Richard A. Muller and Gordon J. MacDonald

Ice Ages and Astronomical Causes

Data, spectral analysis and mechanisms

Springer

Published in association with
Praxis Publishing
Chichester, UK

2

Astronomy

"Probably another reason why many Europeans consider the Chinese such barbarians is on account of the support they give to their astonomers, people regarded by our cultured Western mortals as completely useless. Yet there they rank with heads of Departments and Secretaries of State. What frightful barbarism."

Franz Kuhnert, 1888[1]

2.1 WHY ASTRONOMICAL FORCING?

In the 1840s, James Croll assumed that the ice ages were driven by changes in the orbit of the Earth. The astronomical origin was likewise taken as natural by Milutin Milankovitch, when he revived and improved the Croll theory in the early 1900s. But neither of these theories has proven successful in accounting for the dominant 100 kyr glacial cycle of the last million years, and so there have sprung up theories that ignore astronomy and attribute this cycle to processes confined to the Earth. These are sometimes called "free oscillation" theories, and they argue that the climate system of the Earth is undergoing a natural resonance, like the tone heard when you put your ear to a seashell. Yet in this book, even in the title, the astronomical origins are taken for granted. Why is this?

There are several persuasive reasons to think that astronomy is responsible. The first is the coincidence of astronomical frequencies with those found in the ice age data. In our discussion of the data, we will show that the strongest frequencies that show up often have periods 100, 41 and 23 kyr, and these are readily identified with astronomical periods. The 100 kyr period has historically been identified with the changes in the eccentricity of the Earth's orbit, although it also matches the period of another parameter called orbital inclination. The 41 kyr period is the dominant oscillation of the tilt of the Earth's poles towards the sun. And the 23 kyr period

[1] Quoted by F. R. Stephenson and D. H. Clark, *Sky & Telescope*, Feb. 1977, p. 86.

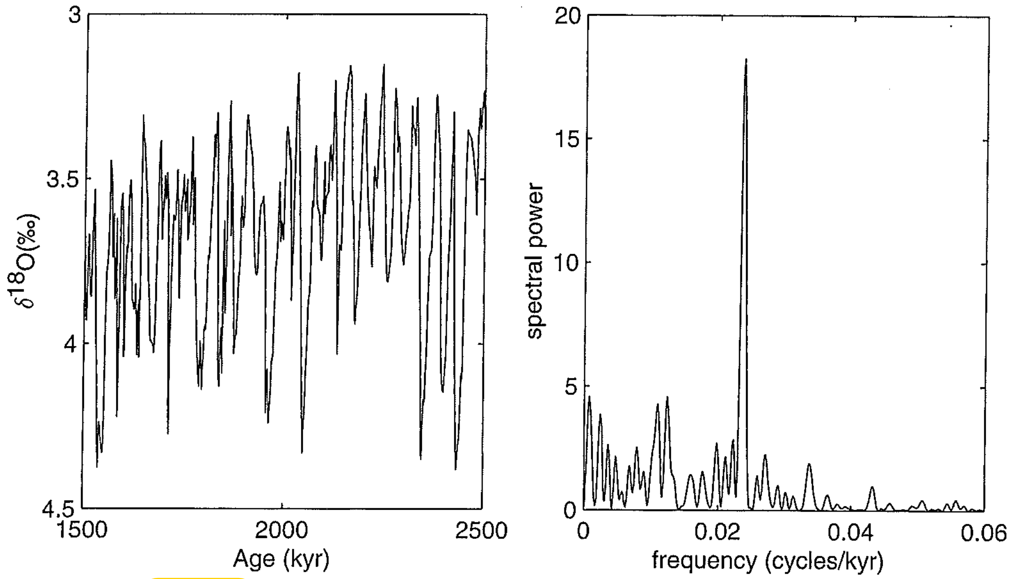

Fig. 2.1. Site 607 from 1.5 to 2.5 kyr: δ^{18}O data and spectrum.

is the dominant oscillation in the parameter that tells us how close the Earth is to the Sun in mid-summer.

A second, and independent, reason to think that the 100 and 41 kyr lines are driven by astronomy can be found in the nature of the oscillations. Over long periods (700 to 1000 kyr), these oscillations remain coherent, i.e. they keep a constant phase. This can be seen by doing a spectral analysis and looking at the widths of the spectral peaks: the peaks are exceedingly narrow. As an example, we show the variations in oxygen isotopes from a sea floor core taken in the North Atlantic in Fig. 2.1. The time scale for the data was derived by assuming that the sedimentation rate was constant. The data and the ages for the end points were taken from Ruddiman et al. (1989), but the ages for points in between were calculated by linear interpolation.

The variations in the signal are believed to represent variations in the global ice; we will discuss the reasons for this in Section 4.1. Accept this for now, and look at the rapid variations. They appear too regular to be random, and this conclusion is confirmed in the spectrum of the data, shown on the right side of the figure. The spectrum is dominated by a single narrow line that appears at the frequency $f = 0.024 \pm 0.001$ cycles/kyr, corresponding to period 42 ± 2 kyr.

The period happens to match the calculated period of the Earth's obliquity, i.e. the tilt of the North Pole towards the Sun, which varies between 22 degrees and 24.5 degrees, with a period of 41 kyr. This match is certainly no coincidence, since we expect that the tilt of the poles towards the Sun could play a major role in the continued growth of glaciation in the polar regions. But aside from that coincidence, note the narrow width of the peak. As we will show, it requires that the signal stay in phase with itself over the entire duration of the data set. If the sedimentation time scale at the end of the data set was inaccurate by as little as 20 kyr (a half cycle) then the single narrow peak would have been broadened and probably split into two adjacent peaks. Thus the sedimentation rate must have been remarkably constant during this period.

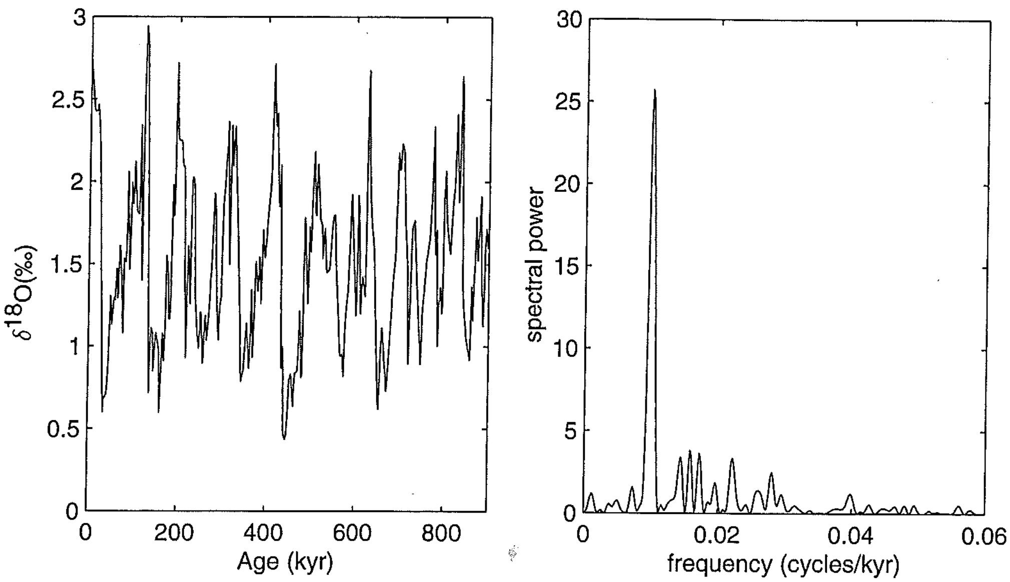

Fig. 2.2. Site 659 $\delta^{18}O$ and spectrum, 0 to 900 kyr.

We argued in 1997 (Muller and MacDonald, 1997a) that the narrowness of the peak implies that the glacial cycles are being driven by an astronomical force, regardless of what the details of the driving mechanism actually are. The reason for this is that natural processes in astronomy virtually always give rise to narrow spectral peaks, and natural processes in geology and climate do not. Narrow peaks are characteristic of processes that have low loss of energy. That means the lack of both friction and those mechanisms through which the oscillation is coupled to other processes that can drain energy away. The connection between low loss and narrow peaks is not well known to many scientists, so we will spend much of the rest of this section describing it in detail.

The argument can also be applied to the data from the last 900 kyr, which are dominated by a 100 kyr cycle. In Fig. 2.2 we show the oxygen isotope data and spectrum for Site 659 in the Atlantic off the coast of Africa. The data are plotted assuming a linear sedimentation rate, with the end point times set according to the time scale of Tiedemann et al. (1994).

Just as for the earlier data, the climate cycles appear to be very regular, but now with a period of 100 kyr. This is verified in the spectrum plot, which is dominated by a single narrow peak with $f = 0.01$ cycles/kyr. The width of this peak is the same that we would obtain with a pure sine wave 900 kyr long. Just as we argued previously that the 42 ± 2 kyr peak was astronomically caused, we argue that the 100 kyr peak must also be astronomically caused.

The subtle part of this argument is the statement that a natural climate oscillation would not show a similarly narrow peak. We will now explain the reasons in some detail. We will consider three cases: a *resonant* oscillator, a *relaxation* oscillator and, finally, a *driven* oscillator. The observed spectra are consistent only with the driven oscillator model. If this argument sounds arcane and uninteresting, then you are welcome to skip it. It is of interest primarily to those who postulate that the basic oscillations of climate are internal to the Earth, and not astronomically driven.

2.1.1 Resonant oscillator

Any oscillation of an object which maintains a single frequency is called, in physics, a *mode*. A simple pendulum has a single mode; two coupled pendula have two modes. The number of modes increases greatly if the internal vibrations of the objects are included. An object that has N molecules (where N is typically Avagadro's number, or greater), each one free to move in three dimensions, typically will have 3N modes. In quantum mechanics, only oscillations in pure modes have well-defined energy, equal to hf, where h is Planck's constant, and f is the frequency of the mode. Heat consists of the simultaneous excitation of all modes. In thermal equilibrium, for example, all modes are excited equally, and for classical systems all have average energy equal to $(1/2)kT$, where k is Boltzmann's constant, and T is the temperature. Friction in an oscillating system can be thought of as a process that takes energy from one mode and redistributes it to many others, turning it into heat. The presence of friction implies that an object will not continue to oscillate as a pure sine wave, but that the amplitude will decay with time. A spectral analysis of a decaying sine wave yields a spectrum with a broadened peak.

To discuss this mathematically, we will have to jump ahead and use some of the Fourier transform equations that will be fully discussed in the next chapter. Let the time behaviour of a decaying oscillation be given, for $t > 0$, by:

$$y(t) = \sin(2\pi f_0 t)e^{-\lambda t} = \frac{e^{2\pi i f_0 t} - e^{-2\pi i f_0 t}}{2i} e^{-\lambda t}$$

The Fourier transform of this function is obtained by multiplying it by $e^{2\pi i f t}$, and integrating (see Section 3.2):

$$\begin{aligned} Y(f) &= \int_0^{+\infty} \frac{e^{2\pi i f t} - e^{-2\pi i f t}}{2i} e^{-\lambda t} e^{2\pi i f t}\, dt \\ &= \frac{1}{2i} \int_0^{+\infty} (e^{2\pi i f t} - e^{-2\pi i f t}) e^{-\lambda t} e^{2\pi i f t}\, dt \\ &= \frac{1}{2i} \left[\frac{1}{2\pi i(f + f_0) - \lambda} + \frac{1}{2\pi i(f - f_0) - \lambda} \right] \end{aligned}$$

Note that we have taken the lower limit of the integral to be zero, to reflect the assumption that the oscillation begins at $t = 0$. The Fourier transform $Y(f)$ has a peak near $f = f_0$, the resonant frequency. The spectral power is given by the absolute square of the Fourier transform (i.e. multiply the Fourier transform by its complex conjugate). Near the peak at f_0, the square of the second term will dominate:

$$P(f) \approx \left(\frac{1}{4}\right) \frac{1}{4\pi^2(f - f_0)^2 + \lambda^2}$$

This functional form for the spectral power is often given the name "Lorenzian". It is the spectral shape usually seen in optical spectra as well as in radioactive decay.

The spectral power will drop to half its maximum amplitude when

$$\Delta f \equiv (f - f_0) = \frac{\lambda}{2\pi}$$

Thus, when a system is highly damped (i.e. λ is large) then it will have a broad peak in the frequency domain. We can call $\Delta t = 1/\lambda$ the characteristic duration of the signal. Then the equation is

$$\Delta f \, \Delta t = \frac{1}{2\pi}$$

This equation is a special case of the *uncertainty principle*, which we will discuss further in Section 3.2.3.

The key point here is that any oscillator that has a mechanism to lose energy with a time constant Δt, will have a broadened spectral peak. The exceptions are objects that do not have coupling between modes, and thus are essentially friction free. Such modes are often called "high Q" oscillations, since the quality factor Q is defined as the number of radians for damping to reduce the energy by a factor of e:

$$Q = \frac{\pi f_0}{\lambda} = \pi \frac{\Delta t}{T}$$

where T is the period of the oscillation.

In nature, high Q systems are found only in those realms in which friction is weak, so there is little decay (i.e. λ is small). The Earth does have high Q modes of oscillation. They are the seismic modes of vibration, which have relatively low loss in the interior of the Earth, and reflect efficiently off the surface and interior layers. But internal climate modes tend to be lossy and therefore low Q. To estimate the Q of oscillatory behaviour, try to imagine how many cycles it will take before the energy is transferred into other phenomena, and then multiply your guess by 2. Q can be estimated from the width of the spectrum from the equation:

$$Q = \frac{f_0}{2\Delta f} = \frac{f_0}{FWHM}$$

where $FWHM$ stands for the full width at half maximum of the resonance response. For the 41 kyr peak shown in Fig. 2.2, the $FWHM = 0.000875$, $f_0 = 1/41$, and we have $Q = 28$, assuming that it is a free oscillation. The challenge for any free oscillation model would be to explain how the lossy behaviour of the climate system is compatible with this high value of Q.

Two natural realms which have low friction, and show high Q behaviour are those of astronomy and quantum mechanics. The laws of physics were discovered in astronomy (by Newton) precisely because solar system motion occurs with very little friction.[2] Likewise, in atomic systems, electrons move without friction; states decay

[2] Galileo had to demonstrate that objects of different sizes fall with the same velocities (his famous Tower of Pisa experiment) by using two round massive objects; if one of them had been a feather then the demonstration would have failed.

only by relatively large, discrete quanta. It is important to note that a forced oscillator mimics the spectrum of the driving force. Thus, if the glacial cycles are driven by orbital oscillations, they too will have narrow spectral peaks. The peaks will be broadened only to the extent that the friction in the Earth causes a drag on the astronomical driving force.

We conclude that the high *Q* of the glacial oscillations system suggests that they are driven either by astronomical forcing or by quantum mechanics. Since we are aware of no mechanisms that couple quantum mechanical phenomena to climate, we conclude (by default) that astronomy is responsible.

2.1.2 Relaxation oscillator

Unlike the resonant oscillator of the previous section, a relaxation oscillator does not have a built-in frequency. Rather, it has a built-in time constant. In an electronic circuit, a relaxation oscillator works, typically, by charging a capacitor until a certain voltage threshold is reached that opens a switch and discharges the capacitor rapidly. The process then begins anew (see, for example, Horowitz and Hill, 1991.) In geology, an analogous process can take place. We might have a glacier that slowly increases in size but, when it reaches a certain level, spontaneously collapses—and then starts again. The switch that causes the collapse could be a related geophysical phenomena such as sea level. For example, we might suppose that an ice-age glacier would continue to grow until sea level has dropped to such a point that methane from clathrates was spontaneously released, causing a sudden warming that melts the glaciers completely. Sea level rises, the methane has a short lifetime in the atmosphere and disappears, and the glaciers begin to grow again. The collapse could not take place again until the clathrates recharge. (For a discussion of the possible role of clathrates on climate, see MacDonald, 1990.)

The ability of a relaxation oscillator to maintain a constant phase depends on the absence of background variability that can cause the oscillation to trigger prematurely or late. In an electronic circuit, this means that the electronic noise must be very low. In climate, it means that other variables, independent of the oscillation under question, must be constant. For example, if our relaxation oscillator has a natural time constant of 100 kyr, then any other climate signal (e.g. a 41 kyr oscillation, or a 23 kyr oscillation) can interfere with its proper triggering.

Many people would not consider that a defect, but rather a feature of the theory. The 100 kyr cycles of the ice ages do not terminate suddenly, but they seem to vary by ± 25 kyr. In fact, this is the basis for precession triggering models (discussed in Sections 6.4.4 and 6.4.8), in which the terminations are not quite free, but linked to precession cycles. Some occur after three precession cycles (i.e. 3×23 kyr $= 69$ kyr), and others after five such cycles (i.e. after 115 kyr).

The problem with this interpretation is that it requires some external force to keep the oscillations in phase. (In the precession triggering model, this is the insolation signal.) Unless there is such a timing mechanism, the phase of the oscillator tends to drift. If there is a variation of ± 25 kyr for each cycle, then after N cycles, the signal will drift out of phase by (on average) $25\sqrt{N}$ years. For $N = 9$, the approximate

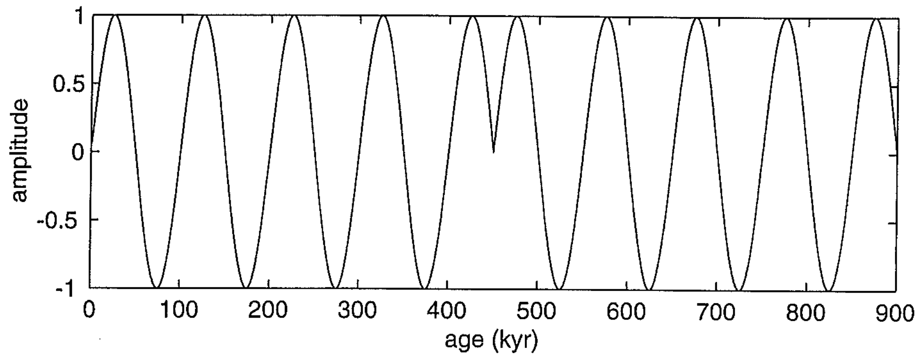

Fig. 2.3. Signal with phase shift.

number of 100 kyr oscillations observed (see Fig. 5.20), we would lose phase by 75 kyr. Yet this is not happening. (Fig. 5.20 also shows the remarkable phase stability.)

Such phase drift will also degrade the spectrum. A phase drift of 75 kyr is more than a half cycle. It means that the last half of the signal is out of phase with the first half, and that results in a broadening of the spectral peak. We illustrate this in Fig. 2.3, which shows a signal that shifts its phase by 180 degrees half-way through its 900 kyr duration. The spectral power for this signal is shown in Fig. 2.4, along with the spectral power of a pure sine wave (dashed line). Note that the width of the phase shifted line (full-width at half-maximum) is about three times that of the unshifted line. It is the presence of narrow spectral lines in the data, equivalent in width to that which we would obtain from a completely unshifted line, that leads us to conclude that the spectrum is not characteristic of a free relaxation oscillator.

We will give other examples of the broadening of the spectrum by phase instability in this book. In Fig. 3.4 we show the spectrum of a sawtooth signal that has a phase variation for each cycle of ±20 kyr. Note, however, that in this example the

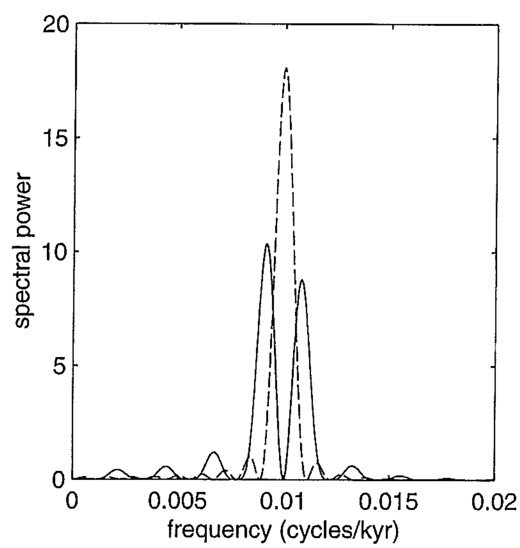

Fig. 2.4. Spectrum of phase shifted signal. The dashed line is the spectrum of a pure sine wave of the same 900 kyr duration.

phase error does not accumulate, but is equally probable on either side of a 100 kyr perfect cycle. In Section 5.6 we show the results on the spectrum of taking a narrow signal (sine wave, or orbital inclination) and putting in an artificial phase shift. In Section 5.9 we show the consequences on the spectrum of introducing a phase shift as part of the process of time-scale adjustment called *tuning*.

A detailed example of the relaxation oscillation model for the 100 kyr cycle can be found in Saltzman and Verbitsky (Saltzman and Verbitsky, 1993). This team was courageous enough to publish a detailed prediction of the signal that would be found in the Vostok CO_2 record *before* the record was measured (Saltzman and Verbitsky, 1995). Unfortunately for their model, their prediction of a maximum at age 250 kyr turned out to be completely wrong; the CO_2 at that age turned out to be a minimum. Nevertheless, they set an example for everybody else by creating a theory that made predictions, a standard that has been matched by very few other theories.

2.1.3 Driven oscillator

In contrast to the free oscillators described in the previous sections, the driven oscillator is a system that has a natural frequency f_0, but which is being driven by a force that has a difference frequency f. Tides are an example of driven oscillations. A bay may have a natural resonant frequency, the frequency of sloshing, but it is driven by the tidal gravitational force of the Moon and the Sun, which has a period of 12 hours. The key feature of driven oscillators is that they operate at the driving frequency, not at the natural frequency. Regardless of the natural period of the bay, the tides maintain a frequency of two cycles per day.

On the other hand, the amplitude and the phase of the driven oscillator depend critically on the match between the natural frequency f_0 and the driving frequency f. It is easily shown (see, for example, Marion and Thornton, 1988, p. 126), that the amplitude of the resulting motion is

$$h(t) = \frac{A\cos(2\pi ft - \delta)}{\sqrt{(f_0^2 - f^2)^2 + 4f^2\beta^2}}$$

In this equation, δ is a phase shift, and β is a damping parameter that describes the natural decay of a free oscillation. β is related to the Q of the resonance by the equation

$$Q = \frac{\pi f_0}{\beta}$$

The critical thing to note here is that $h(t)$ oscillates at frequency f, not at the natural frequency f_0. If f is close to f_0, then the denominator becomes small, and the amplitude of $h(t)$ becomes very large—but the frequency doesn't change.

If there are two frequencies in the driving force, then the situation becomes more complicated. Both frequencies appear in the response $h(t)$, but with different amplitudes. The result is that the shape of $h(t)$ is not the same as the shape of the driving force, although the frequency content is unchanged. This means the spectrum of the data will show peaks at the same locations as the spectrum of the driving force, but with different amplitudes.

An additional consequence of this fact is that if the spectrum of the driving force has a single, narrow spectral feature, then that same feature will show up in the response function $h(t)$. If there are several narrow peaks (as there will be for some of the orbital parameters we will discuss, including eccentricity and precession) then the response will also show several narrow peaks, although their amplitudes may have a different ratio.

The driving forces from astronomy have extremely narrow spectral peaks, because of the lack of damping in the motion of the planets. If these forces are responsible for glacial cycles, then we expect the spectrum of glacial cycles to show narrow spectral peaks. This is what is observed.

It is also possible, perhaps even likely, that the response of the climate to the driving force is not resonant at all, but is simply linear in the driving force. This is mathematically identical to the response that you would have if the resonant frequency f_0 is very far removed from f, and so the conclusions are the same. But with such a linear response, not only are the frequencies preserved, but so are the relative amplitudes.

If we need to enhance one frequency and suppress others, then we can do so with a resonant response. For example, we will show that the orbital parameter known as eccentricity has three frequencies inherent in it, corresponding to periods of 95, 125 and 400 kyr. However, the glacial cycles have one only frequency present, consistent with the period 95 kyr. Could eccentricity still be driving the behaviour? Yes, if a natural resonance in the climate response function is close to 95 kyr, and has very small damping (represented by β in the previous equation). This requires the resonance to have a very high Q so that the 125 is suppressed relative to the 95. If there were such a narrow resonance in the climate response, we would hardly need an astronomical driving force, since the natural oscillations of such a resonance would account for the narrow peak. The challenge in such theories is to find a physically plausible mechanism that would have such a resonance. As explained earlier, the free relaxation oscillator will not work, since it is too vulnerable to background noise.

When a driving force happens to match the resonant frequency of the response, i.e. when $f = f_0$, then there is another interesting consequence. The phase shift is given by the equation (Marion and Thornton, 1988, p. 125):

$$\delta = \tan^{-1}\left(\frac{2\omega\beta}{\omega_0^2 - \omega^2}\right) = \tan^{-1}\left(\frac{2f\beta}{\pi(f_0^2 - f^2)}\right)$$

where $\omega = 2\pi f$. Near resonance, the argument goes to infinity, and δ is $\pi/2$, i.e. there is a 1/4 cycle delay. At frequencies far below resonance, the phase delay δ is approximately zero, and at frequencies far above resonance, the phase delay is π, i.e. a half cycle. So, for example, if the 95 kyr peak is enhanced by a resonance at this same frequency, then we expect a 95/4 = 24 kyr lag between the eccentricity cycle and the glacial cycle. The presence of this phase shift has not been widely appreciated by those who have invoked resonance models.

We conclude this section with our conclusion: that the narrow spectral features seen in glacial cycles can be understood in terms of a driven oscillator, and only astronomy provides plausible driving forces with sufficiently narrow spectral features to account for the data.

2.2 ORBITAL PARAMETERS

It may surprise some people that we *can* calculate solar system orbits back millions of years. But remember that physics was discovered when Newton recognised that the laws of gravity that affect behaviour on Earth were identical to the laws that determined the celestial motions. Newton's laws work extremely well in space, because of the relative absence of friction.

The field known as celestial mechanics, in which these calculations are done, was once one of the forefronts of scientific research. Perhaps surprisingly, it has now returned to that position. The original interest grew out of astrology. Those who believed that the locations of the stars, and in particular the planets, influenced life on Earth, saw great potential power in the possibility of predicting where those planets would be in the future. (We suspect that Franz Kuhnert, whose quote was used at the beginning of this chapter, confused astronomy with astrology.) Recently, interest in celestial mechanics has revived because of the space programme. One must calculate the orbits of spacecraft to the exquisite precision to send a mission to Mars. And interest in celestial mechanics has revived in recent years because of its value in testing General Relativity theory, and in accounting for some of the surprising discoveries made by the planetary probes. Celestial mechanics, like spectral analysis, has developed into an extremely arcane field. It is subtle and difficult if you want to calculate orbital parameters to sufficient precision to send a probe to pass a satellite of Jupiter, or to find the orbit of the Earth ten million years ago.

The basic calculations are not difficult, and can be learned in a single undergraduate course. However, just as in spectral analysis, there are many traps awaiting the incautious. In 1999, a probe to Mars failed because some calculations were mistakenly done using feet rather than metres. And it was discovered in 1991 that the published calculations of orbits older than 1.5 million years were incorrect, because of accumulating roundoff errors in the computer programs used (Quinn et al., 1991).

2.2.1 Methods of calculation

The earliest detailed calculations of changes in the Earth's orbit, those carried out by d'Alembert, Leverrier and Pilgrim, were done using a method known as perturbation theory. In this approach, the orbits of the planets were assumed to be ellipses, and the mass of the planet was distributed along the orbit of the ellipse. This approximation was based on the insight that the average effect of the planet was well described by the average mass distribution along the orbit. This insight is still extremely useful for thinking about the perturbations. The mathematics of this perturbation approach is described in detail in the text by Danby (1992).

To calculate the changes in the orbital parameters, the parameters are first expanded in a trigonometric series. For example, the eccentricity e is written

$$e = e_0 + \sum_i E_i \cos(\lambda_i t + \Phi_i)$$

and then the coefficients e_0, λ_i, and Φ_i are calculated from the equations of motion (i.e. from $F = ma$) using perturbation theory. The number of terms used in the sum depends on the precision required. In the 1970s, the most complete tables of coefficients, with up to 47 terms, were published by Berger (1978). Insolation for any date and latitude could then be easily calculated from these coefficients; for the equations see, for example, MacDonald (1990).

The issue became more complicated with the recognition that the orbits can be chaotic in character, so that a solution in terms of periodic functions misses important behaviour. This, coupled with the development of special purpose computers, led to the extensive use of "direct integration" which avoided the perturbation approach by working directly with the equations of motion. In the direct integration method, the positions and locations of the planets are put into the program. The force on each planet is then calculated from the gravitational force equation, and the acceleration is calculated from $F = ma$. A small step in time is taken (Quinn et al. (1991) used step sizes of 0.75 days to cover 3 million years) and the new location and velocity of each planet is calculated. The process is continued. Special mathematical methods are used in the calculations to avoid mathematical artifacts. These solutions are not readily described by a few equations, but the internet has come to our rescue, by allowing the detailed numeric values to be posted for all to use.

The use of direct integration led to an interesting discovery by Quinn et al. (1991), who uncovered an error in the previous perturbation calculations. They showed that the previous perturbation methods of Berger and colleagues (1978) had correctly described the main features of the data for the past 1.5 million years, but that they were quite inaccurate for older ages. Quinn et al. showed that error could be traced to roundoff error in the prior computer calculations. The ironic result is that, although Berger et al. had recommended that all of their expansion coefficients be used for accurate calculations, only the first few were accurate. Quinn et al. state that "typical [roundoff] error in the obliquity in the last 1 Myr is at least 0.001 radians, which is larger than all but 6 of the 47 terms in Berger's series". Berger and Loutre corrected the error, and their new results (Berger and Loutre, 1992) are now in agreement with those of Quinn et al. Detailed calculations going back 20 million years have now been published by Laskar et al. (1993), and 100 million years by Sussman and Wisdom (1992).

For the purposes of this book, it is not necessary to understand these calculations in detail. However, it is important to know what the relevant orbital parameters are, and it is interesting to identify the phenomena responsible for the changes in them.

The orbital parameters that are central in understanding climate are called eccentricity, obliquity, the precessional parameter (or just 'precession' for short), and the orbital inclination. The first three of these are critical in describing the insolation; the inclination is important for other processes, including extraterrestrial accretion. We will now describe each of these in turn.

2.2.2 Eccentricity

The present orbit of the Earth around the Sun is shown in Fig. 2.5, to scale. The small dot in the centre is the Sun, also drawn to scale. The Earth is too small to be seen under the line representing its orbit. The orbit appears nearly circular, but it is not. Note that the orbit is not centred on the Sun, but comes closer on the right side, and further away on the left, as can be seen by comparing it to the grid.

Eccentricity measures the departure of the Earth's orbit from that of a perfect circle. If a is the major axis of the orbit (the horizontal diameter in Fig. 2.5), and b is the minor axis (the vertical width), then the eccentricity e is

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2}$$

The present eccentricity of the Earth's orbit is 0.016722. From the equation for e, you can calculate from this that $(a/b) = 1.00014$. Thus the orbit is circular to within 0.014%. That's why it looks like a circle; the deviation is less than the width of the line. (It is a true circle only if $e = 0$.) However (and this is the important part) the

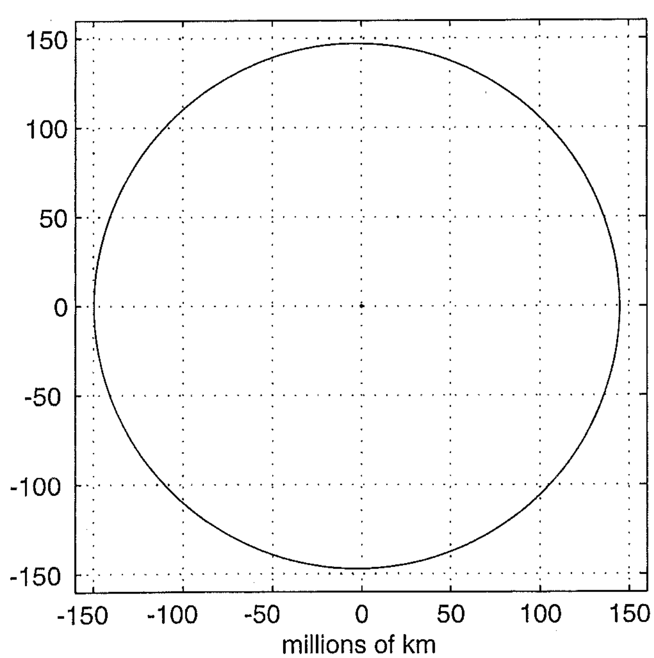

Fig. 2.5. Earth's orbit, to scale.

centre of the orbit is displaced from the location of the Sun. The distance of closest approach of the Earth to the Sun is $r_{MIN} = a(1 - e)/2$, the maximum distance is $r_{MAX} = a(1 + e)/2$, and so the fractional difference as the Earth goes around the Sun is $(r_{MAX} - r_{MIN})/r = 2e = 3.3\%$. This small variation, which can be seen in Fig. 2.5 (although just barely), is at the heart of every insolation model.

The eccentricity is not constant, but changes slowly with time. To see why this is true, it is helpful to recognise that the eccentricity is related to the angular momentum L of the Earth in its orbit. We write Newton's law of gravity in the form

$$F = -\frac{GMm}{r^2}$$

where M is the mass of the Sun, m is the mass of the Earth, and r is the distance between them. The energy of the orbit is the sum of kinetic and potential energies:

$$E = \frac{1}{2}mv^2 - \frac{GMm}{r}$$

To simplify the equations we are about to write, we define the constant k by

$$k = \frac{-2E}{G^2M^2m^3}$$

For orbits in which the planet is bound to the Sun (closed orbits), the energy E is negative, so the constant k will be positive. (The zero energy orbit, for which the planet can escape to infinity, is parabolic). Note that k is a constant only to the extent that the energy of the orbit is unaffected by perturbations. But for perturbations of the kind that occur in our solar system, this is a very good approximation. (See, for example, Danby, 1992.) With k approximately constant, it can be shown that the eccentricity is related to the orbital angular momentum L by

$$e^2 = 1 - kL^2$$

L=\sum_i r_i x p_i

(For a derivation, see, for example, Landau and Lifshitz, 1976, p. 36. Note: in the reference, M represents angular momentum, not mass.) This is the equation that relates eccentricity to orbital angular momentum. It is the key for understanding why and how perturbations of other planets affect the eccentricity of the Earth's orbit. The equation says, for example, that in extreme eccentricity ($e = 1$, the limiting value for which the ellipse becomes a parabola), the angular momentum is zero. This makes sense; this is the value you get when the object is falling directly into the Sun, with no transverse velocity. Eccentricity reaches its minimum value of zero when L reaches its maximum value of $1/\sqrt{k}$. You can easily verify that this is for a circular orbit, i.e. an orbit for which the centrifugal force equals the centripetal acceleration:

$$\frac{mv^2}{r} = \frac{GMm}{r^2}$$

Any force which removes angular momentum from a circular orbit makes that orbit more eccentric. Any force which increases angular momentum makes the orbit more circular. The rate of change of angular momentum is related to the torque τ on the Earth–Sun system by:

$$\frac{dL}{dt} = \tau$$

Torque on the Earth–Sun system is produced by any planet that pulls on the two asymmetrically. Major causes of such torque are the planet Jupiter, because it is massive, and Venus, because it comes so close. So the torques from these planets cause the eccentricity of the Earth's orbit to change. The fact that a few planets dominate the torque leads to the relatively simple variation observed in eccentricity. A plot of the Earth's eccentricity for the last 800 thousand years is shown in Fig. 2.6.

Over this time interval, the curve is approximately fit by the sum of a constant and three sine waves:

$$e(t) = a_0 + \sum_{k=1}^{3} a_k \sin(2\pi f_k t + \theta_k)$$

The three frequencies are 0.0105, 0.0080 and 0.0025 cycles/kyr, corresponding to periods of 95, 125 and 400 kyr. The presence or absence of these frequencies in paleoclimate has been a central issue in the analysis of the data, so it is worth looking at them closely. In Fig. 2.7 we plot the sum of these waves, with amplitudes $a_0 = 0.025$, $a_k = (0.0083, -0.0069, -0.0118)$, and phases $\theta_k = (0.0561, 0.1162, 1.8108)$ radians. These three frequencies account for 90% of the variance of the eccentricity signal.

It is worthwhile to examine this plot when the 400 kyr component is removed, leaving only the 95 and 125 kyr components, as shown in Fig. 2.8. What we see here is the beating of the 95 kyr cycle with the 125 kyr cycle. They nearly cancel near 0 years, 400 kyr and 800 kyr. The near cancellation near 400 kyr occurs near a climate period known as Stage-11 when the climate modulation was large. We will discuss

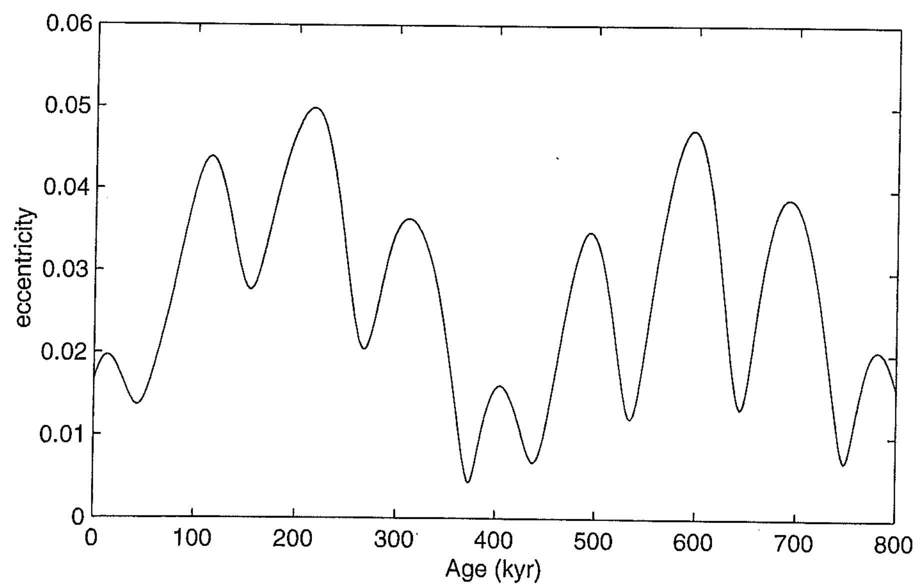

Fig. 2.6. The eccentricity of the Earth's orbit.

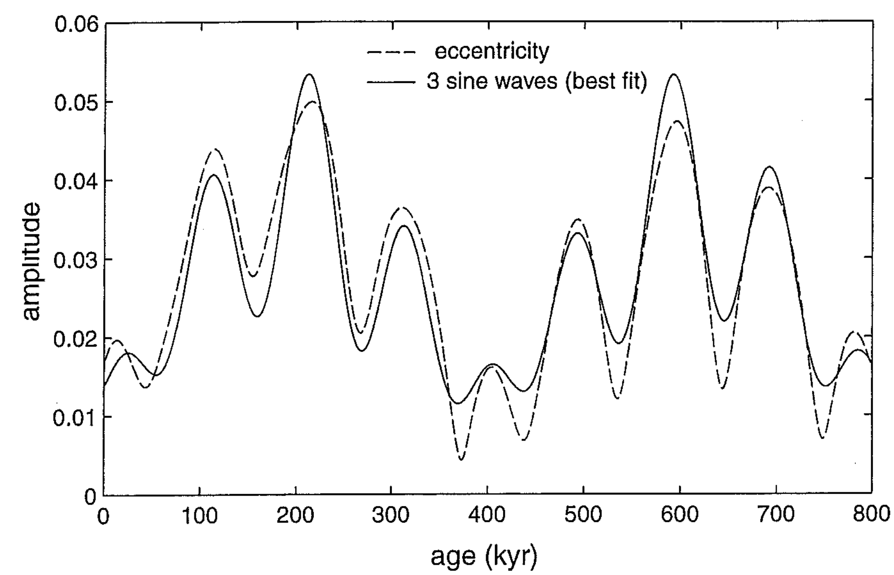

Fig. 2.7. Sum of sine waves with periods 95, 125 and 400 kyr.

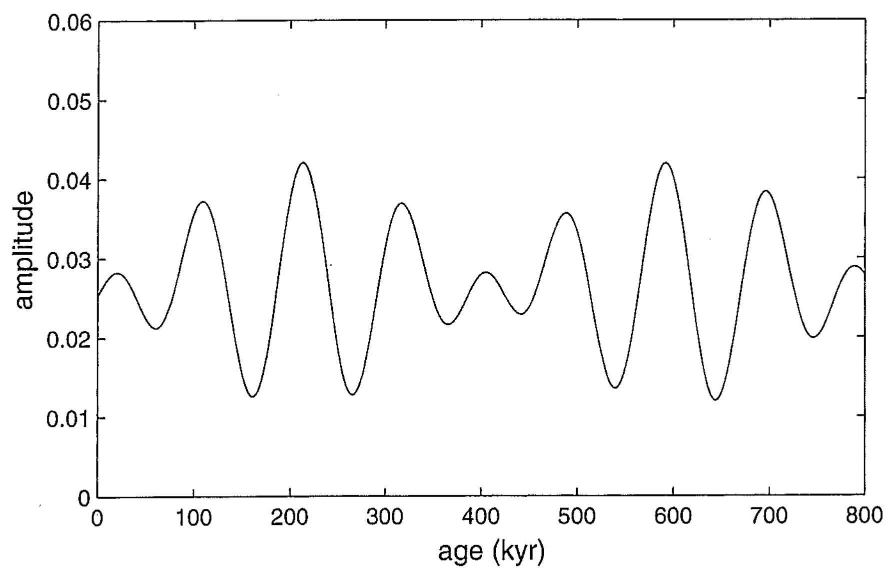

Fig. 2.8. Sum of sine waves with periods 95 and 125 kyr.

this at length in Section 8.2. The fact that the eccentricity modulation is small at this time has been considered a difficulty for the standard Milankovitch theory, and was termed by Imbrie et al. (1993a) as *The Stage-11 Problem*. Comparing Fig. 2.8, which has two frequencies, with Fig. 2.7, which has all three frequencies present, we see that the 400 kyr component takes this beat signal and raises the average value near 200 and 600 ky, lowering it near 0, 400 and 800.

The spectrum of eccentricity for two time intervals is shown in Fig. 2.9. Note that the ratio of the peaks varies with time. Because of such variations, comparisons with data should always use spectra derived from matched time intervals.

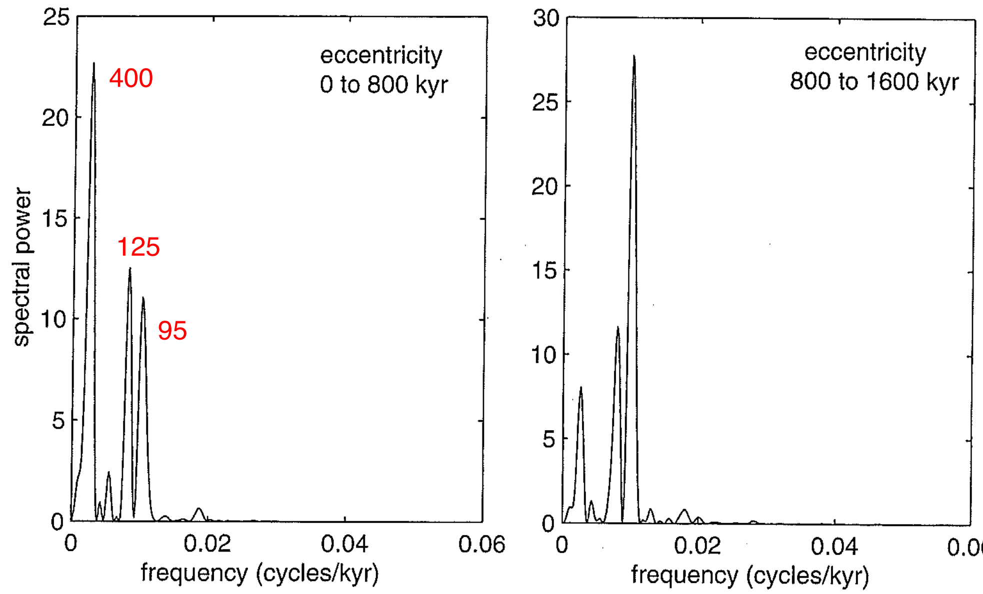

Fig. 2.9. Spectrum of eccentricity.

The three peaks correspond to periods of approximately 400, 125 and 95 kyr. These three frequencies are not really independent. They obey the relationship:

$$\frac{1}{95} - \frac{1}{125} \approx \frac{1}{400}$$

This relationship does not reflect a nonlinear process in climate. The eccentricity signal is purely astrophysical, and the relation derives from the nonlinear generation of eccentricity from the perturbations of the other planets on the orbit of the earth.

2.2.3 Precession

Precession is the slow change of the direction of the North Pole, as the Earth wobbles under the torque that the Moon and Sun exert on its equatorial bulge. The effect is completely analogous to the precession of a tilted top, whose axis rotates from the torque of gravity. The "ellipticity" of the Earth is 0.0034. (Ellipticity is mathematically similar to eccentricity, but it describes the shape of the Earth's surface rather than the shape of its orbit.) It is defined as the difference in the radius of the Earth at the equator vs. at the poles, divided by the average radius. Thus the equator is further from the centre by about $0.0034 \times 6371\,\text{km} = 21\,\text{km}$.[3] Because of this ellipticity, the gravitational pull of the Moon and the Sun exerts a torque on the Earth. The subsequent behaviour of the Earth is similar to that of a spinning gyroscope, which precesses because of the torque of gravity. (The torque arises in the gyroscope

[3] This makes Mt. Huascaran in Peru, which is only 11 degrees from the equator, further from the centre of the Earth than Mt. Everest, which is 28 degrees north, a fact which leads some people to argue that it is "higher". But it is not higher if the local reference of sea level is taken as the standard.

when the centre of mass is not directly above the point of support.) The net effect, as we discussed in Section 1.1, is that the direction of the Earth's axis precesses with a period of 25.8 kyr. This is the period measured with respect to the stars. In 12.9 kyr (a half period) the North Pole will be half way around the circle and will point towards Vega, 47 degrees away from the present Pole Star. In 25.8 kyr, the Earth's axis will again point towards the present Pole Star.

Precession affects climate because, when it changes, it means that the Earth's closest approach to the Sun, perihelion, occurs in different seasons. At present, perihelion occurs on 4 January, close to the Winter Solstice. In the Milankovitch theory, this is related to the fact that we are now in the midst of an interglacial. We are between ice ages because the Earth is too close to the Sun during Northern winter, and the glaciers cannot endure.

The date of perihelion also changes, however, and calculations must take this into effect. If we measure the precession of the Earth's poles with respect to this moving perihelion, the period is shorter, not 25.8 kyr, but closer to 23 kyr. Actually, because the perihelion does not move smoothly but lurches, the perihelion in this moving system is a superposition of several periods, the most important being 24, 22 and 19 kyr. The lurches in the perihelion motion are due to changes in the eccentricity of the Earth's orbit. The perihelion moves most rapidly when the eccentricity is greatest. The triplet of frequencies present in the precession parameter reflects the triplet of frequencies present in the eccentricity.

The effect of precession on climate also depends on how close the Earth comes to the Sun, and this depends on the eccentricity of the Earth's orbit. Thus the key parameter for insolation calculations turns out to be the product of the sine of the precession angle (measured with respect to the moving perihelion), multiplied by eccentricity. This combination is called the "precession parameter", often referred to as simply "precession". We give it the symbol p. We write the precession parameter p as

$$p = e \sin \omega_M$$

where e is the eccentricity, and ω_M is usually taken (by convention) as the angle between the *spring* solstice and perihelion. But it is necessary to recognise that it is not the same thing as the simple precession of the poles that we discussed in the introduction, which has a very regular period of 25.8 kyr. Because of the moving perihelion, as we described above, the period of the precession parameter is a combination of 24, 22 and 19 kyr. The amplitude of p is modulated by eccentricity, but it is not this modulation that gives rise to the triplet of peaks. For short time intervals, the peaks may not be resolved (see Section 3.15), and the precession may appear simply as a single peak near period 23 kyr.

Let's illustrate p by calculating its present value. Spring solstice typically occurs (these days) on 20 April, and perihelion occurs on 4 January. There are 106 days between these dates. This makes $\omega_M = 105$ degrees[4] = 4.5 radians. Using the current value of $e = 0.0167$, this gives $p = 0.0163$.

[4] Since there are 365 days per year, and 360 degrees, there is approximately one degree per day.

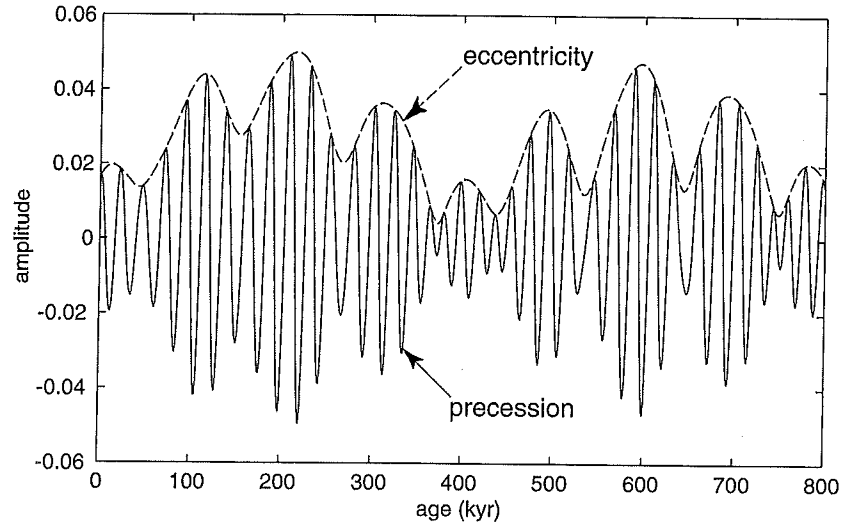

Fig. 2.10. Precession parameter p and eccentricity.

A plot of the precession parameter p for the last 800 kyr is shown in Fig. 2.10. Eccentricity is also plotted. From this diagram it is clear why we sometimes say that the *envelope* of the precession parameter is eccentricity. We can also say that the precession angle is *modulated* by eccentricity. This fact will prove very important, especially in nonlinear climate models. It is this modulation that must provide the 100 kyr oscillation, according to modern versions of the Milankovitch theory.

The spectrum of the precessional parameter is shown in Fig. 2.11 for two different time intervals, 0 to 800 kyr, and 800 kyr to 1600 kyr. The spectrum is dominated by three frequencies, corresponding to periods of 24, 22 and 19 kyr. The relative amplitudes of the three peaks vary with time. To compare a spectrum of a paleoclimate proxy with that of precession, the time interval for the two sets must be chosen to be identical.

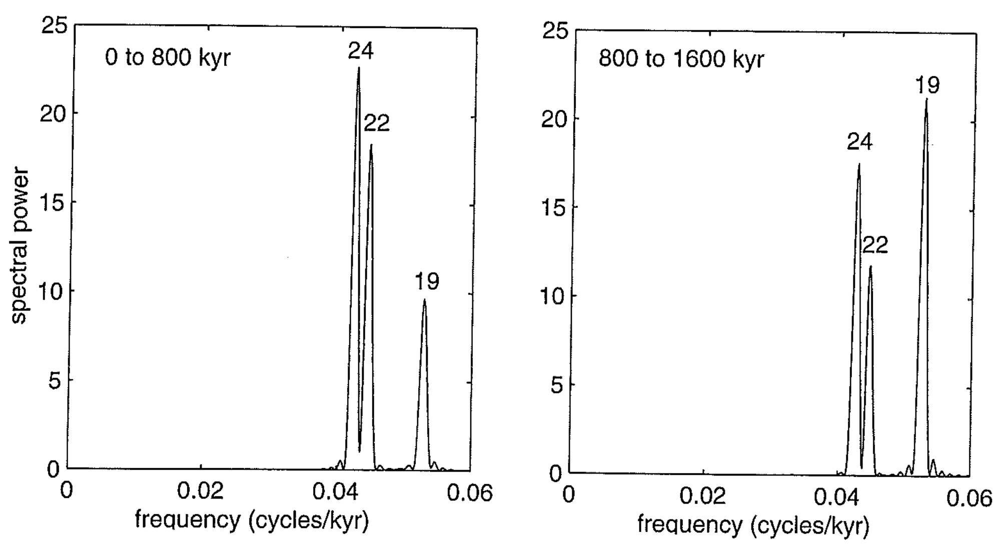

Fig. 2.11. Spectrum of precession.

2.2.4 Obliquity

Obliquity is defined as the angle of the tilt of the Earth's pole towards the Sun—i.e. how much the North Pole is tilted towards the sun in summer. If obliquity were zero, then the Sun would pass overhead at the equator every day. At the poles, half of the Sun would be visible above the horizon at all times, and the Earth would not have strong seasons. (There would be weak seasons from the fact that the distance between the Earth and the Sun varies by 3% throughout the year.) The present obliquity of the earth is 23.5 degrees. This is also the latitude of the Tropic of Cancer—the furthest north that you can go and still see the Sun directly overhead.

But obliquity is not constant: because the plane of the orbit of the Earth is constantly changing, the obliquity must be measured with respect to that plane. The Earth's polar axis precesses with a period of $1/f_P = 26$ kyr. The plane of the Earth's orbit precesses with a period $1/f_\Omega = 70$ kyr. (This variation is called Ω, and we describe it in more detail in Section 2.2.6.) The obliquity measures the differences between these two planes, and so it varies with a frequency equal to the difference of the two precession frequencies: $f_O = f_P - f_\Omega = 0.024$, corresponding to a period of $1/0.024 = 41$ kyr. Since both precession periods are quite regular, the obliquity variation is also very regular, and the 41 kyr period dominates. The current rate of change is 0.128 degrees per thousand years. That may not seem like much, but the Tropic of Cancer obelisk in Chiayi, Taiwan, erected about a century ago, is now 1.4 km distant from the correct latitude, because of the change in obliquity since it was constructed.

The obliquity for the past 800 kyr is shown in Fig. 2.12, and its spectrum is in Fig. 2.13. The variation consists of a strong 41 kyr cycle, with very small side peaks at periods 29 and 53 kyr.

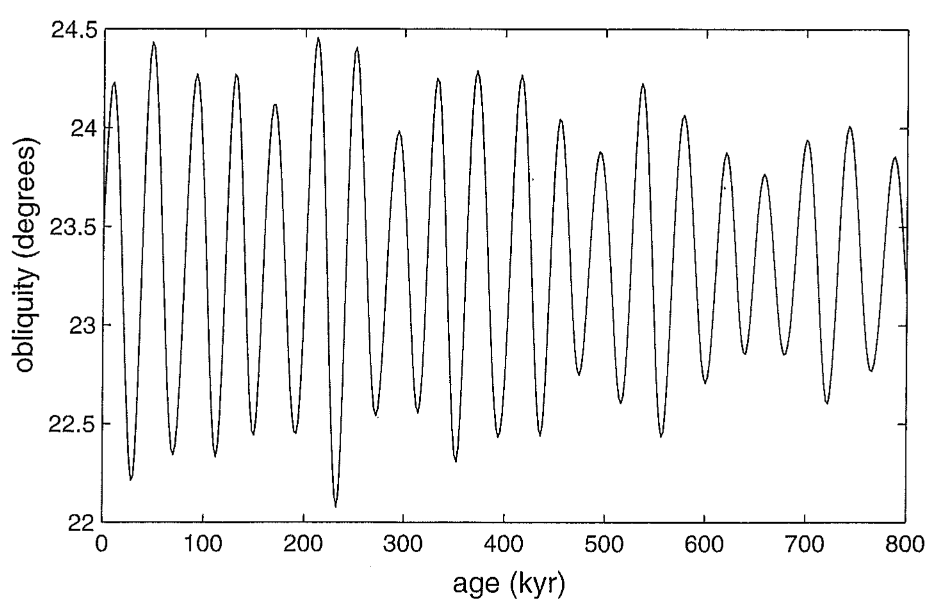

Fig. 2.12. Obliquity.

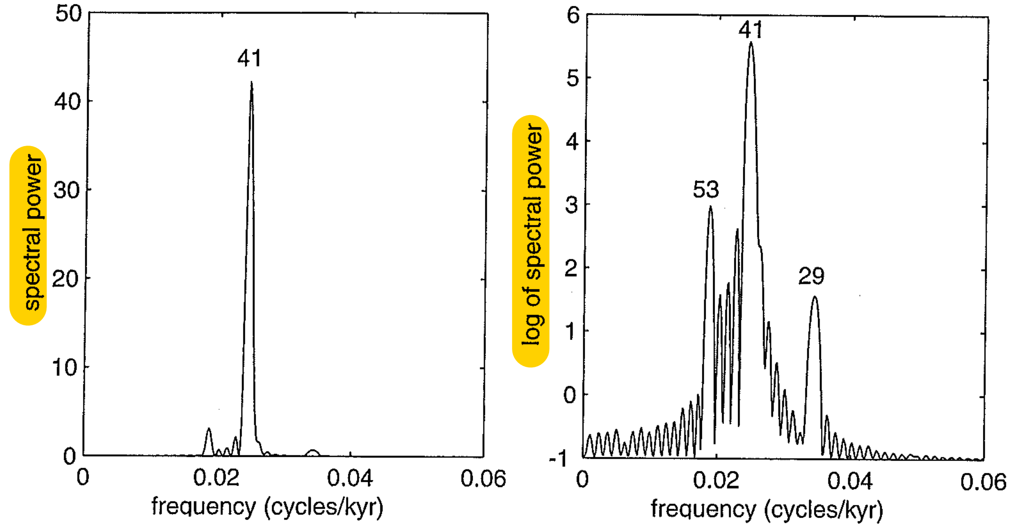

Fig. 2.13. Spectrum of obliquity, linear and log plots.

In the log plot (on the right), the additional periods at 53 and 29 kyr are most easily seen. However, they are small compared to the main peak at 41. The small rapid oscillations in the log plot are an artifact; they would be removed if we applied a taper (discussed in Section 3.4).

2.2.5 Insolation

The average flux of solar energy at the top of the Earth's atmosphere is approximately $S = 1360$ Watts/m^2, but the energy per unit area, the insolation, depends on the tilt of that surface with respect to the incident radiation. The average daily insolation over the Earth is 1/4 of this value, $I = 340\,\mathrm{W/m^2}$. The easiest way to see this is by recognising that the total insolation hitting the Earth is $\pi R^2 S$, where R is the radius of the Earth, but that this radiation is spread out over an area of $4\pi R^2$.

It has become standard in papers related to climate, past or present, to talk about the "equivalent radiative forcing" of various phenomena, i.e. to compare them to the insolation forcing. We show this in Table 2.1, based on a similar table of Svensmark.

Some older papers give insolation in units of calories/(cm^2 day). This unit has even been given a special name, the "Langley". We are thankful that this unit has

Table 2.1. Radiative forcing.

S	(solar constant at Earth's location)	1360 W/m^2
S/4	average insolation at top of atmosphere	340 W/m^2
A	Earth's albedo, the average reflected insolation	−53.5 W/m^2
CO_2	estimated effect from increase 1750 to present	1.5 W/m^2
CO_2	estimated effect if atmospheric CO_2 doubles	4 W/m^2
Clouds	net effect	−28 W/m^2

fallen into disuse (no disrespect to Langley intended). To convert to W/m^2, multiply Langleys by 0.4843.

According to the insolation theory of Milankovitch, the key driving force of the glacial cycles is the summer insolation at northern latitudes. For simplicity, most analysts pick a date and latitude, and calculate the average daily insolation at that place on that day. (The insolation is independent of longitude.) The key orbital parameters that are needed in the calculation are the eccentricity of the Earth's orbit (since the distance to the Sun is important), the obliquity at that time (since this tells how much the northern pole was pointing towards the Sun), and the precessional parameter, which depends on where in the elliptical orbit the Earth was on the date of interest. For the detailed equations, see MacDonald (1990) or Laskar et al. (1993).

The insolation for 21 July at 65N is shown in Fig. 4.38. The spectra of insolation for that date at various latitudes are shown in Fig. 2.14. There are several interesting features to note in these plots:

1. The main frequencies present are those of obliquity (0.024 cycles/kyr, period 41 kyr) and the triplet of precession (0.042, 0.045 and 0.053, corresponding to periods 24, 22 and 19 kyr).
2. Eccentricity is virtually absent. (Its frequencies are 0.0025, 0.008 and 0.0105 cycles/kyr, corresponding to periods of 400, 125 and 95.)
3. At northern latitudes the precession signal is strongest.
4. At far southern latitudes, obliquity dominates. (Note that in July, 65S is experiencing winter, not summer.)

In the insolation plots shown, only one is a good match to real $\delta^{18}O$ sea floor data. That is the last one, July insolation, 65S, which matches the spectrum of the data seen during the period between 1 and 3 million years ago, when it was dominated by obliquity (see Fig. 2.1). A similar spectrum is obtained from the January insolation at 65N, shown in Fig. 2.15.

These results suggest that it is not summer insolation that determined the cycle of the ice ages during this time interval, but winter insolation. This supports a contention that was once argued by George Kukla. He is quoted in Imbrie and Imbrie (1979) as having said, in 1968:

> *When this problem has been clarified, and the importance of winter accepted, the chance selection of summer will probably be considered the most serious mistake in Quaternary research in recent years.*
>
> (Kukla, 1968)

Note that total insolation on the Earth depends only on eccentricity. Obliquity and precession determine distribution, not the total amount. However, the *total* yearly insolation at the poles depends only on obliquity—provided that we include both poles. This is obvious, but nevertheless, it is shown in Fig. 2.16.

Thus total insolation is a contender for the driving force for the 1 to 3 Ma period ice ages. However, the absence of a 23 kyr cycle in this spectrum is the result of a

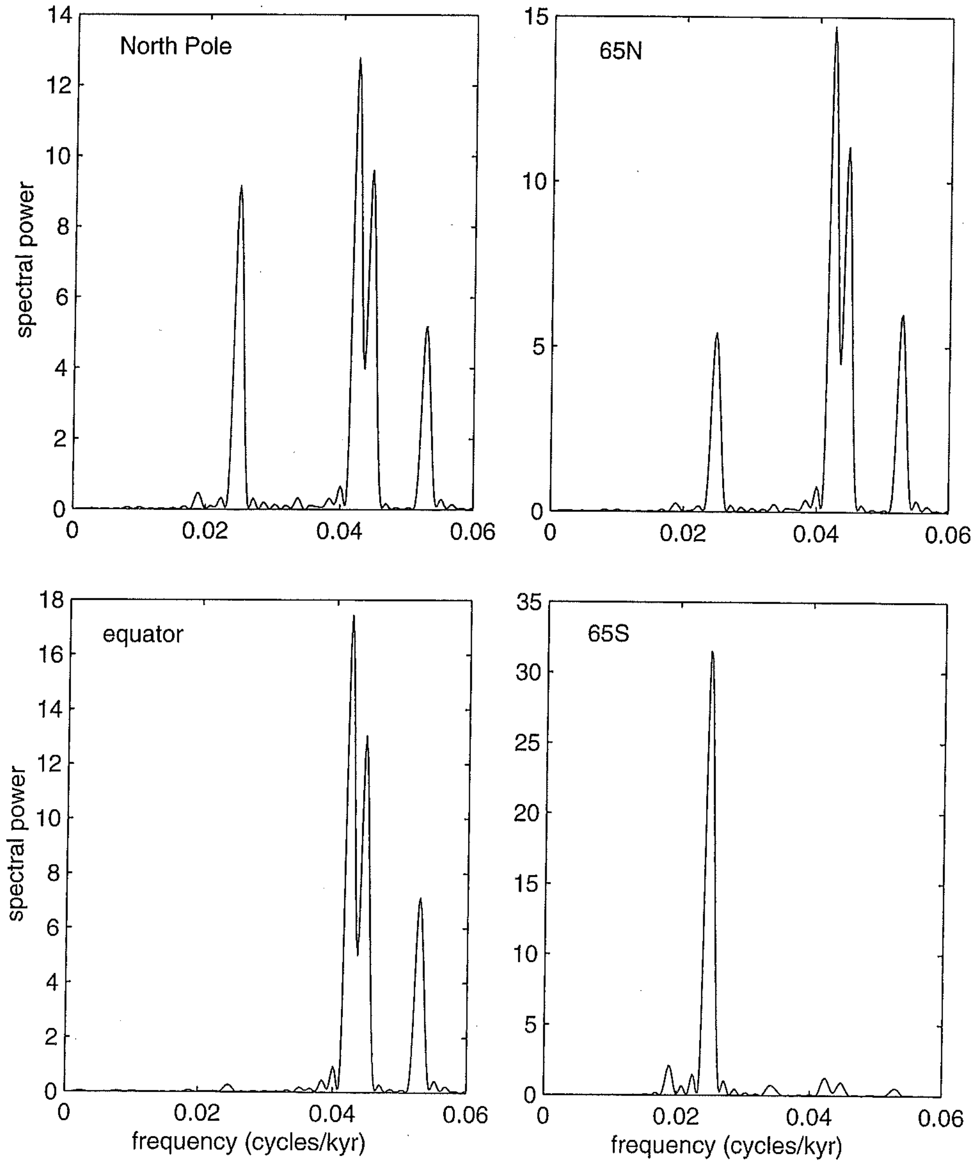

Fig. 2.14. Spectra of July insolation, at latitudes 90N, 65N, 0 and 65S.

precise cancellation of these terms in the Northern and Southern Hemispheres. Such an exact cancellation appears unlikely, so we favour the alternative explanation that the driving force is winter insolation.

Nevertheless, the insolation models for the past million years are usually based on Northern Hemisphere summer insolation. It is this insolation that Milankovitch theorists believe can best explain the features of the ice ages during that period. We will discuss these theories at length in Chapter 6.

2.2.6 Orbital inclination

The orbit of the Earth lies in a plane, but this plane is not identical to that of other planets. It differs from that of Venus by 3.4 degrees, from that of Jupiter by 1.3

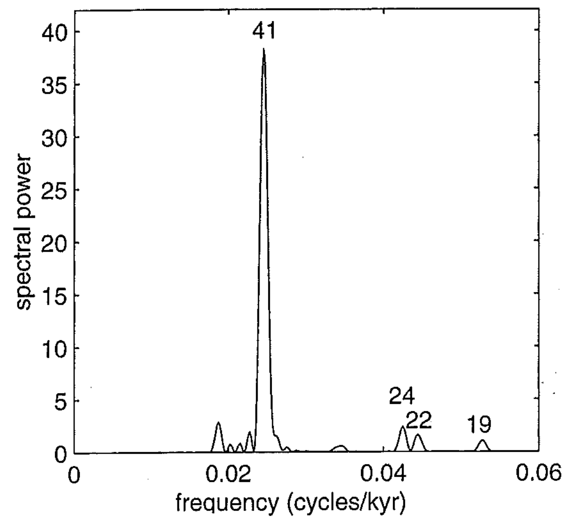

Fig. 2.15. Spectrum of January insolation, 65N.

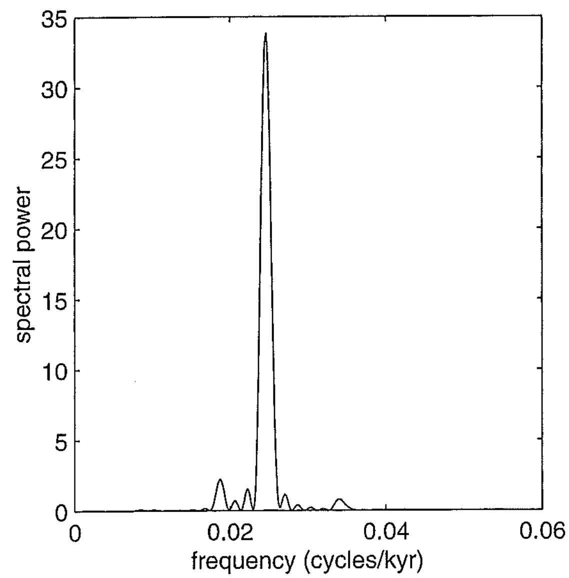

Fig. 2.16. Spectrum of total yearly insolation at poles.

degrees, and from the average of all the other planets by 1.7 degrees. In astronomy, the plane of the Earth's orbit has often been used as a convenient reference. It is called the "ecliptic". As viewed from the Earth, the Sun moves in this plane, and so it is also the plane of the signs of the Zodiac. But this is not a good reference plane for paleoclimate, because it changes with time. It is perturbed by the torques of Jupiter, Venus and Saturn, which give the Earth's motion a horizontal component of orbital angular momentum. (We distinguish between the orbital angular momentum of the Earth's motion, and that of the Earth's spin about its axis.) When we refer to the orbital plane, we must state the year that we are referring to. Croll used the plane of 1850, and present-day paleoclimate studies have generally stuck with this convention, so old calculations can be compared with newer ones. By this convention, the tilt of the Earth's present orbit is very close to, but not quite, zero.

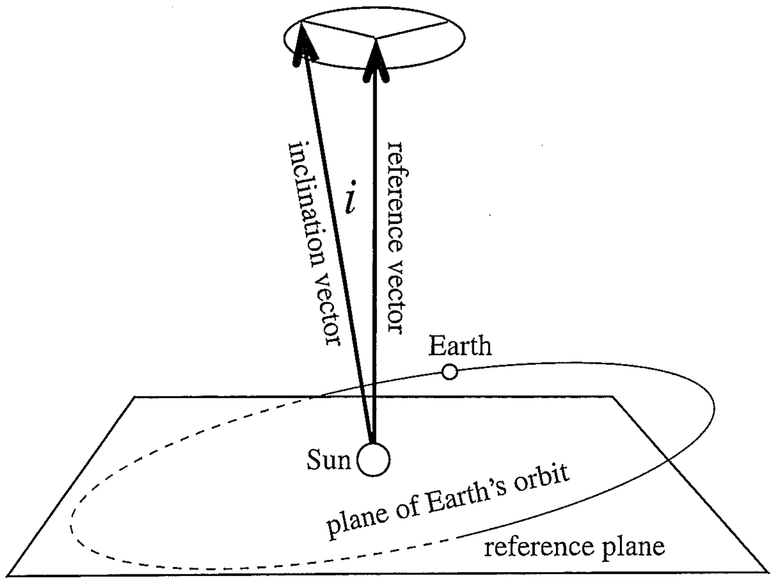

Fig. 2.17. i and Ω for the Earth's orbit.

The orbital inclination of the Earth is the angle between the reference plane and the plane of the Earth's orbit. It is also the angle between vectors that are perpendicular to these two planes, as illustrated in Fig. 2.17.

If our reference vector is the one that is perpendicular to the plane of 1850, then we get the simple result that the present orbital inclination is nearly zero. As we will show below, this apparently trivial choice of reference vector has led to important confusion in the field of paleoclimate. In Fig. 2.18 we plot the orbital inclination i for the last 600 ky, based on orbital calculations by Quinn et al. (1991). Note that it goes to zero *only* at the present (actually, in 1850). It is very unlikely that the orbital vector of the Earth was ever at this exact location before, so although the inclination approaches zero (going back in time), it never quite touches zero. Also notice that it is always positive. That, too, is a trivial result of the convention that you can always state the angle between two vectors as a positive number.

As can be seen from the plot, the inclination oscillates with a period of about 70 kyr. Little attention was paid to this variation, because Northern-Hemisphere

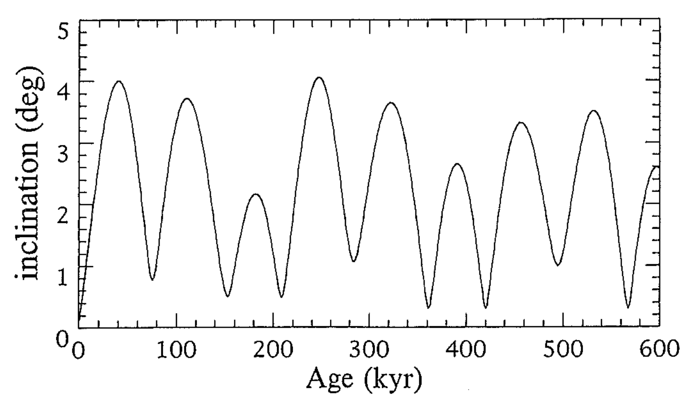

Fig. 2.18. Orbital inclination, referred to Zodiac, for the past 600 kyr.

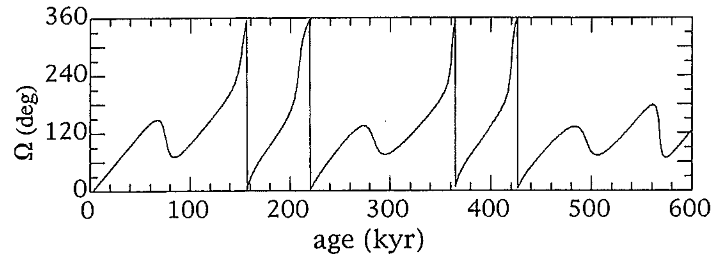

Fig. 2.19. Ω, in Zodiac frame.

insolation does not depend directly on this parameter and, according to the standard Croll/Milankovitch theory, such insolation is the only physical parameter that affects climate.

Of course, the direction of the inclination vector requires a second angle to describe its relation to the reference vector. The second angle is called Ω, and it can be defined in terms of the line of intersection of the two planes. This line is traditionally called a node, and the node on the side of the orbit for which the Earth is rising above the reference plane is called the "ascending node". This terminology will be found in the literature. Ω is then the angle between this line and a reference direction. Equivalently, we can define Ω as the azimuthal angle of the inclination vector, as shown in Fig. 2.17. The value of Ω vs time is shown Fig. 2.19.

This variation, at first look, appears irregular and chaotic, unlike many of the other plots we have seen above. However, we will show this is just an artifact of the choice of reference vector! If we plot the position of the tip of the inclination vector vs. time, we get the graph shown in Fig. 2.20.

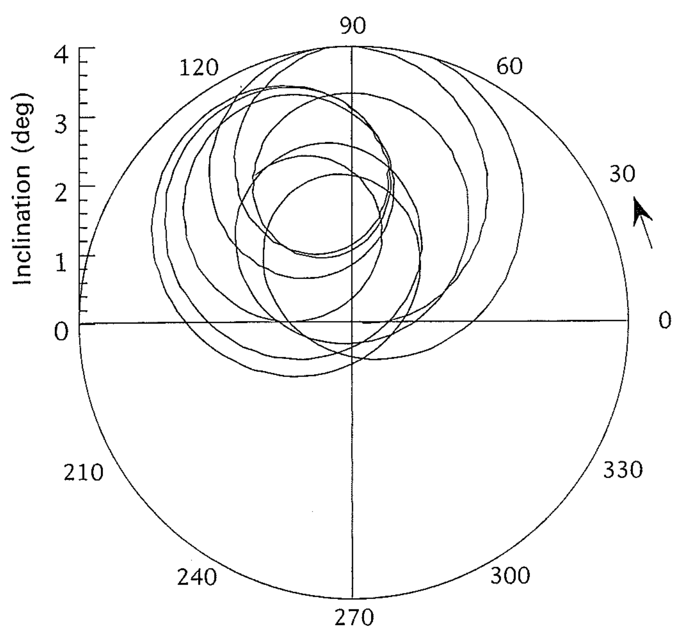

Fig. 2.20. Polar plot of inclination, referred to the Zodiac.

In this plot, i is the radial coordinate and Ω is plotted as the azimuth angle. In this plot, the behaviour of the inclination vector appears relatively simple. It goes in a circle with a slowly varying radius. The sudden jumps in Fig. 2.19 occur during those times when the vector passes close to the origin (the present inclination vector of the Earth), and thereby makes a sudden (but unphysical) jump in Ω. The inclination vector changes slowly and smoothly, and the jump is just a result of the fact that we chose the present orbit of the Earth as our reference.

A much better reference frame for paleoclimate work is the "invariable plane", defined as the plane perpendicular to the angular momentum vector of the planets. This plane is fixed in time, unlike the plane of the Earth's orbit. The invariable plane is close to the plane of Jupiter's orbit, since Jupiter dominates the total angular momentum. The invariable plane is also the average plane of the Earth's orbit. The point that the vector circles turns out to be essentially identical to the vector that defines the invariable plane. If we chose the invariable plane as the reference vector, instead of the 1850 vector, then the polar plot is simply shifted from that shown in Fig. 2.20 to that shown in Fig. 2.21.

This is a simple transformation, but it is a nonlinear transformation. The new value of i is a nonlinear combination of the former i and Ω; likewise for the new value of i. The new values of i and Ω in this new frame are plotted in Fig. 2.22, as a function of time.

Note that, with this new reference frame, the behaviour of the inclination vector is simplified. The value of i is nearly sinusoidal, with a period of about 100 kyr. This fact was first noted by us (Muller, 1994; Muller and MacDonald, 1995) and led us to

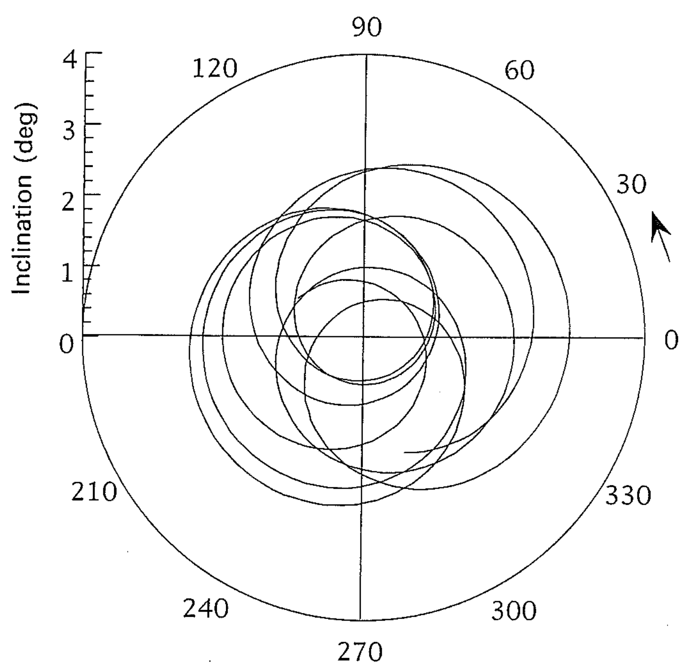

Fig. 2.21. Polar plot of inclination, referred to invariable plane.

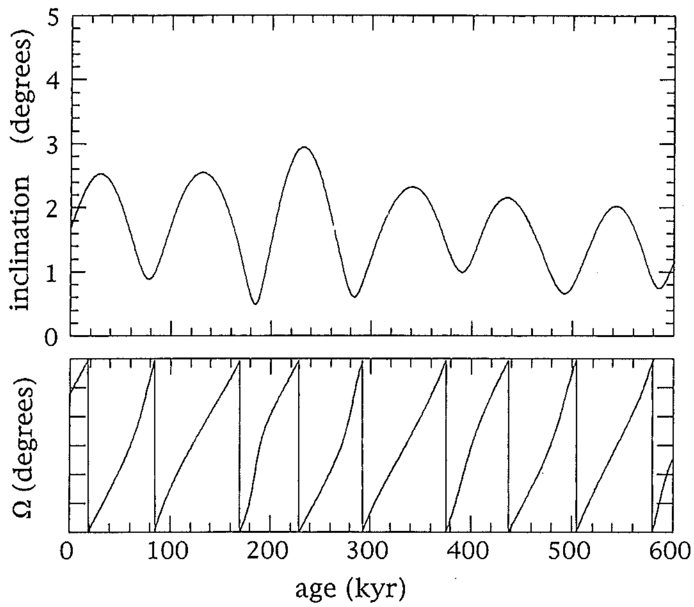

Fig. 2.22. Inclination and Ω, referred to invariable frame, for the last 600 kyr.

speculate that this variation could play a role in the 100 kyr cycle of the ice ages. The behaviour of Ω is also much more regular. It shows a nearly constant increase of 360 degrees every 70 kyr. The sudden jumps are simply the return of the azimuth angle to 0 every time it surpasses 360 degrees.