lies a small angular distance (order of degrees) from the spacecraft as seen from the Earth. This is illustrated in the right panel of figure 14.4. Typically quasars with positions known to a fraction of a milliarcsecond are used. Quasars don't send pre-arranged signals, of course. But their intensities are not constant, and although their variations are essentially white noise over short timescales, it is nonetheless the same signal that is received at the two ground stations, just delayed in time. So by correlating the two signals, the difference in arrival times can be determined. By taking a DOR measurement on the spacecraft and then on the quasar, then taking the difference, many of the systematic error sources cancel. This technique is known as *delta-DOR*.

Delta-DOR has been used in various interplanetary missions, for example NASA's Mars Science Laboratory (the Curiosity rover), which arrived at Mars in 2012. This achieved a timing uncertainty of 0.06 ns over an 8000 km baseline between ground stations, which corresponds to an angular uncertainty of 2.2 nrad, or 0.5 mas (using the error propagation of equation 14.1). This in turn corresponds to a transverse position accuracy of 350 m for a spacecraft 1 au from the Earth (the distance is found by two-way ranging).

Beyond determining the position and velocity of the spacecraft relative to the ICRF, there will be times when we also need to accurately determine the position and velocity relative to the target, for example during flyby or at rendezvous. In that case the spacecraft must observe the target directly. As an example, New Horizons used its onboard cameras to determine the direction to Pluto (and its largest moon Charon), during its flyby in 2015. Together with onboard IMU data and Doppler ranging information from the Earth, this led to an improved determination of the spacecraft velocity and position relative to Pluto. The velocity in three dimensions was then established to within 1 mm s^{-1}, and the two-dimensional position transverse to the line of sight to Pluto was established to an accuracy of $1'' \times d$ in each direction, where d is the Pluto–spacecraft separation, $1''$ (5 μrad) being the resolution of the camera. This corresponds to 5 km when $d = 10^6$ km, for example. The sixth component – the distance to Pluto – was less well determined. Of particular importance was the remaining time to encounter, and here the acceleration of the spacecraft due to Pluto's gravity had to be taken

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

into account. The imaging data taken from the spacecraft allowed the uncertainty in this to be reduced to about 20 s, one day before the encounter.

The Pluto flyby was a success, passing just 12 500 km above the surface of Pluto at a relative velocity of 14 km s^{-1} when it was 34 au from the Sun. Four years later in 2019 this feat was bettered, when New Horizons steered to within 3500 km of the Kuiper belt object 486958 Arrokoth.

14.4 Stellar astrometric navigation

Interstellar spacecraft will be too far from the Earth for the use of ranging methods. They instead have to navigate autonomously. Here we will see how, by measuring only the angular positions of stars, a spacecraft can establish its 3D position and 3D velocity.

14.4.1 Principle

To determine our position in a coordinate system, we need to determine our location relative to one or more objects of known position in that coordinate system. Consider two such objects, or beacons, on the surface of the Earth, shown as A and B in figure 14.5. By measuring the angle to these relative to a reference direction (e.g. north, perhaps found using a compass), we can determine our position S uniquely. If we do not know the reference direction, then a third beacon – C in the figure – can effectively take its place (it does not have to lie in the reference direction). In three dimensions, the number of beacons required depends on the type of measurement we make. If we can only measure the angles between pairs of beacons, then provided we do not lie in a common plane with all three beacons, we can measure three independent angles and from this determine our 3D position. If we instead can measure the two angular coordinates (like right ascension and declination) to a beacon, then from two beacons we can acquire four measurements, which suffice for us to determine our position.

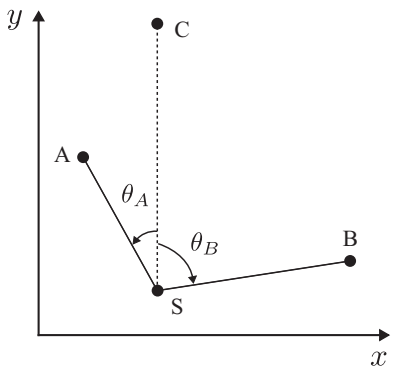


Fig. 14.5 Navigation by beacons in two dimensions to find the position of S. If we know the positions of two beacons, A and B, and measure the angles θ_A and θ_B between them and a reference direction (dotted line), the location of S can be found.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

We can use these ideas to determine our position in interstellar space. Specifically, by measuring just the angular positions of stars with known 3D positions in the ICRF, we can determine our position in the ICRF. For this to work we need accurate 3D positions of those stars.

For over two thousand years, astronomers have catalogued the two angular coordinates – right ascension and declination – of stars in the sky. Only much more recently could the third dimension – the distance – be determined with any precision. The fundamental way to do this is with parallax, which exploits the annual motion of the observer. As the Earth moves around the Sun, the apparent position of a nearby star shifts with respect to very distant stars (or even better, extragalactic quasars). As we know the size of the Earth's orbit very accurately, the measured angular shift – the parallax – gives the distance to the star via simple trigonometry.

The nearest stars have parallaxes below $1''$, so their shifts were unobservable to early astronomers. It was not until the 1830s that the first statistically significant measurements were reported, first by Friedrich Bessel in 1838 with a measurement of 314 ± 20 mas for the binary star 61 Cygni. Parallaxes of more stars with ever-increasing accuracy were collected during the twentieth century, culminating in the first mission dedicated to astrometry from space. The satellite Hipparcos measured parallaxes for 120 000 stars to an accuracy of around 1 mas. This was followed by Gaia, which between 2014 and 2025 observed around two billion stars, achieving parallax uncertainties as low as 10 μas, and effectively extended the horizon of good parallaxes from around 100 pc to a few kiloparsecs.[2]

As a result of these surveys, we now know the distances of all but the faintest stars within 100 pc to within 0.3% (median accuracy), with a higher accuracy for the nearer and brighter stars. A 0.1% distance accuracy for a star at 10 pc is about 2000 au, for example. The angular positions of the stars are typically known to within a few microarcseconds, so their transverse positions are known much better than their distances: 10 μas at 10 pc corresponds to 10^{-4} au, or just 15 000 km. By the time interstellar spacecraft are built, we can assume that the distances to potential navigation stars will be determined even more accurately.

These nearby stars can be used as beacons by an interstellar spacecraft to determine its position in space. This is complicated, however, by the fact that the spacecraft is moving with respect to the stars. As we saw in section 9.1, even a small non-relativistic velocity introduces a significant aberration. As we would not know the velocity of the spacecraft with respect to each star, it seems we could not account for this. However, if we know the velocities of the stars relative to the ICRF, then we can turn the problem around and use the measured aberration to determine the velocity of the spacecraft (if we also know the position of the spacecraft, that is; we'll get to this circular problem shortly).

Fortunately, we do know the 3D velocities for many stars. As part of its astrometric observations, Gaia also determines the proper motions – angular velocities – of the stars. It has to do this, in fact, in order to account for the motion of the stars over the several years of its observations. The median accuracy for each proper motion component for stars within 20 pc is about $30\,\mu\text{as yr}^{-1}$, which for a star at 20 pc corresponds to a transverse velocity accuracy of just $3\,\text{m s}^{-1}$. Gaia also measures the radial velocities for many stars to an accuracy of a few kilometres per second, although these can also be obtained from ground-based observations, which can achieve an accuracy of a few tens of metres per second.

[2] In its third data release (DR3), Gaia determined the parallaxes of the two components of 61 Cygni to be 285.995 ± 0.060 mas and 286.005 ± 0.029 mas. Bessel was off by less than two sigma.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

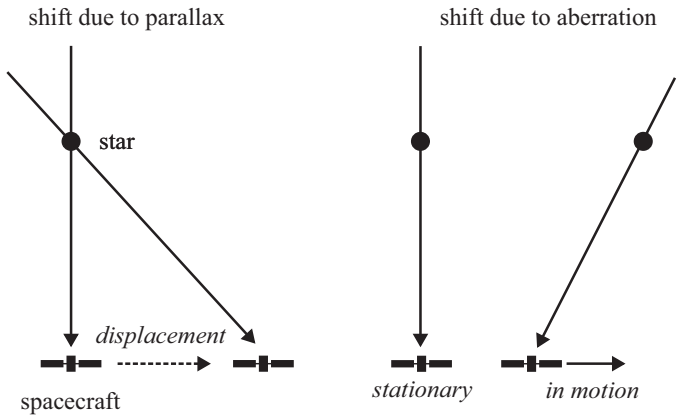


Fig. 14.6

Parallax and aberration induce different shifts in the observed star location. In both cases the star initially appears straight up from the stationary spacecraft. Left: When the spacecraft has moved to the right (and is again at rest), the star appears to move to the left, by an amount that depends on the size of the displacement and the distance to the star. Right: If the spacecraft acquires a velocity, then before it has even moved, the star appears to shift to the right due to aberration (see figure 9.1) by an amount that depends on the velocity but is independent of distance.

The overall problem, then, is to determine both the 3D position and 3D velocity of the spacecraft in the ICRF from the known 3D positions and 3D velocities of a number of stars in the ICRF. When a spacecraft is stationary relative to the ICRF, but displaced some distance away from its origin (the SSB), the observed positions of the stars shift because of the parallax effect. But the finite velocity of the spacecraft also creates shifts due to aberration. To interpret the shifts as due to aberration, and thus velocity, we would need to know the spacecraft's position. And to interpret the shifts as due to parallax, and thus position, we would need to know the spacecraft's velocity. However, the effects of aberration and parallax are not the same. The size of the parallax depends on the distance to the source, whereas the size of the aberration does not. Furthermore, a displacement of the spacecraft in a given direction, and a velocity of the spacecraft in the same direction, give rise to parallax and aberration shifts with opposite signs, as illustrated in figure 14.6. By measuring the apparent shifts in the positions of several stars, we can disentangle the effects of parallax and aberration and determine both the 3D position and 3D velocity of the spacecraft simultaneously.

14.4.2 Implementation and simulation results

The technique of navigation outlined in the previous section has of course never been applied to interstellar travel. We therefore look at the results of a simulation.

Because of the interdependence of the parallax and aberration, and the inherent nonlinear relationship between the measured angles and the spacecraft position and velocity, an iterative approach is used, as summarized in figure 14.7. It works as follows. We start off by assuming a 3D position ($\mathbf{x}$) and 3D velocity ($\dot{\mathbf{x}}$) of the spacecraft. Then, from the catalogued ICRF coordinates of the beacon stars, we calculate the effects of parallax and aberration to predict the angular positions of the stars that we expect to observe from the spacecraft. We then compare these predicted angular positions to the measured angular positions (which are

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

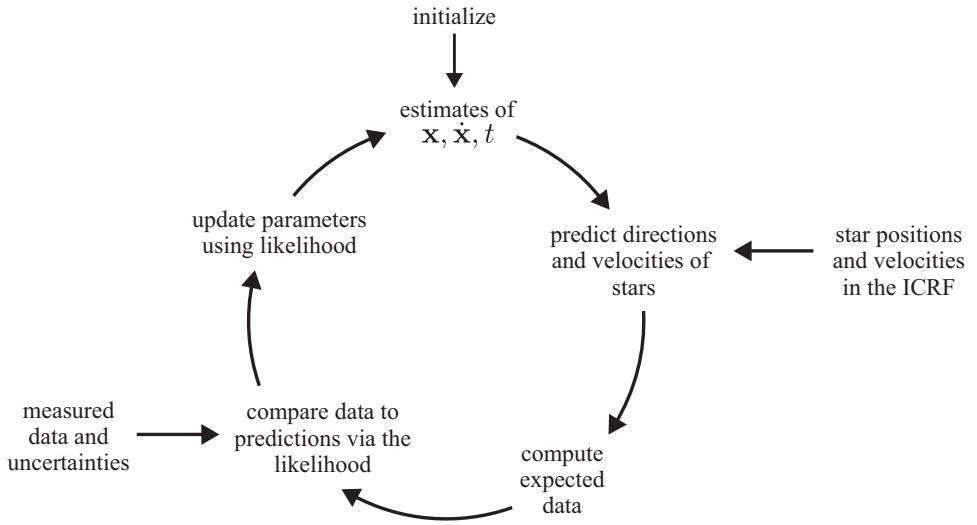


Fig. 14.7 Stellar astrometric navigation. An iterative approach is used to estimate the 3D position ($\mathbf{x}$) and 3D velocity of the spacecraft ($\dot{\mathbf{x}}$) from measurements of the angular separations of pairs of stars (and/or their radial velocities).

simulated here from their true positions with measurement noise added). Then we adjust the spacecraft position and velocity in such a way that aims to minimize the difference between the predictions and measurements. The whole process is iterated until the predictions converge on the measurements, which occurs when the true spacecraft position and velocity has been found (to within the noise).

As there is no angular coordinate grid printed on the sky, we cannot directly measure angular positions of stars, in the sense of right ascension and declination in a coordinate system. In practice we can only measure with a high degree of accuracy the angular separation θ between a pair of stars, as illustrated in figure 14.8. We therefore convert our predicted angular coordinates into predicted angular separations, and compare these to measurements. The angular separations can be measured using a pair of telescopes, or better, using a single camera with two different fields of view. This is in fact the principle of a sextant, traditionally used to determine the angle of the Sun or stars above the horizon (figure 14.9).

There are various ways to implement the iterative scheme, but a good approach is to compare the predictions and the data via a probabilistic likelihood, as this takes into account the measurement uncertainties. In the simplest case, the likelihood is just the product of one-dimensional Gaussian distributions, each one describing the probability of the measurement given its true but unknown value (the one we are trying to infer). The space of the parameters – $\mathbf{x}$ and $\dot{\mathbf{x}}$ – can then be explored with a Monte Carlo method to determine the relative probabilities of different solutions. This produces a probability distribution over the parameters, which we can summarize with a point estimate and uncertainties. An alternative approach is to use an optimization algorithm to search the space to locate just the maximum of the likelihood.

To determine $\mathbf{x}$ and $\dot{\mathbf{x}}$ we need to measure at least six independent angular separations. We can do this by measuring the angular separation between six stars and a seventh reference star. The reference star is arbitrary, but the Sun is a natural choice because for communication

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

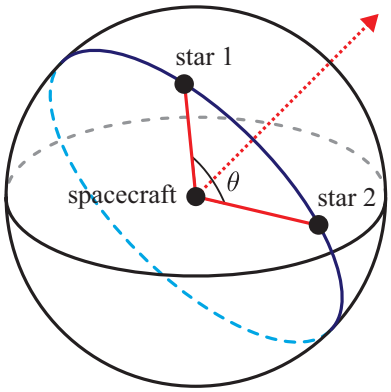


Fig. 14.8 Stellar astrometric navigation relies on measurements of the angular separation θ between pairs of stars. It does not need the separate angular coordinates (right ascension and declination) of the stars.

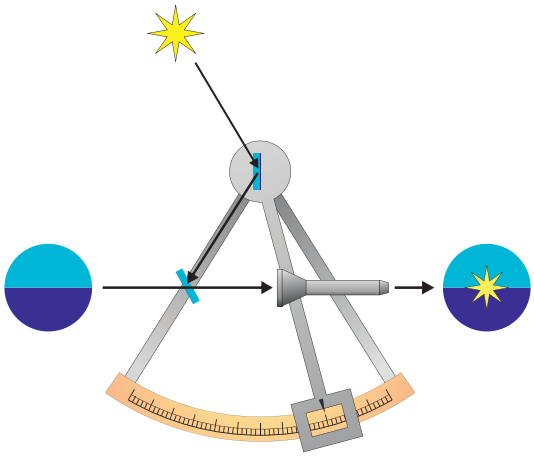


Fig. 14.9 A sextant uses a half-silvered mirror to superimpose the images of two objects – here the horizon and the Sun – onto a common focal plane in order to measure the angular separation between them. Credit: Public domain, from Wikimedia Commons, modified by the author.

purposes, at least, we always need to know where it is. But there is no lack of suitable – bright, nearby – stars, and having more stars spread out evenly over the sky is beneficial.

Note that in this method we do not need to determine either the distances to, or the velocities of, the stars from the spacecraft. These cannot anyway be determined instantaneously from onboard the spacecraft. Only the radial velocity component of the stars can be measured quasi-instantly, using a spectrograph to measure the Doppler shift of the stellar spectral lines. This could, in fact, be used as an alternative measurement to the angular separations (or in addition to them), as it too encodes the position and velocity of the spacecraft via parallax and aberration. But we won't consider it here.

How well does this method perform? Bailer-Jones (2021a) performed various simulations to determine this. As expected, the performance depends on both the number of stars used and

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

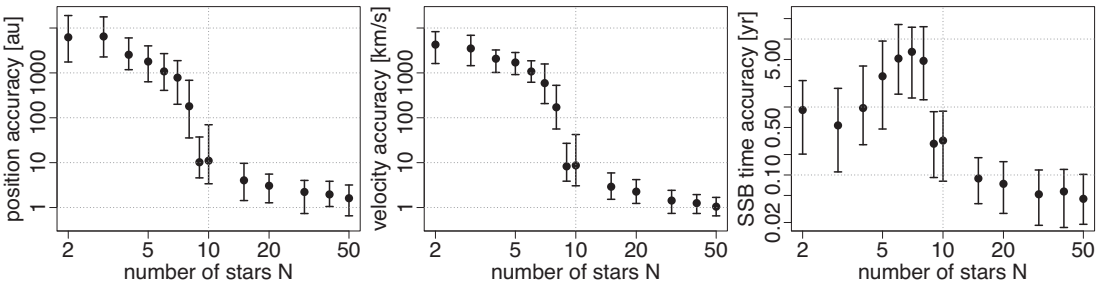


Fig. 14.10 Astrometric stellar navigation. This shows how the spacecraft navigation accuracy (the error in the estimation) varies with the number of stars used, for the overall spacecraft position (left), the magnitude of the 3D velocity (centre), and the SSB (= ICRF) time (right). This assumes angular separations between pairs of stars are measured at a very modest accuracy of 1″. The navigation accuracies scale linearly with this measurement accuracy, so for example they would all improve by a factor of 1000 with measurements accurate to 1 mas. Credit: Adapted from figure 11 of Bailer-Jones (2021a) published under a CC-BY-3.0 licence.

the accuracy of the onboard measurements. These simulations first assume that the angular separations between the N nearest stars to the Sun are measured to a very modest accuracy of 1″. This can easily be achieved with a very small telescope aperture ($\sim$ 1 cm). For a given N, many simulations were made with the spacecraft at different positions up to 10 ly away, and moving at velocities up to 0.5 c. For each simulation, the iterative method was then used to infer the position and velocity of the spacecraft and this compared with the true values, from which the error in the position and velocity was computed.

Figure 14.10 shows how the performance varies with N. The points show the median error over a set of simulations, and the bars show the spread over the set (16th and 84th quantiles). We see that with $N = 10$, the position of the spacecraft can be determined to within 10 au (left panel) and the magnitude of the velocity to within 10 km s^{-1} (centre panel). With more stars the performance improves roughly as $N^{-1/2}$. For smaller N the behaviour is more complex. As there are seven unknowns to determine – three positions, three velocites, and the time (as I will explain shortly) – we need at least seven measurements, that is, $N = 8$ stars. With fewer stars we see a big increase in the errors, as we formally cannot solve for all the parameters (although there is still some information, and as the problem is not linear there is not a complete degeneracy).

A positional knowledge of 10 au may be adequate for some interstellar navigation purposes (Proxima Cen is 269 000 au away) but is too poor for injecting a spacecraft into a star system, where we may want to pass within 0.1 au or less of a planet. One solution is to use more stars, as figure 14.10 shows. However, further simulations show that the position and velocity improve directly in proportion to the onboard angular measurement accuracy. So if angular separations can be measured to 1 mas between 10 stars, then we could determine the position and velocity to within 0.01 au and 10 m s^{-1}, respectively.

The above explanation of the method glosses over one aspect. The stars are moving relative to the ICRF, so their expected positions relative to the spacecraft depend on the time elapsed since their ICRF positions were recorded in the catalogue. Over the course of 50 years, a nearby star moving at 30 km s^{-1} relative to the ICRF moves 0.0015 pc. If the star is at 20 pc and its movement is entirely in the transverse direction, then even if the spacecraft is still at the SSB (the origin of the ICRF), the star is displaced by 16″, which is much more

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

than our measurement accuracy. To compensate for this we need to know the time of the spacecraft's observations.

Time can be measured precisely using an onboard clock (section 14.2.3). But if our spacecraft has been moving at relativistic speeds, the onboard clock will no longer show ICRF time due to time dilation. So if we cannot keep track of all the velocity changes, it may not be possible to know ICRF time onboard our spacecraft with sufficient accuracy. It turns out, however, that if we observe one additional star, we can solve for the ICRF time as part of the iterative method. The right panel of figure 14.10 shows that the time is not actually determined very precisely in this way: to the order of just a few weeks once we have more than about 10 stars with $1''$ measurement accuracy. This is approximately the time it takes the stars to move the distance of the angular measurement uncertainty, so we cannot – and need not – resolve the time to better than this. When we have onboard measurement accuracy of 1 mas, then we can determine the ICRF time to within about an hour. This will prove useful for tracking the position of the Earth for communication purposes (see section 15.3).

It turns out that this astrometric navigation method is very insensitive to the initially assumed position, velocity, and time required to initialize the calculations. The performance is also largely independent of the true positions or velocity of the spacecraft. It should, therefore, be quite reliable.

This method of navigation works by using the known positions and velocities of stars to determine the position and velocity of the observer (spacecraft). The Gaia astronomical survey spacecraft mentioned earlier does the opposite. It uses the known position and velocity of the spacecraft – its orbit in the ICRF – to determine the positions and velocities of the stars. It does this too by measuring angular separations between the stars.

One practical aspect we have not yet considered is the need to identify the correct stars so we can match them with their catalogue data. This is a standard task for satellites, and is solved by the use of star trackers. These are wide-field cameras coupled with software that can identify patterns of the brightest stars (asterisms, typically as triangles). If our spacecraft travels far compared to the typical distances to the beacon stars, the asterisms will change significantly and eventually become unrecognizable (and it is in fact this change that we are exploiting in the navigation). This could be addressed by keeping track of the directions to the relevant stars during the course of the mission. If there is a spectrograph on the spacecraft, the spectral signature can also help identify the stars from their characteristic line features and continuum shapes.

14.5 Pulsar navigation

We now turn to another way of using astronomical objects to navigate autonomously. This makes use of the fact that some objects send out very regular pulses, although any highly regular predictable signal would do. Observations of several such objects in different directions can be used to constrain the position of the spacecraft.

14.5.1 Pulsars

Once nuclear fusion ceases in stars, they evolve in to one of several end states. Stars with masses between $8\,M_\odot$ and $30\,M_\odot$ explode as supernovae, shedding much of their mass. With

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

insufficient radiation pressure to counter the pull of gravity from the remaining mass, they collapse. The gravity is strong enough to overcome the electromagnetic forces that otherwise prevent less massive objects such as planets from collapsing. The electrons create a large degeneracy pressure that, if the stellar mass were low enough, would balance gravity to leave what is known as a white dwarf. But in higher-mass stars, the gravity is large enough even to overcome the electron degeneracy pressure, and the temperature increases enough to split the nuclei via photodisintegration, forcing protons to interact with electrons to produce neutrons. Further collapse is finally stopped by the repulsive strong nuclear force and the neutron degeneracy pressure to create a neutron star. If the mass were even larger, then even this balance would eventually be overcome, and the end state would be the singularity of a black hole.

Neutron stars typically have masses between $1.2\,M_\odot$ and $2.0\,M_\odot$ and radii of 10 km, which gives them an enormous average density of the order of $10^{18}\,\mathrm{kg\,m^{-3}}$. Each cubic centimetre has a mass of nearly a billion tonnes. All stars rotate, and now that the star is orders of magnitude smaller than before it collapsed, conversation of angular momentum produces a very rapid rotation. Typical rotation periods range from milliseconds to seconds. Some neutron stars also possess strong magnetic fields – 10^4 to 10^8 T, up to 10^{11} T for magnetars – and emit radiation at a wide range of wavelengths (gamma to radio) along their magnetic poles, with a spreading angle of a few degrees. If the magnetic axis and spin axis are misaligned, the magnetic pole will sweep out a circle over the sky, and any observer positioned in its path will see regular pulses of radiation from the neutron star (figure 14.11). This is a pulsar. The emitted electromagnetic energy is derived from the rotation of the neutron star, so pulsars are losing rotational energy and slowing down, albeit very slowly. The Crab pulsar (PSR B0531+21) with a period of 33 ms, for example, is spinning down at a rate of 4.2×10^{-13} (a dimensionless number, e.g. seconds per second). This is a very young pulsar, formed from the supernova that occurred in the year 1054 CE (for which historical observing reports exist), and is the central star in the Crab nebula that was produced by this supernova.

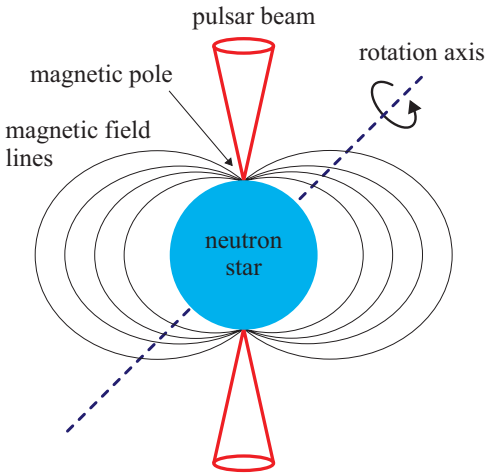


Fig. 14.11 The concept of a pulsar. A neutron star emits X-rays and radio waves in a narrow cone along the magnetic axis. When this axis is inclined to the rotation axis, the electromagnetic radiation is received in pulses, once per rotation, by a distant observer.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

Ninety per cent of rotation-powered pulsars have periods above 20 ms. The rest with shorter periods are called millisecond pulsars. These are older, with weaker magnetic fields, and so smaller spin-down rates, of order 10^{-18} to 10^{-21}. This rate is so small that the constancy of their periodic pulses is even better than the stability of atomic clocks (see section 14.2.3). Just over 2000 rotation-powered pulsars are known, of which 150 are detected in X-rays, and 50 of those are millisecond pulsars. The periods of pulsars can be measured to an accuracy of about 10^{-15}. This is crucial for being able to use their pulses in navigation, as we will see shortly.

The are others classes of neutron stars that emit pulses and so appear as pulsars. Accretion-powered neutron stars are neutron stars in binary systems in which mass, and therefore angular momentum, is accreted by the neutron star. This causes strong X-ray emission (but not radio emission), the pulses being caused by the rotation of the accretion hotspot. However, the accretion and the pulses are irregular, so they are not suitable for navigation. Another type of neutron star, a magnetar, is a single star with a very strong magnetic field. Their pulses vary in ways that are not understood well enough for them to be reliable beacons for navigation at this time.

14.5.2 Navigation principle

The rough idea of pulsar navigation is as follows. If we can precisely measure the arrival times of pulses from known pulsars, this constrains the position of the spacecraft in the radial direction to the pulsar, with one important caveat that we'll get to. Repeating this for at least two more pulsars in different known directions, we can locate the spacecraft in three dimensions.

Pulsar navigation requires us to know the period and phase of the pulses, such that we can predict when pulses will arrive at the origin of our coordinate system, which we take to be the SSB. We then compare the measured arrival time of a pulse at the spacecraft with the expected arrival time at the SSB. Suppose at first that the pulse period is very long, so we only have to deal with one pulse, and that the difference in arrival times at the spacecraft and the SSB is Δt. In figure 14.12, the known direction to the pulsar is $\hat{\mathbf{x}}_p$ (a unit vector) and the unknown position vector of the spacecraft S relative to the SSB is $\mathbf{x}_s$. Assuming the pulsar is infinitely far away, the pulse will arrive simultaneously at points S and A. The extra time needed for the pulse to travel from A to the SSB is $a = c\Delta t$. From the scalar product we see that

$$\hat{\mathbf{x}}_p \cdot \mathbf{x}_s \;=\; |\hat{\mathbf{x}}_p||\mathbf{x}_s| \cos\theta \;=\; a, \tag{14.2}$$

where θ is the angular separation between the spacecraft and pulsar as seen from the SSB. If we have three pulsars in different directions, then we can write down the three corresponding equations together in matrix form as

$$\mathbf{X}_p\, \mathbf{x}_s \;=\; \mathbf{a} \tag{14.3}$$

where $\mathbf{X}_p$ is the 3×3 matrix containing $\hat{\mathbf{x}}_p$ for each pulsar in each row, and $\mathbf{a}$ is the vector of distances computed from the three time delays. Provided the three pulsars and the spacecraft do not lie in a common plane, the matrix is invertible and we get a unique solution for the spacecraft position. If we have more than three pulsars, we can solve equation 14.3 by least squares to also find a unique solution.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

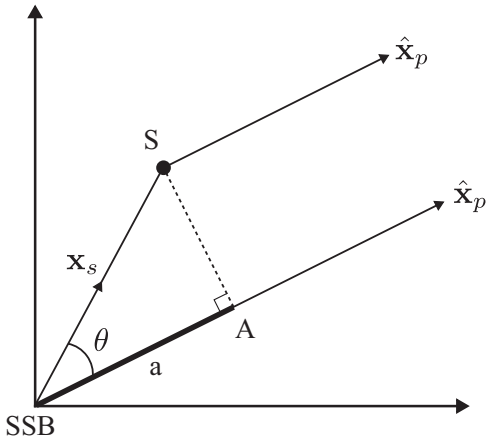


Fig. 14.12 The geometry of pulsar navigation when assuming the pulsar is infinitely distant in direction $\hat{\mathbf{x}}_p$ (unit vector). The spacecraft S is at position $\mathbf{x}_s$ relative to the SSB.

Now to the caveat. Pulsars emit pulses continuously. These pulses are indistinguishable, so we cannot know when a particular pulse is received at the SSB and compare that to the corresponding pulse at the spacecraft. All we know is that pulses are received at the SSB every P seconds, where P is the period, relative to some known reference time t_0. That is, pulses arrive at the SSB at times $t_0 + nP$ where n is an integer. When the spacecraft receives a pulse, it does not know the value of n. Thus we cannot measure the actual time delay Δt, but instead only the delay with respect to the nearest cycle, expressed as a phase shift f (fraction of a cycle; see top of figure 14.13). Thus in figure 14.12 we have $a = cP(f+n)$ for unknown n.

The consequence of this phase ambiguity is that measuring the first pulsar only places the spacecraft on one of an infinite number of parallel planes with separation cP, as shown in the lower left of figure 14.13. If we know the initial position of the spacecraft to within a distance $cP/2$, then we know which cycle we are in – the value of n – and thus which plane the spacecraft is on. A second pulsar in another direction constrains the position to a second set of infinite planes, inclined to the first, and the intercept of these planes with the first set of planes constrains the position to a set of lines (pointing in/out of the page). A third pulsar constrains the set of possible spacecraft positions further (bottom right of figure 14.13), and if this pulsar lies out of the plane of the page, it limits the set of points in the third dimension too. The distance between these solutions will be smaller or larger depending on the orientation and periodicity of the planes. With sufficiently accurate knowledge of the initial position of the spacecraft (from astrometric navigation, or dead reckoning), a single most-probable solution can be found. On the other hand, if the initialization is less precise than the separation between the solutions, a phase ambiguity, and therefore a spacecraft position ambiguity, will remain.

Reality is of course more complicated. The pulses are not perfectly narrow spikes, and the signals are weak, so the spacecraft has to measure more than one pulse in order to get a sufficient signal-to-noise ratio to determine the phase shift. Consequently the planes in figure 14.13 have finite width and their interceptions are finite-sized blobs. There are also uncertainties in the pulsar periods and directions. All these propagate into uncertainties in

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

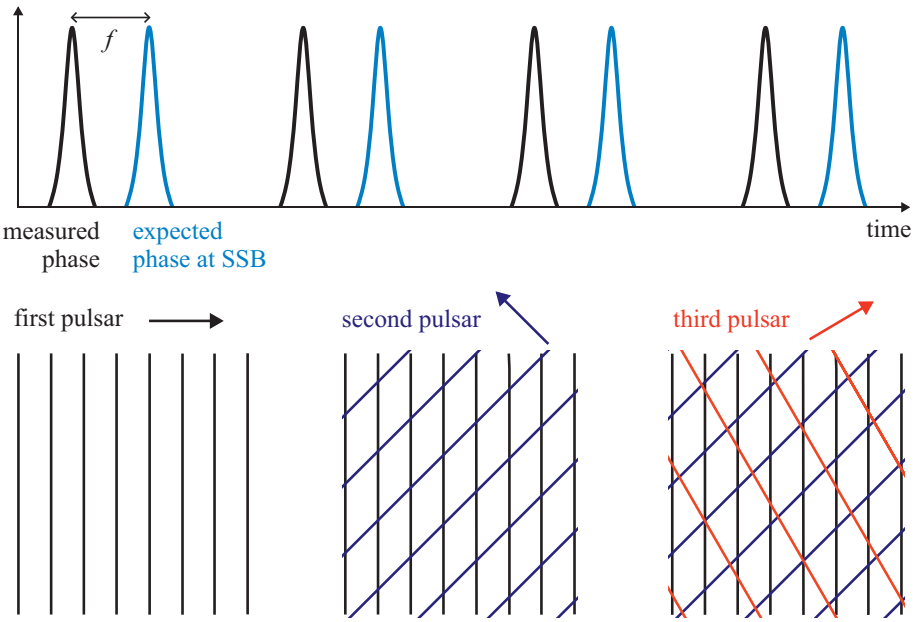


Fig. 14.13

Pulsar navigation. Top: the times of arrival of the pulses at the spacecraft ('measured phase') are offset in phase from their arrival times at the SSB by a large number of integer periods plus a fraction of a period, but only that fraction – the phase difference f – can be measured. Bottom: Measuring the arrival times from one pulsar would therefore place the spacecraft on any one of an infinite number of parallel planes (bottom left). Adding a second pulsar (middle) and then a third pulsar (right) helps to solve resolve the ambiguity by reducing the volume of plausible solutions where the planes intercept.

the spacecraft position. Because the spacecraft has to measure multiple pulses, it will move during the course of its measurements, and this too has to be taken into account. Some knowledge of the spacecraft's velocity is therefore necessary. Note that the pulses have a different period at the spacecraft than in the SSB due to the Doppler effect, and the pulsars will also appear to lie in different directions due to aberration, so the velocity of the spacecraft must be known or determined also to account for these.

In practice an iterative procedure is required. As with astrometric navigation (section 14.4), this is initialized with an assumption of the spacecraft's position and velocity relative to the SSB. These are used to predict the phase offset, and from the difference with respect to the measured offset we can identify the best solution. This solution involves a trade-off between reproducing the measured phases as well as possible to within the uncertainties, but also being as close to the initialization as possible (within its uncertainties). As with astrometric navigation, we can regard this as a probabilistic inference problem, with a likelihood defined by the phase measurements, and the prior probability corresponding to the initialization.

This navigation method relies on measuring the time of arrival of pulses on a reference clock that is keeping track of SSB time. An accurate onboard clock is therefore essential. However, as the spacecraft has been moving and accelerating, we have variable time dilation, and it will be almost impossible to compute precisely enough all the relativistic corrections

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

required to keep the clock synchronized with the SSB. However, by observing a fourth pulsar we can also solve for the SSB time, or rather a correction to the onboard clock that resynchronizes it with SSB time.

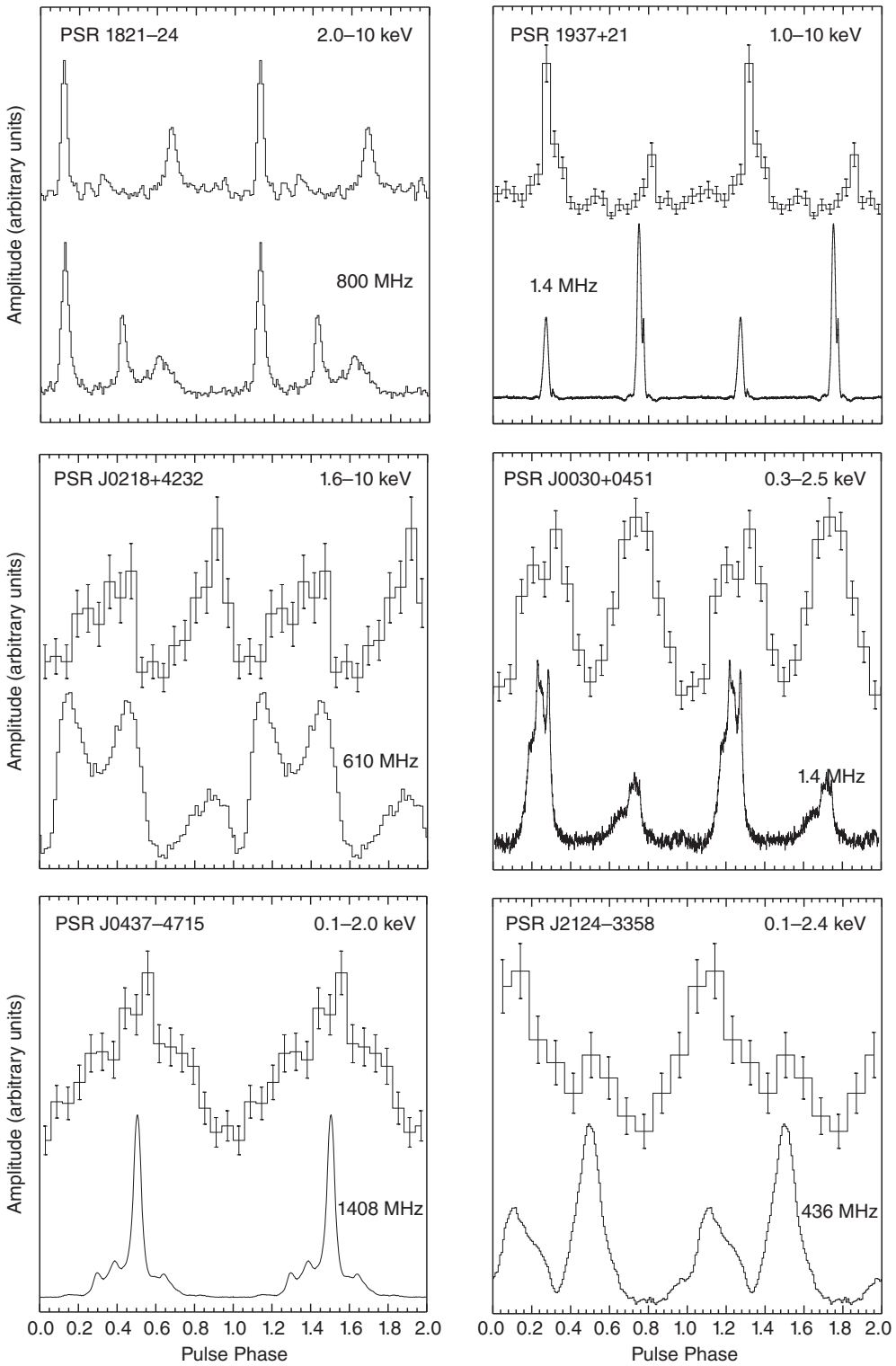
An important part of the determination of the pulse arrival times is measuring and fitting the profiles of the received pulses. Pulsars give out periodic signals at both radio and X-ray wavelengths, and we could measure them at either wavelength. The pulses are not always simple shapes or even single profiles, as can be seen in figure 14.14, and they can also change with time. They must therefore be fitted against a reference pulse profile for that particular pulsar. X-ray pulsars are relatively faint, so the determination of the pulse profile is limited by the photon statistics (note the error bars in figure 14.14). This is less of an issue in the radio, although even here a large telescope (10 metres or more in diameter) is needed, which may not be compatible with low-mass interstellar spacecraft. Instead of a single dish one could use a phased array antenna (see section 15.8.3). This too would be large and also requires a lot of computer power for the processing. But phased arrays have the advantage of being able to see half the sky – to observe multiple pulsars – without having to be physically steered. Another aspect that needs to be taken into account is the dispersion of electromagnetic waves by charged particles in the ISM, which cause the pulses to spread in time, making it harder to determine an accurate time of arrival. This affects the radio more than the X-ray, which is another reason to prefer the use of X-ray wavelengths.

14.5.3 Application within the solar system

Several groups have considered how pulsar navigation could work in practice, and using simulations based on real pulsars have estimated how accurately the spacecraft position could be determined. This accuracy depends on the size of the instrument that records the pulses, the brightness of the pulsars, how accurately their positions are known, and also on their distribution across the sky (all three lying in the same direction gives little information on two of the dimensions, for example). Shemar et al. (2016) assume an X-ray receiver with an effective area of 50 cm^2 operating at 1 keV that observes each pulsar for 10 hours. They identify a set of 10 suitable pulsars with periods between 2 ms and 150 ms. Using three pulsars and an accurate atomic clock, they conclude that a 3D positional accuracy of between 30 km and 100 km can be attained depending on which pulsars are used. A velocity accuracy of the order of 0.5 $m\,s^{-1}$ for the spacecraft can also be obtained. These results assume that the spacecraft is no more than 30 au from the Sun. At larger distances, the parallax of the pulsars becomes an issue (see next section). For comparison, the transverse (plane of sky) position errors for a spacecraft at 30 au obtained with a 70 m antenna on the Earth via ranging is around 4.5 km.

The instrument must have a good time resolution in order to record the pulses; here it is assumed to be 1 μs. The X-ray telescope could either be a photon counter with a collimator tube to select the source, or an imaging instrument. The latter potentially offers a better signal-to-noise ratio by allowing the sky around the target to be removed. But these are quite bulky, because a long focal length is generally needed to focus the X-rays. For example, the MIXS instrument on the BepiColombo mission to Mercury, which also has a collecting area of 50 cm^2, is about 1 m long and 0.5 m in diameter. Shemar et al. (2016) outline an analogous instrument optimized for pulsar timing of similar size with a mass of around 12 kg and power consumption of 16 W.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015



Pulsar profiles in the X-ray (upper curve in each panel) and radio (lower curve) for the six brightest millisecond pulsars. The pulse amplitude in arbitrary units vs phase is plotted. Two identical phases are shown. Credit: Used with permission of Springer Nature BV, from Becker (2009), conveyed through Copyright Clearance Center, Inc.

Fig. 14.14

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

Pulsar navigation has been tested in the real world from Earth's orbit. NASA attached an X-ray experiment called NICER to the International Space Station in 2017. Part of this is the SEXTANT (Station Explorer for X-ray Timing and Navigation Technology) experiment. Using several millisecond pulsars, it could determine its position to within 10 km (Yu et al. 2020). China launched an experimental X-ray navigation satellite called XPNAV-1 into an Earth orbit in 2016. From observations of the Crab nebula it was able to determine its position at certain control points in its orbit to within an average of 38 km over an 85 day period (Huang et al. 2019).

So far we've been assuming that pulsars have a constant period, or at least that the period is changing in a known way. This is generally true for millisecond pulsars, which is why they are selected for navigation. Other pulsars, and at least one millisecond pulsar, have exhibited 'glitches', which are sudden changes in the rotation period by a factor of around 10^{-6}, although sometimes much smaller. These are believed to be a result of a redistribution of mass, and thus moment of inertia, within the star, at constant angular momentum. Such glitches could be catastrophic for navigation if they were not noticed.

14.5.4 Pulsars for interstellar navigation

Up until now we have been assuming that the pulsar is so distant that the direction to the pulsar as seen from the spacecraft is the same direction to the pulsar as seen from the SSB (as in figure 14.12). This approximation no longer holds once the spacecraft is far from the SSB, as we can see in figure 14.15. The difference in the spacecraft–pulsar and SSB–pulsar distance is, for small angle α,

$$\Delta x = \frac{x_p}{\cos\alpha} - x_p = x_p(\sec\alpha - 1) \simeq \frac{x_p\alpha^2}{2} \simeq \frac{x_s^2}{2x_p}\,. \tag{14.4}$$

The Crab pulsar, for example, is about $x_p = 2$ kpc away. When the spacecraft is $x_s = 30$ au from the SSB, then $\Delta x = 160$ km (corresponding to $\alpha = 15$ mas). Once Δx is more than the uncertainty in the spacecraft position derived from the navigation, the assumption of a common direction no longer holds. As we expect navigation uncertainties of tens of kilometres, the assumption breaks down once the spacecraft is more than a few tens of astronomical units from the SSB.

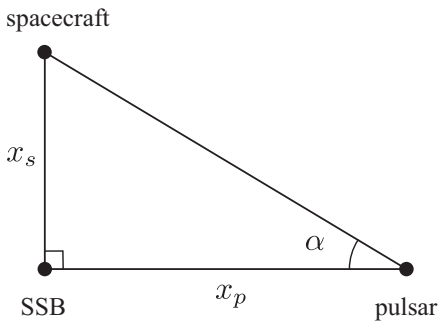


Fig. 14.15 Navigation using pulsars at finite distances.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

To use pulsars for interstellar navigation, we therefore need to have some idea of their distances, although not necessarily as accurately as we did when measuring the angle between stars for navigation in section 14.4. Knowing the distances, we modify the expected phase difference when solving for the spacecraft position (and we would use the modified directions to the pulsars too). Unfortunately, other than the Crab pulsar, all other known rotation-powered pulsars are too faint in the optical to have had their parallaxes (and thus distances) measured by Gaia. Distances to some pulsars have been determined to a greater or lesser degree of precision in the radio, however. As the spatial resolution of a telescope aperture is of the order of λ/d, where λ is the observing wavelength and d is the diameter of the telescope, then the parallax precision from a single dish ($d \leq 100\,\mathrm{m}$) is very low in the radio. However, by linking many dishes over thousands of kilometres via VLBI, it is possible to obtain parallax precisions at the milliarcsecond level.

There are other complicating factors when the spacecraft is at large distances from the SSB. The path through interstellar space to the spacecraft is now different from the path to the SSB, so the pulses will be distorted by charged particles in the ISM in different ways, potentially leading to systematic errors in the measurement of the arrival times. A further issue is that a pulsar is only seen as a pulsar because the Earth happens to lie in the rotating beam of the neutron star (figure 14.11). Once the spacecraft is sufficiently far from the Earth, it will no longer fall in the beam. However, as the opening angle of pulsars appears to be on the scale of degrees, this is unlikely to be a problem for early interstellar travel.

Finally, the finite proper motion of a pulsar will change its observed direction over time. Some pulsars have been ejected from binary systems at velocities of hundreds of kilometers per second, and for pulsars 1 kpc away, $500\,\mathrm{km\,s^{-1}}$ correspond to a proper motion of $100\,\mathrm{mas\,yr^{-1}}$. Thus within a year the position will change by more than the positional uncertainty. Of course, if the proper motion has been measured, it too can be incorporated into a time-dependent model of the pulsar.

14.6 Relation to global navigation satellite systems

Pulsar navigation bears some relation to global navigation satellite systems (GNSS), such as the Global Positioning System (GPS), for determining locations on the Earth. Yet there are important differences.

GNSS works by a navigator receiving signals from several satellites. These satellites are in low-Earth orbits that are known to high accuracy. The signal from a given satellite encodes both the satellite's position at the time the signal was sent, as well as this time, which is provided by an accurate clock onboard the spacecraft. If the navigator had an equally accurate clock synchronized to the satellite's clock, then the difference between the transmission and reception times gives the distance to the satellite on account of the constancy of the speed of light (neglecting atmospheric effects). This information allows the navigator to place themselves somewhere on the surface of a sphere centred on the satellite. The signal from a second satellite places the navigator on the surface of a second sphere. The intercept of two spheres is a circle (see figure 14.16). A third satellite gives another sphere, and its intersection with the circle gives two points. Only one of these will be near to the Earth's surface, thus giving a unique 3D position of the navigator. This has been achieved without measuring the directions to the satellites, just the signal arrival times.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

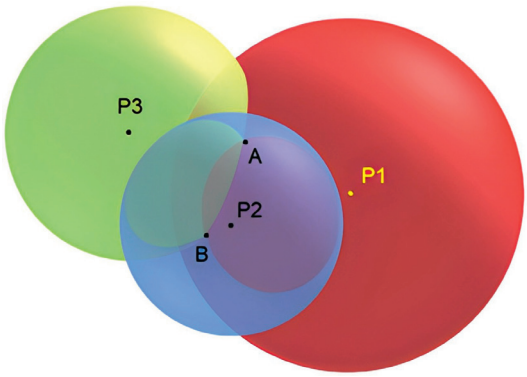


Fig. 14.16 The principle of global navigation satellite systems, here with three satellites ('trilateration'). The signal received from each satellite places the observer on the surface of a sphere with the satellite at the centre (centres marked as P1, P2, P3). The intersection of two spheres (if they intersect at all) is a circle. The intersection with a third sphere generally gives two points, shown here as A and B. Credit: Tomasz Ludyga, licensed under CC-BY-4.0, from Wikimedia Commons.

For the above to work as stated, the navigator would need to have a clock that is accurately synchronized to the satellite clocks. With a synchronization error of Δt, the positional error from one satellite is $c\,\Delta t$. Thus to achieve a positional accuracy of 3 m (typical for GNSS), the navigator's clock would have to be desynchronized by less than $\Delta t = 10\,\text{ns}$ over long durations. This is an unrealistic expectation for most clocks (see section 14.2.3) and also ignores relativistic time dilation. Therefore, a signal from a fourth satellite is used to solve for the time t at which the signals are simultaneously received at the navigator's position $\mathbf{x}$. If $\mathbf{x}_i$ is the position of the ith satellite when it sends its signal at time t_i, navigation involves solving the set of equations

$$|\mathbf{x} - \mathbf{x}_i| = c(t - t_i) \quad \text{for} \quad i = 1, 2, 3, 4 \tag{14.5}$$

for the unknowns $\mathbf{x}$ and t. As the receiver usually has a clock of low accuracy (e.g. a GPS wristwatch), then instead of solving for the absolute time, the receiver solves for an offset from the common satellite time (the different satellite clocks must be synchronized to one another to high accuracy). Furthermore, the signals from the satellites will not be received simultaneously, so the receiver has to measure the differences between the arrival times. A watch with an accuracy of 10 s per year accumulates the 10 ns error after just 30 ms, implying it needs to receive the different GPS signals within 30 ms in order to achieve the nominal 3 m position accuracy. Thus while the receiver's clock does need a high time resolution (better than Δt), and must be accurate over the duration of the observations, it does not need long-term accuracy.

To determine our position with GNSS on the Earth, we therefore need signals from at least four satellites, or three if we know our altitude. More satellites will give higher accuracy. One example of GNSS is the Galileo system, which comprises nominally 24 satellites orbiting in three different planes at an altitude of 23 000 km above the Earth's surface. This ensures that more than four satellites are visible from any point on the Earth at any time, to allow for possible satellite outages. The system can deliver a positional accuracy of around 20 cm.

To maintain the accuracy of the satellite's clocks, they must be regularly synchronized with a reference clock on the Earth's surface, and they must be sufficiently accurate to

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

maintain this until the next synchronization. If they are synchronized once per orbital period (just over half a day), then adopting 1 ns as the target accuracy, we see that the clock must be accurate to about one part in 10^{13}. This can only be achieved with atomic clocks.

Because the navigation satellites are moving relative to the reference clock (at around $3.5\,\mathrm{km\,s^{-1}}$), there is a significant special relativistic time dilation (section 7.2). This makes the clocks on the satellites appear to run slow by about $7\,\mu\mathrm{s}$ per day. There is an additional general relativistic time dilation arising from the fact that the satellites are moving through a lower gravitational potential than the reference clock on the Earth's surface. This correction leads to the satellite clock advancing by $45\,\mu\mathrm{s}$ per day. The net effect is that clocks at the GNSS altitude advance by about $38\,\mu\mathrm{s}$ per day.[3] These relativistic effects are large ($38\,\mu\mathrm{s} \times c = 11\,\mathrm{km}$) and orders of magnitude larger than the required timing accuracy, so must be corrected for when solving for the navigator's position. There are additional relativistic effects, such as the Sagnac effect arising from motion in a rotating coordinate system. These are smaller in magnitude, but must also be corrected for.

Deep space pulsar navigation is similar to GNSS in that both use beacons of known positions, transmitting unique identification signals. With GNSS, the satellites transmit their continuously changing 3D locations, whereas with pulsars we assume that the direction to the pulsar is fixed but the distance is unknown. (Once we get to interstellar space then we do need to know the distance to some degree of accuracy.) The pulsar does not transmit its position, but we can identify it from the shape of its pulse and so look up its direction in a catalogue. In both cases we use the timing of the received signal to determine our position accurately in the radial direction to the beacon, but without constraining the tangential position much, or at all. Multiple beacons lying in different direction then allow us to determine our 3D position.

Another major difference between GNSS and pulsar navigation is that GNSS broadcasts a universal timescale (e.g. Galileo System Time, GST) whereas pulsars of course do not. The pulses from pulsars are instead markers on a universal timescale (SSB time), but with a phase ambiguity that is solved to within some distance range by observing multiple pulsars.

Stellar navigation as outlined in section 14.4 is very different from both pulsar navigation and GNSS, because it observes the directions to the beacons by measuring the angular separations between them. It is therefore using the parallax displacement, which is not used in either GNSS or pulsar navigation. The variation on stellar navigation mentioned earlier using the stellar radial velocities, on the other hand, does not measure angular positions. It relies instead on the displacement of the stars to change their relative directions of travel, and thus observed radial velocities. So this too is a radial, rather than tangential measurement, and again multiple sources will reveal the 3D position.

[3] The net timing correction from these relativistic effects is a function of orbital radius. For satellites in low-Earth orbits (up to 2000 km) such as the International Space Station, the time dilation from the faster orbital speed dominates over the smaller gravitational effect, and the clocks lose time. For higher orbits the gravitational effects dominate, and so the orbiting clocks gain time.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

14.7 Spacecraft control

Knowledge of the position, velocity, and orientation of a spacecraft is required for two distinct tasks. First, it is needed to tag observations. For example, as the spacecraft crosses from the solar system into the ISM, mapping the change in plasma density, we need to know where the spacecraft is. Similarly, when arriving at the target system, the spacecraft will need to know where it is in order to point its instruments in the right directions.

The second use of navigation is to change course by changing the velocity of the spacecraft. Interplanetary spacecraft typically achieve this with rockets. Changes in spacecraft orientation are often achieved using cold gas thrusters (diametrically opposed so that they generate a torque but not a lateral acceleration), although for space observatories in particular this is also done with gyroscopes.

What options a spacecraft has for controlling its velocity or orientation depends very much on the overall propulsion design and mass budget. A spacecraft with a fusion drive (chapter 8) could be designed to reuse this during the mission to adjust its course. Even if that has been used up, or jettisoned, it is likely that the spacecraft has enough mass to host thrusters for small corrections. A massive laser sailcraft (chapter 11) may likewise have enough mass budget for microthrusters. Another option for steering might be the use of magnetic and electric sails (chapter 13), by changing the orientation of the loop or grid with respect to the direction of travel.

The situation is very different for ultralight (gramme-scale) sailcraft, which would not be able to accommodate any kind of matter rocket. One option here is to vary the reflectivity of different parts of the sail. When illuminated by the star in the target star system, this will induce a torque that can be used to change the orientation and thus direction of the travel of the sailcraft (as tested by IKAROS; see section 10.6.1). The same could be done using the Sun when the spacecraft is still near the beginning of its journey (after the laser has been turned off). The sail could also be equipped with semiconductor light-emitting diodes on the sail rim to reorient it via their emitted photon momentum.

14.8 Summary

Navigation addresses the need to determine the position and velocity of a spacecraft within a coordinate system. When an interstellar spacecraft is within a few tens of astronomical units of the Earth, it can use Earth-based beacons to help it navigate, just as interplanetary spacecraft do today. This is done with two-way ranging and the Doppler effect, in which signals are sent from the Earth to the spacecraft and back again. This can achieve an accuracy in the radial distance and velocity of 1 m and 1 mm s^{-1}, respectively (but worse for large distances). DOR (differential one-way ranging) and delta-DOR, which rely on two widely spaced ground stations receiving signals from the spacecraft, may be used to determine the angular coordinates of the spacecraft (figure 14.4). Accuracies of a few tenths of a milliarcsecond have been achieved.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015

Inertial techniques use gyroscopes, accelerometers, and clocks to monitor the change in velocity, and thus position, over short time periods. They provide a differential navigation between receiving external updates of the position and velocity of the spacecraft.

At larger distances from the Earth, autonomous methods of external navigation are required that do not rely on receiving signals from the Earth. Stellar astrometric navigation uses a catalogue of stars with accurately known 3D positions and 3D velocities relative to the Sun. A spacecraft at another position with another velocity will see these stars displaced by the effects of parallax and aberration (figure 14.6). By measuring the angular separation between multiple pairs of stars, it is possible to untangle these effects to determine the 3D position and 3D velocity of the spacecraft. By measuring the angular separations of 10 star pairs to an accuracy of 10 mas, we could achieve accuracies of 0.01 au and $10\,\mathrm{m\,s^{-1}}$ in position and velocity, respectively, even at interstellar distances and relativistic speeds.

Another autonomous method is the use of highly regular pulses from pulsars to determine the position of the spacecraft along the line of sight to the pulsar, albeit with a phase ambiguity (figure 14.13). Using multiple pulsars, the 3D position can be determined. With sufficiently accurate initial knowledge of the spacecraft position and velocity, a final accuracy of 10–100 km should be achievable. For navigation near to the Sun, only the directions to the pulsars need to be known accurately; beyond about 30 au, their approximate distances also need to be known.

14.9 Exercises

1. Convince yourself – with a sketch, and mathematically – that the statement in the footnote on page 255 about hyperbolic navigation is true.
2. The DOR technique in section 14.3 assumes that the spacecraft is infinitely far away when interpreting the time delay at the ground stations as a position angle (equation 14.1). Suppose instead that the spacecraft is at a large but finite distance qb from the midpoint of the ground stations, where b is the distance between the ground stations, and q is positive real number. Denote the angular position of the spacecraft from this midpoint as ϕ. What angular position would equation 14.1 predict? That is, find an expression for θ as a function of ϕ. Plot $\theta - \phi$ vs ϕ for a range of values of q for $b = \mathrm{R}_\oplus$. Below what distance is the difference between θ and ϕ significant, such that equation 14.1 is no longer a good approximation? (Note that for $q \gg 1$, $1/q$ is the parallax in radians of the spacecraft between the ground stations.)
3. Show that neglecting the special relativistic time dilation between a receiver on the Earth's surface and a GNSS satellite orbiting at 22 300 km above the Earth's surface introduces a timing error of around 7 μs per day.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 17 Nov 2026 at 05:31:56, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009700429.015