ELSEVIER

Journal of Neuroscience Methods 83 (1998) 15–34

JOURNAL OF NEUROSCIENCE METHODS

The Fourier analysis of biological transients

Christopher M. Harris *

Department of Ophthalmology, Great Ormond Street Hospital for Children NHS Trust, and Institute of Child Health, University College London, London WC1N 3JH, UK

Received 7 April 1998; accepted 9 April 1998

Abstract

With modern computing technology the digital implementation of the Fourier transform is widely available, mostly in the form of the fast Fourier transform (FFT). Although the FFT has become almost synonymous with the Fourier transform, it is a fast numerical technique for computing the discrete Fourier transform (DFT) of a finite sequence of sampled data. The DFT is not directly equivalent to the continuous Fourier transform of the underlying biological signal, which becomes important when analyzing biological transients. Although this distinction is well known by some, for many it leads to confusion in how to interpret the FFT of biological data, and in how to precondition data so as to yield a more accurate Fourier transform using the FFT. We review here the fundamentals of Fourier analysis with emphasis on the analysis of transient signals. As an example of a transient, we consider the human saccade to illustrate the pitfalls and advantages of various Fourier analyses.

Keywords: Biological transients; Fast Fourier transform; Discrete Fourier transform

1. Introduction

The Fourier transform (FT) has become a basic tool for analysis of many biological signals. The popularity of the FT lies partly in its intuitive appeal; we all know what sines and cosines look like, and the notion that a signal can be broken down into a spectrum of sinusoids with different frequencies is familiar to many of us, just as we can see light broken down into its spectrum by a prism. The FT is also fundamental to linear systems theory where, via the convolution theorem, the spectrum of the output is simply the product of spectrum of the input and the transfer function of the system under study. Indeed, the first line of investigation of a biological system is often to model it as a linear system.

Fourier analysis has its intuitive appeal for long periodic signals, where there are many repetitions or cycles of some temporal pattern. However, measured biological signals are always finite in time. One reason is that we cannot observe or collect data forever and must decide how long to measure the phenomenon to best capture the desired information. Usually recording as many cycles as possible is the best policy, although this may not always be the case for non-stationary signals, but in any case the decision is at least under control of the experimenter. Another reason, which forms the focus of this article, is that the biological phenomenon is itself a brief event or 'transient', in which there is no, or very few, repetitions of an underlying cycle. For example, fast muscular movements, or the underlying neuromuscular signals that drive such movements are single events in time. In such cases, Fourier analysis can be physiologically informative, but it is not a 'natural' approach and it is neither intuitive nor trivial. This is because a transient has a beginning and an end, and so there must be discontinuities at the beginning and end. The Fourier spectrum of a transient depends strongly on the temporal separation and type of these discontinuities, and may be completely domi-

* Tel.: +44 171 4059200 (ext. 0284); fax: +44 171 8298647; e-mail: chris@vissci.ion.ucl.ac.uk

0165-0270/98/$19.00
PII S0165-0270(98)00080-6

nated by them rather than the signal during the transient. Unfortunately the use of discrete Fourier techniques, as epitomized by the FFT, only leads to further obfuscation.

This article is divided into four sections in which we review Fourier theory and apply it to transients. In Section 2 we describe the basics of Fourier theory by starting with the Fourier series (FS), rather than with the Fourier transform (FT) per se. Unfortunately, the FS tends to be under-represented in textbooks, only to re-emerge unrecognized as the discrete Fourier transform (DFT). In Section 3, some general Fourier properties of transients are introduced. Here we discuss the effects of discontinuities on the spectrum and how they are affected by filtering (convolution). In Section 4 we describe some of the basic principles of sampling. The popularity of the FT has grown because of the increasing availability of computer software packages that can churn out DFTs at the press of a mouse-button. These computer algorithms have been based mostly on the FFT, which has become so common as to be considered almost synonymous with the Fourier transform—this is far from true.

In the last section we apply Fourier theory to an actual biological transient signal, namely the human saccade. The saccade has been chosen because: (a) its Fourier transform (when estimated correctly) reveals physiological information that cannot be seen in the temporal signal; (b) simple application of the FFT does not reveal these features; and (c) a previous attempt at the Fourier analysis of saccades reveals how even sophisticated analysis can go astray if the FT is incorrectly computed.

2. Basics

2.1. The trigonometric Fourier series

A periodic signal is one that repeats itself precisely end-on-end from-ever forever, such as an infinitely long sine wave. As first shown by Fourier, periodic signals can be expressed as an infinite sum of sines and cosines at discrete harmonic frequencies[1] of the fundamental period of the signal. Even discontinuous periodic signals, such as square waves, can be represented as a FS.

[1] The scientific use of the word 'harmonic' usually describes an integer multiple of a fundamental frequency, where the 1st harmonic is at the fundamental frequency, and the 2nd harmonic occurs at twice the fundamental frequency, etc. In music, by distinction, the 1st harmonic is usually an octave above the fundamental frequency. We shall use the former.

Assume that some function $g(t)$ is periodic with a fundamental frequency of f_g, then this function can be expressed as an infinite series:

$$\begin{aligned} g(t) &= a_0 + a_1\cos(\omega_g t) + a_2\cos(2\omega_g t) \\ &\quad + a_3\cos(3\omega_g t) + \ldots \\ &\quad + b_1\sin(\omega_g t) + b_2\sin(2\omega_g t) + b_3\sin(3\omega_g t) + \ldots \\ &= a_0 + \sum_{n=1}^{\infty}\left[a_n\cos(n\omega_g t) + b_n\sin(n\omega_g t)\right] \end{aligned} \tag{1}$$

where $\omega_g = 2\pi f_g$ is the *angular* frequency of the fundamental, measured in rad/s. In general an infinite number of terms are required, but if $g(t)$ can be expressed as a finite series then $g(t)$ is bandwidth-limited. Two numbers (or degrees of freedom) are needed for each harmonic frequency, a_n and b_n, because a sinusoid is only fully defined when both its amplitude and phase are specified. Sinusoids can be expressed in different ways (Table 1), and an alternative expression of the FS is:

$$\begin{aligned} g(t) &= c_0 + c_1\cos(\omega_g t + \Psi_1) + c_2\cos(2\omega_g t + \Psi_2) \\ &\quad + c_3\cos(3\omega_g t + \Psi_3) + \ldots \\ &= c_0 + \sum_{n=1}^{\infty} c_n\cos(n\omega_g t + \Psi_n) \end{aligned} \tag{2}$$

where the amplitude and phase of each harmonic are now explicit as c_n and Ψ_n.

For a given $g(t)$, the coefficients of the trigonometric FS are given by:

$$\begin{aligned} a_0 &= \frac{1}{T_g}\int_{\tau}^{\tau+T_g} g(t)\,\mathrm{d}t, \\ a_n &= \frac{2}{T_g}\int_{\tau}^{\tau+T_g} g(t)\cos(n\omega_g t)\mathrm{d}t, \text{ and} \\ b_n &= \frac{2}{T_g}\int_{\tau}^{\tau+T_g} g(t)\sin(n\omega_g t)\mathrm{d}t, \end{aligned} \tag{3}$$

Table 1
Different ways to write a sinusoid

From	To	Coefficients
$\cos(\omega t)$	$(e^{i\omega t}+e^{-i\omega t})/2$	
$\sin(\omega t)$	$(e^{i\omega t}-e^{-i\omega t})/2i$	
$e^{i\omega t}$	$\cos(\omega t)+i\sin(\omega t)$	
$e^{-i\omega t}$	$\cos(\omega t)-i\sin(\omega t)$	
$\sin(\omega t+\Psi)$	$A\sin(\omega t)+B\cos(\omega t)$	$A=\cos\Psi;\ B=\sin\Psi;\ A^2+B^2=1;\ \tan\Psi=B/A$
$\sin(\omega t+\Psi)$	$D\,e^{i\omega t}+E\,e^{-i\omega t}$	$2D=\sin\Psi-i\cos\Psi,\ 2E=\sin\Psi+i\cos\Psi,\ \sin\Psi=D+E,\ \cos\Psi=i(D-E)$
$A\sin(\omega t)+B\cos(\omega t)$	$D\,e^{i\omega t}+E\,e^{-i\omega t}$	$2D=B-iA,\ 2E=B+iA;\ B=D+E,\ A=i(D-E)$

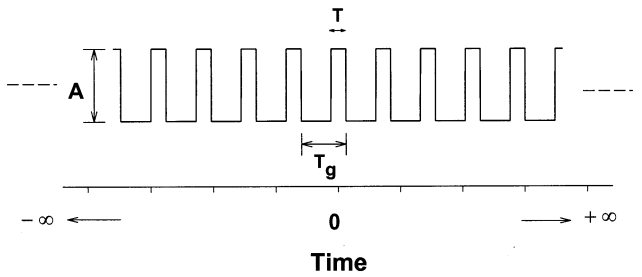

Fig. 1. A symmetric rectangle wave as an example of an infinitely long periodic function of amplitude A, period T_g, and duty cycle of T/T_g. The Fourier series of this function is given in Eq. (4).

where, $n \geq 1$, T_g is the period of $g(t)$, and τ is arbitrary, that is, the integration can be over any convenient whole period of $g(t)$. An important example is the rectangle wave with period T_g, and height A, where the wave is 'on' for a time T and 'off' for $T_g - T$ (i.e. a duty-cycle of T/T_g) (Fig. 1). Assume that the wave is symmetrical about the origin (that is, $g(t)$ is an even function), then the b_n will be zero and the a_n will be given by:

$$a_0 = \frac{AT}{T_g},$$

$$a_n = \frac{2A}{T_g}\int_{-T/2}^{T/2} \cos(n\omega_g t)\,\mathrm{d}t = \frac{2A\,\sin(n\pi T/T_g)}{n\pi}$$

$$= \frac{2AT}{T_g}\,\mathrm{sinc}(n\pi T/T_g) \tag{4}$$

where, because the function $\sin x/x$ is so important in Fourier theory, it is given the shorthand label: $\mathrm{sinc}(x)$ (pronounced sink of x). The square wave is the most commonly described rectangle wave in textbooks, and its FS is given by Eqs. (1) and (4) when $T = T_g/2$:

$$g(t) = \frac{A}{2} + \frac{2A}{\pi}\left\{\cos(\omega_g t) - \frac{\cos(3\omega_g t)}{3} + \frac{\cos(5\omega_g t)}{5} - \ldots\right\} \tag{5}$$

which has no energy at even harmonics. The traditional focus on the square wave tends to hide the role of the sinc function in Eq. (4), as illustrated for the even function when $T = T_g/4$, in which case the FS becomes:

$$g(t) = \frac{A}{4} + \frac{2A}{\pi}\left\{\begin{array}{l} \dfrac{\cos(\omega_g t)}{\sqrt{2}} + \dfrac{\cos(2\omega_g t)}{2} + \dfrac{\cos(3\omega_g t)}{3\sqrt{2}} \\ -\dfrac{\cos(5\omega_g t)}{5\sqrt{2}} - \dfrac{\cos(6\omega_g t)}{6} - \dfrac{\cos(7\omega_g t)}{7\sqrt{2}} \\ +\ldots \end{array}\right\} \tag{6}$$

2.2. The exponential Fourier series

Another way to express the FS is as an exponential series:

$$g(t) = \ldots + d_{-2}\,\mathrm{e}^{-i2\omega_g t} + d_{-1}\,\mathrm{e}^{-i\omega_g t} + d_0 + d_1\,\mathrm{e}^{i\omega_g t} + d_2\,\mathrm{e}^{i2\omega_g t} + \ldots \tag{7}$$

$$= \sum_{n=-\infty}^{\infty} d_n\,\mathrm{e}^{in\omega_g t}$$

sometimes called the exponential Fourier series. Algebraically this is more elegant, but less intuitive. Here, each harmonic is still defined by two numbers, d_n and d_{-n}, but these are now coefficients of a positive and a negative frequency; they can be derived from Euler's relations (Table 1):

$$a\cos x + b\sin x = a(\mathrm{e}^{ix} + \mathrm{e}^{-ix})/2 + b(\mathrm{e}^{ix} - e^{-ix})/2i$$

$$= \left(\frac{a+ib}{2}\right)\mathrm{e}^{-ix} + \left(\frac{a-ib}{2}\right)\mathrm{e}^{ix}, \tag{8}$$

alternatively:

$$c\cos(x+\Psi) = c(\mathrm{e}^{i(x+\Psi)} + \mathrm{e}^{-i(x+\Psi)})/2$$

$$= \left(\frac{c}{2}\mathrm{e}^{-i\Psi}\right)\mathrm{e}^{-ix} + \left(\frac{c}{2}\mathrm{e}^{+i\Psi}\right)\mathrm{e}^{ix}. \tag{9}$$

The relationships among these coefficients are shown in Table 1.

Note that the coefficients d_i are now complex numbers. However, provided our original function, $g(t)$, is real (as biological data usually are), then the coefficients of corresponding positive and negative frequencies are conjugate pairs (Eqs. (8) and (9)): $d_i = d^*_{-i}$; thus we still need two numbers to describe each frequency. The coefficients are given by:

$$d_n = \frac{1}{T_g}\int_{\tau}^{\tau+T_g} g(t)\,\mathrm{e}^{-in\omega_g t}\mathrm{d}t \tag{10}$$

(note the absent factor of 2). It should be emphasized that we will obtain exactly the same series whether we use Eqs. (1), (2) or (7). It is merely a matter of how we wish to express sinusoids.

2.3. The Fourier transform (integral)

We could plot the coefficients of the exponential FS along a continuous frequency axis, where values would only be defined at discrete frequencies that are harmonics of the fundamental of $g(t)$. However, if we made the period very long, the points along the frequency axis would become ever closer, and in the limit as the period approached infinity, the points would eventually become continuous. This is one way to derive the Fourier transform (FT)—a limiting case of the exponential FS.

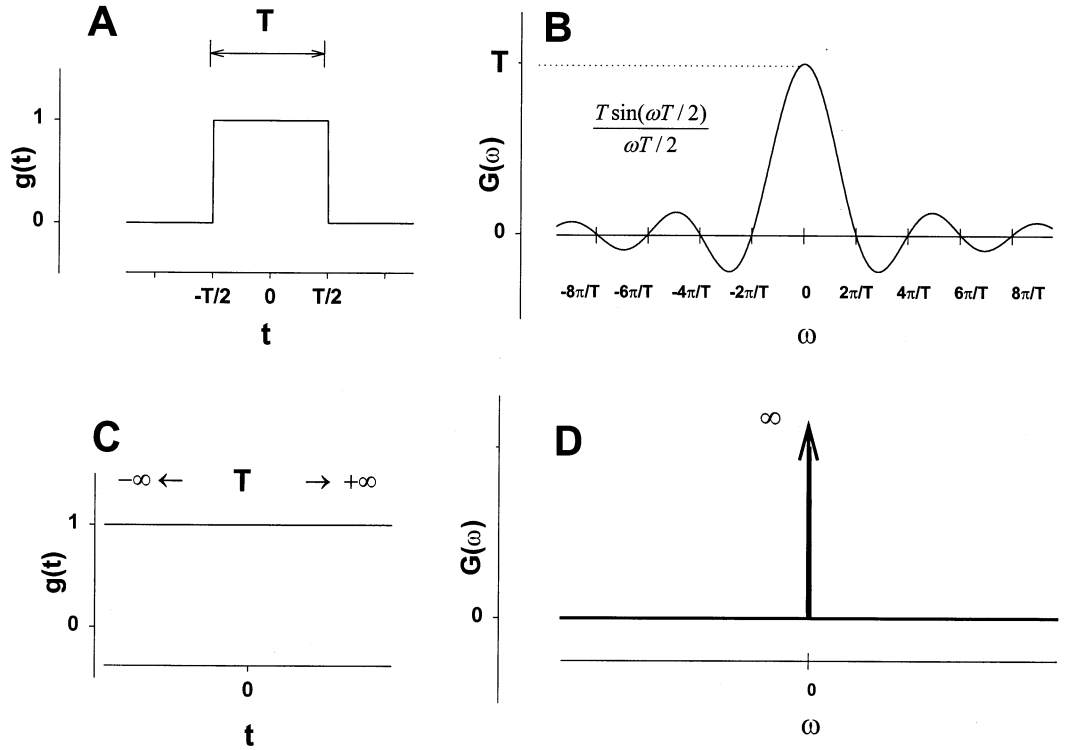

Fig. 2. (A) A rectangle of unit height and width T. (B) Fourier transform of rectangle in (A) given by the sinc function of height T with zero crossings at harmonic frequencies; (C) a unit constant is the limiting case of a rectangle with infinite width (D) Fourier transform of limiting case is a 'delta' or 'impulse' function of infinite height and infinitesimal width.

The Fourier transform FT of a function $g(t)$ is defined as:

$$G(\omega) = \int_{-\infty}^{\infty} g(t)\, \mathrm{e}^{-i\omega t} \mathrm{d}t, \tag{11}$$

and the inverse is:

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega)\, \mathrm{e}^{+i\omega t} \mathrm{d}\omega; \tag{12}$$

note the asymmetric exponential sign and the scale factor of $1/2\pi$. As in most texts, the temporal function will be written in lower-case, and its FT in upper-case. It should be noted that the FT can be written in slightly different ways as seen in various texts. Thus, the $1/2\pi$ may scale the forward transform instead of the inverse, or $1/\sqrt{2\pi}$ may occur on both; or no $1/2\pi$ occurs if $2\pi f$ is used instead ω; the forward transform may have a positive exponent instead of negative. Therefore, some care is needed in comparing FTs from different sources.

As in the exponential FS, the FT contains negative frequencies. In some applications negative frequencies have a physical representation, as in diffraction patterns, but in most there are no physical counterparts. Their existence must simply be accepted as an intrinsic part of Fourier analysis to allow amplitude and phase to be coded, as in Eq. (7).

It is important to be aware that $G(\omega)$ is a complex function: $G(\omega) = G_R(\omega) + iG_i(\omega)$, where $G_R(\omega)$ and $G_I(\omega)$ are the real and imaginary components, respectively. As with the exponential FS, provided $g(t)$ is a real function $G(\omega) = G^*(-\omega)$, so that the negative frequencies do not carry information additional to that in the positive frequencies (this is not true if $g(t)$ is complex).

Another way to write the FT is as: $G(\omega) = |G(\omega)|\mathrm{e}^{i\Psi(\omega)}$, where the modulus is given by:

$$|G(\omega)| = +\sqrt{G(\omega)G^*(\omega)} = +\sqrt{G_R^2(\omega) + G_I^2(\omega)}$$

and the phase by:

$$\Psi(\omega) = \tan^{-1}\left[\frac{G_I(\omega)}{G_R(\omega)}\right]. \tag{13a}$$

The energy spectrum is the square of the modulus:

$$E_G(\omega) = |G(\omega)|^2, \tag{13b}$$

where the word 'energy' is used rather liberally. Some authors use the term 'power', but this should be reserved for infinite periodic and stochastic signals.

2.4. Fourier pairs

Clearly, we can swap between $g(t)$ and $G(\omega)$ with no loss or gain of information, and they are called 'Fourier pairs'. There are many important relationships between Fourier pairs that help to calculate as well as understand FTs. These are described in many textbooks (e.g.

Table 2
Fourier-pair relationships

Operation	Time domain	Frequency domain
Fourier transform	$g(t) = FT^{-1}\{G\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(\omega)\, e^{+i\omega t}\, d\omega$	$G(\omega) = FT\{g\} = \int_{-\infty}^{\infty} g(t)\, e^{-i\omega t}\, dt$
	(inverse Fourier transform)	(Fourier transform)
Linearity	$ag(t)+bh(t)$	$aG(\omega)+bH(\omega)$
Scaling	$g(at)$	$\frac{1}{\lvert a\rvert} G\left(\frac{\omega}{a}\right)$
Energy	$\int_{-\infty}^{\infty} \lvert g(t)\rvert^2 dt$	$\frac{1}{2\pi}\int_{-\infty}^{\infty} \lvert G(\omega)\rvert^2\, d\omega$
Area	$\int_{-\infty}^{\infty} g(t)\, dt$	$G(0)$
	$2\pi g(0)$	$\int_{-\infty}^{\infty} G(\omega)\, d\omega$
Shifting	$g(t+T)$ (backwards in time by T) $g(t-T)$ (fowards in time by T) $g(t)\, e^{-i\omega\Omega}$ $g(t)\, e^{+i\omega\Omega}$	$G(\omega)\, e^{+\omega T}$ $G(\omega)\, e^{-i\omega T}$ $G(\omega t+\Omega)$ (to lower frequency by Ω) $G(\omega t-\Omega)$ (to higher frequency by Ω)
Differentiation	$\frac{d^n g(t)}{dt^n}$	$(i\omega)^n G(\omega)$
	$(-it)^n g(t)$	$\frac{d^n G(\omega)}{d\omega^n}$
Convolution	$\int_{-\infty}^{\infty} g(t-\tau)h(\tau)\, d\tau = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)\, d\tau = g\otimes h$	$C(\omega) = G(\omega)H(\omega)$
Product	$c(t) = 2\pi g(t)h(t)$	$\int_{-\infty}^{\infty} G(\omega-\omega_1)H(\omega_1)d_1 = \int_{-\infty}^{\infty} G(\omega_1)H(\omega-\omega_1)\, d\omega_1 = G\otimes H$

Champeney, 1973; Bracewell, 1986), and some of the more common ones are summarised in Table 2.

A useful property of FTs is the ease with which a function can be differentiated. Differentiating both sides of Eq. (12) with respect to time yields:

$$\frac{dg}{dt} = \frac{1}{2\pi}\int_{-\infty}^{\infty} (i\omega)G(\omega)\, e^{+i\omega t}\, d\omega,$$

and we see that differentiating in time is equivalent to multiplying the FT by $i\omega$, so in general:

$$\frac{d^n g}{dt^n} \Leftrightarrow (i\omega)^n G(\omega). \quad (14)$$

Thus, differentiation augments relatively the high frequencies. Consequently, and not surprisingly, differentiation makes functions less smooth. Similarly integration is equivalent to dividing the FT by iw (although some care is needed with the constant of integration), which emphasizes low frequencies and makes functions smoother. For energy spectra, differentiating or integrating n times is equivalent to multiplying or dividing by ω^{2n}.

2.5. Relation to the Laplace transform

The Laplace transform is commonly used in the analysis of control systems, and is very useful in finding the transfer functions of systems with feedback. The Laplace transform of $g(t)$ is usually written as

$$G(s) = \int_0^\infty g(t)\, \mathrm{e}^{-st}\, \mathrm{d}t,$$

where s is the Laplace variable. By making the substitution $s \rightarrow i\omega$, the FT $G(\omega)$ can be found for functions that are zero for negative time.

2.6. Generalised functions

It can be shown that the FT of $g(t)$ can only be found if the integral of its absolute value (IAV) i.e.:

$$\int_{-\infty}^{\infty} |g(t)|\, \mathrm{d}t,$$

is finite. Clearly, the IAV of a constant or a periodic function will not be finite. To circumvent this limitation, Fourier theory has been extended to include so-called generalised functions. Let us assume we want to find the FT of the simple function $g(t) = \mathrm{e}^{i\omega_g t}$ (which is a useful building block for many other functions). Only its phase changes with time, $\Psi(t) = \omega_g t$, and it has a constant unit modulus; thus it does not have a finite IAV since

$$\int_{-\infty}^{\infty} |g(t)|\, \mathrm{d}t = \int_{-\infty}^{\infty} 1.\mathrm{d}t$$

is not finite. We define $f(t)$ as being equal to $g(t)$ over the range $(-T/2 \le t \le T/2)$ and zero elsewhere, ie:

$$f(t) = \begin{cases} \mathrm{e}^{i\omega_g t} & |t| \le T/2 \\ 0 & |t| > T/2 \end{cases},$$

then we can write the FT of $g(t)$ as a limit:

$$\begin{aligned} G(\omega) &= \lim_{T\rightarrow\infty} \left[\int_{-\infty}^{\infty} f(t)\, \mathrm{e}^{-i\omega t}\, \mathrm{d}t \right] \\ &= \lim_{T\rightarrow\infty} \left[\int_{-T/2}^{T/2} (\mathrm{e}^{i\omega_g t})\, \mathrm{e}^{-i\omega t}\, \mathrm{d}t \right] \\ &\quad \lim_{T\rightarrow\infty} [T\, \mathrm{sinc}\{T(\omega_g - \omega)/2\}] \end{aligned} \tag{15}$$

As T become ever larger the sinc function becomes ever narrower and $G(\omega)$ becomes ever higher. In the limit,

$$T\rightarrow\infty,\ G(\omega)\rightarrow 2\pi\delta(\omega_g - \omega),$$

where $\delta(x)$ is called a 'delta' function or 'impulse' (Fig. 2). The delta function has infinite height, infinitesimal width, unit area at x and is zero elsewhere (there are many other ways to derive a delta function, see Champeney (1973)). The validity of generalised functions cannot always be taken for granted (although we shall); see Bracewell (1986) for an informative discussion.

Now consider a general sinusoid expressed in exponential form: $g(t) = d_{-1}\, \mathrm{e}^{-i\omega_g t} + d_1\, \mathrm{e}^{i\omega_g t}$. From Eq. (15), its FT is $G(\omega) = 2\pi d_{-1}\delta(-\omega_g - \omega) + 2\pi d_1\delta(\omega_g - \omega)$; i.e. two delta functions at $\pm\omega_g$. The amplitude and phase of the sinusoid are determined by the coefficients d_{-1} and d_1. For a cosine of unit amplitude $d_{-1} = +1/2$ and $d_1 = +1/2$, for a sine of unit amplitude $d_{-1} = -1/2i$ and $d_1 = +1/2i$.

For a constant, when ω_g and $g(t) = 1$, the FT is $G(\omega) = 2\pi\delta(\omega)$, i.e. a delta function at the origin.

If $g(t)$ is the sum of a constant and sinusoids with frequencies at ω_g, $2\omega_g$, $3\omega_g$,..., then its FT will clearly be a series of delta functions at 0, $\pm\omega_g$, $\pm 2\omega_g$, $\pm 3\omega_g$,..., that is, we have reassuringly recovered the Fourier series.

Similar relationships occur for delta functions in the time domain. Thus, the FT of a delta function $g(t) = \delta(t-\tau)$ is $G(\omega) = \mathrm{e}^{i\omega\tau}$. If the delta function occurs at $t = 0$, the FT of $g(t) = \delta(t)$ is simply a constant $G(\omega) = 1$. For two delta functions separated in time by τ, $g(t) = a\delta(t\text{-}\tau/2) + b\delta(t + \tau/2)$, the FT will be sinusoid in frequency (Table 1):

$$G(\omega) = \sin(\omega\tau/2 + \Psi). \tag{16}$$

An important function in sampling is the infinite series of delta functions, or a 'comb' of delta functions

$$g(t) = \sum_{n=-\infty}^{n=\infty} \delta(t - n\tau),$$

which we shall denote by the symbol $\mathrm{III}(t,\tau)$. It can be shown that the FT is also a delta comb

$$G(\omega) = \frac{2\pi}{\tau} \sum_{m=-\infty}^{m=\infty} \delta\left(\omega - \frac{2\pi m}{\tau}\right) \text{ or } \frac{2\pi}{\tau}\, \mathrm{III}\left(\omega, \frac{2\pi}{\tau}\right).$$

By integrating a single delta function at the origin we obtain a step function at the origin with unit height, $u(t)$. Therefore, the FT of a step function is:

$$U(\omega) = \frac{1}{i\omega} \text{ or } -\frac{i}{\omega} \tag{17}$$

(which could have been derived directly from Eq. (11)).

2.7. Bandwidth-limited functions and smoothness

An important aspect of Fourier theory is whether a function is limited in the frequency or time domain. A function whose FT is zero beyond some finite frequency is said to be bandwidth-limited. Similarly, a function that is zero outside a finite time range is said to be time-limited. A function cannot be both bandwidth-limited and time-limited; if it is finite in one domain it must be infinite in the other. On the other hand, a function that is infinite in one domain may or may not be finite in the other domain. For example, the FT of a Gaussian function is another Gaussian, and therefore, a Gaussian is infinite in both the time and frequency domains.

It is fundamental to Fourier analysis that, as a function becomes more concentrated in time (not necessarily just by truncation), the more spread there is in the frequency domain; the delta function is the extreme example. Similarly, concentrating a function to low frequencies inevitably makes the function longer in time and makes it 'smoother'. In physics, this is called the 'Uncertainty Principle', but it applies to any Fourier pair, in any discipline.

3. Transients

Broadly speaking, a transient is a signal that changes its behaviour in some relatively brief interval, such as a neural impulse, a blink, a reaching movement, or a saccadic eye movement etc. We consider a transient to mean a signal that has a defined start and a defined end. There is some question as to what constitutes the start and an end in measured signals. Some functions, such as a Gaussian, clearly change in some interval, but do not have an abrupt beginning or end because they extend from minus to positive infinity. Other functions may have a definite start but no end, such as the impulse response of a first-order filter, and at least theoretically, functions could have no start but a definite end. In practice, the start/end of asymptotic signals would be defined by the resolution of the recording equipment. However, whether asymptotic functions are just a mathematical abstraction or genuinely exist in real signals is a matter of debate (Slepian, 1983). In any case, we believe that the brain is capable of being decisive in time, and we shall assume that transient biological signals have a genuine start and end.

We will define a transient as a function that has discontinuities at the beginning and end (and may have other discontinuities during the transient). A discontinuity occurs when a time derivative of the function becomes infinite. If we differentiate a transient again and again, eventually a derivative will become infinite at the beginning and end of the function. In this article we shall denote the 'order' of a discontinuity as the number of times a function must be differentiated to yield the infinite derivative. For example a rectangle has first-order discontinuities because its derivative yields delta functions at the start and end (a delta function is the derivative of a step). A delta function is itself discontinuous and hence is zero-order. As another simple example consider the parabolic transient of duration T:

$$g(t) = \begin{cases} 1 - 4\left(\frac{t}{T}\right)^2 & |t| \le T/2 \\ 0 & |t| > T/2 \end{cases} \qquad (18)$$

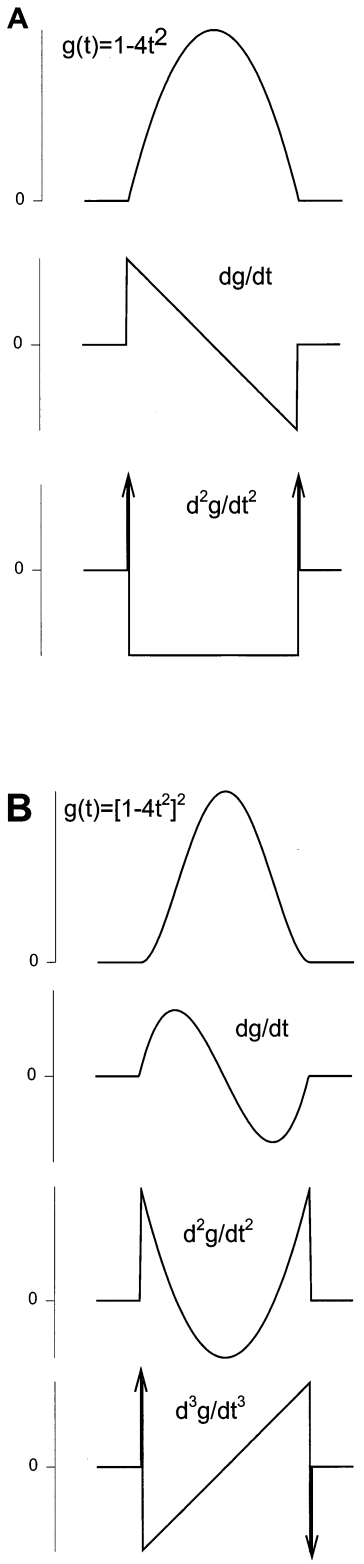

Fig. 3. Illustration of impulsive discontinuities in some simple transients revealed by repeated differentiation. (A) A parabola given by Eq. (18) with unit width and height truncated at $t = \pm 0.5$. After differentiating twice, two delta functions (impulses) are revealed indicating second-order discontinuities. (B) A squared parabola (Eq. (21)) requires three differentiations before delta functions are revealed, indicating a third-order discontinuities.

which is symmetrical about the origin (Fig. 3A). Differentiating once yields:

$$\frac{dg}{dt} = \begin{cases} -\frac{8t}{T} & |t| \le T/2 \\ 0 & |t| > T/2 \end{cases}$$

and differentiating again, we obtain:

$$\frac{d^2g}{dt^2} = [-8 + 8\delta(t + T/2) + 8\delta(t - T/2)]/T$$

which has delta functions indicating discontinuities at $t = \pm T/2$, i.e. the beginning and end of the transient. Thus, this function has second-order discontinuities. In some transients the order of the discontinuities at the start and end may differ, and moreover, not all transients develop delta functions on repeated differentiation, such as $\log(t)$ or t^p where p is not a positive integer. These types of functions tend to have complicated FTs and will not be considered here.

3.1. The Fourier transform of a transient

As shown earlier (Eq. (14)), differentiating n times corresponds to multiplying the Fourier transform by $(i\omega)^n$. From Eq. (16), we know that the FT of two delta functions separated in time is a sinusoid. Therefore, a transient that develops two delta functions after differentiating n times will have a FT that obeys:

$$G(\omega) \propto \frac{\sin(\Psi + \omega T/2) + R(\omega)}{(i\omega)^n}, \tag{19}$$

where $R(\omega)$ is the FT of a function $r(t)$ given by the nth derivative of $g(t)$ without the delta functions. Thus, we see that $G(\omega)$ will tend to be oscillatory, and for an nth order discontinuity the FT decays at high frequencies at least as fast as $1/\omega^n$.

From Eq. (19) we see that through the interaction of the start and end discontinuities, the FT of a transient will oscillate depending on the temporal separation of the discontinuities, T, and $R(\omega)$. For functions where $r(t)$ is relatively smooth (between the start and end), $R(\omega)$ will decay at high frequencies, so that as $\omega \to \infty$,

$$G(\omega) \propto \frac{\sin(\Psi + \omega T/2)}{(i\omega)^n} \tag{20}$$

Thus, the FT of the transient at high frequencies will oscillate and will pass through zero twice in each cycle, and the frequency separation of these zeroes will be given by the reciprocal of the duration of the transient. In the energy spectrum, these zeroes will lead to sharp minima at frequencies given when $G(\omega) = 0$. For many illustrations of this phenomena see Champeney (1973) and Percival and Walden (1993).

A special case is when the transient is perfectly symmetrical (even) or anti-symmetrical (odd), where it follows that for high frequencies either

$$G(\omega) \propto \frac{\cos(\omega T/2)}{(i\omega)^n} \text{ or } G(\omega) \propto \frac{\sin(\omega T/2)}{(i\omega)^n}$$

depending on whether the discontinuity delta functions are of the same or opposite sign. Then the zeroes at high frequencies will occur at $\omega = 2\pi(m + 1/2)/T$ or $\omega = 2\pi m/T$, respectively, where m is an integer. The converse is not true, however: a transient with zeroes at $\omega = 2\pi(m + 1/2)/T$ or $\omega = 2\pi m/T$ is not necessarily symmetrical/anti-symmetrical. For small ω, the minima will occur at frequencies that depend on the shape, particularly the smoothness, of the function, $r(t)$, and will not in general be simple multiples of π/T. For example, the FT of the parabola in Eq. (18) develops delta functions of the same sign for $n = 2$ (Fig. 3A); the FT of the parabola is

$$G(\omega) = 8T\left[\frac{2\sin(\omega T/2)}{(\omega T)^3} - \frac{\cos(\omega T/2)}{(\omega T)^2}\right],$$

which is dominated by the last term at high frequencies. The zeroes occur at approximately $\omega = 1.44(2\pi/T)$, $2.46(2\pi/T)$, $3.48(2\pi/T)$, $4.48(2\pi/T)$, $5.50(2\pi/T)$, $6.50(2\pi/T)$, etc. showing the asymptote to $\omega = 2\pi(m + 1/2)/T$. For convenience we shall express the zeroes in temporal frequency, which occur at approximately $f = 1.44/T$, $2.46/T$, $3.48/T$, $4.48/T$, $5.50/T$, $6.50/T$.

As another example, consider the transient given by the quartic equation:

$$g(t) \begin{cases} \left[1 - 4\left(\frac{t}{T}\right)^2\right]^2 & |t| \le T/2 \\ 0 & |t| > T/2 \end{cases}, \tag{21}$$

which is the classic 'minimum jerk' equation that has been used to fit arm movement trajectories (Hogan, 1984) (jerk is the rate of change of acceleration). This function develops delta functions with opposite signs after three differentiations (Fig. 3B). Thus, for high frequencies we expect

$$G(\omega) \propto \frac{\sin(\omega T/2)}{(i\omega)^3},$$

with zeroes at $f = m/T$. It can be shown that FT of the minimum jerk profile is

$$G(\omega) = 64T\left[\frac{12\sin(\omega T/2)}{(\omega T)^5} - \frac{6\cos(\omega T/2)}{(\omega T)^4} - \frac{\sin(\omega T/2)}{(\omega T)^3}\right],$$

which is dominated by the last term at high frequencies and has zeroes at approximately $f = 1.84/T$, $2.90/T$, $3.94/T$, $5.96/T$, $6.96/T$, $7.98/T$… being close to and eventually reaching integer multiples of $1/T$.

3.2. Convolution theorem and transients

An important application of the FT is in unravelling the convolution in linear systems. The output of a linear filter, $o(t)$, at some instant t depends not only on the input at time t, but also on the previous history of the input. For a linear system this is written as a 'convolution':

$$o(t) = \int_{-\infty}^{\infty} g(t-\tau)h(\tau)\mathrm{d}\tau = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)\mathrm{d}\tau \equiv g \otimes h, \tag{22}$$

where $g(t)$ is the input and $h(t)$ describes how the filter responds to an impulse, called the impulse response function of the filter. The symbol $\otimes$ indicates convolution. The FT of Eq. (22) is:

$$O(\omega) = G(\omega)H(\omega) \tag{23}$$

where $H(\omega)$ is the FT of $h(t)$ and is called the transfer function of the filter. This is a very useful transformation because the filtering problem reduces to a multiplication (of generally complex functions). Another way to write this relationship is in terms of the modulus (or energy) and phase:

$$O(\omega) = |O(\omega)|\, \mathrm{e}^{i\Psi_O(\omega)} = |G(\omega)|.|H(\omega)|\, \mathrm{e}^{i[\Psi_G(\omega) + \Psi_H(\omega)]},$$

where we see that phases add and moduli or energy spectra multiply:

$$E_O(\omega) = E_G(\omega)E_H(\omega) \tag{24}$$

where $E_H(\omega)$ is the energy transfer function of the filter.

An obvious though not well-known property of the convolution theorem is that the zeroes in the FT of a transient are invariant to a linear filter. Should the input FT be zero at some frequency,ω_0, then the FT of the output must also be zero at ω_0 regardless of the transfer function (actually, as long as $H(\omega_0)$ is finite). Equally, if the FT of the output is zero at ω_0, then either the FT of the input or the transfer function (or both) must be zero at ω_0. Thus we see that the zeroes of a transient will be unaffected by a linear filter (provided of course they fall within the pass band of the filter). This is not surprising because clearly a linear filter cannot change the temporal separation of the start and end discontinuities.

In contrast, a linear filter will affect the order of the transient discontinuities. In control systems engineering it is usual to write the transfer function of a linear system as the ratio of two polynomials in Laplace notation:

$$H(s) = \frac{b_Z s^Z + b_{Z-1}s^{Z-1} + \cdots + b_2 s^2 + b_1 s + 1}{a_P s^P + a_{P-1}s^{P-1} + \cdots + a_2 s^2 + a_1 s + 1},$$

where the numerator has Z roots, called 'zeroes' (not to be confused with the zeroes of $G(\omega)$), and the denominator has P roots, called 'poles'. Thus the system acts like a filter, and by making the substitution of $s \to i\omega$, we see that at high frequencies the transfer function of the system is $H(\omega) \propto (i\omega)^{Z-P}$. Therefore the order of a discontinuity will be increased by $P-Z$, which we call the 'order' of the system. From this we see that a low-pass filter ($P > Z$) will place a lower limit on the order of a discontinuity at the output. This is important for transients, because recording equipment always has a finite bandwidth, and although the bandwidth may be sufficient to record the main features of a transient including the zeroes, it cannot record the discontinuities faithfully, and will increase the order of the discontinuity.

At first it may seem paradoxical that the minima are invariant to a linear filter, yet they are determined by the order of the transient. For example if we passed the symmetrical parabola with second-order discontinuities through a first-order filter, the output discontinuities would be third-order yet the minima would still occur at $f = (m + 1/2)/T$ instead of at m/T. The reason is that the output would no longer be symmetrical or anti-symmetrical. A symmetrical/anti-symmetrical output can only arise from a symmetrical/anti-symmetrical input if the filter has a symmetrical impulse response function, or equivalently the transfer function, $H(\omega)$, is a function of only even powers of ω.

3.3. Energy spectra minima

When $G(\omega)$ is zero, the energy spectrum $E_G(\omega)$ is also zero, so that zeroes will appear as minima in the energy spectrum. In practice, because of noise and measurement artifacts, these minima do not go to zero. However, sharp minima in the input energy spectrum will be transferred to the output energy spectrum with negligible change in frequency provided the filter transfer function is reasonably smooth at the minima frequency. This can be easily seen by taking the derivative of the energy spectrum with respect to frequency and setting it to zero for a minimum:

$$\frac{\mathrm{d}E_C}{\mathrm{d}\omega} = E_G \frac{\mathrm{d}E_H}{\mathrm{d}\omega} + E_H \frac{\mathrm{d}E_G}{\mathrm{d}\omega} = 0.$$

$$\text{If } \left|\frac{\mathrm{d}E_G}{\mathrm{d}\omega}\right| \gg \left|\frac{E_G}{E_H}\frac{\mathrm{d}E_H}{\mathrm{d}\omega}\right|$$

then the solution will be very close to the solution of $\mathrm{d}E_G/\mathrm{d}\omega = 0$. Finding these minima is practical, as we shall see in Section 5, but there is a limitation imposed by wide-band noise. From Eq. (19), we see that the energy of the transient minima will fall off at $1/\omega^{2n}$, so that eventually the minima will become buried in noise.

4. Discrete and fast Fourier transforms

The storage and analyses of data by digital computer means inevitably that the raw analog data must be sampled. Fourier analysis is then based on the sampled data, usually in the form of the discrete Fourier transform (DFT), of which the fast Fourier transform (FFT) is a commonly used algorithm. Although there is a self-consistent Fourier theory of discretely sampled signals, this does not in general correspond directly to the continuous case. This is important because we are usually concerned with interpreting continuous biological signals—the sampling is only a tool to get the data into the computer.

4.1. Sampling

Sampling a biological signal $g(t)$ creates a discrete infinitely long data set g_1, g_2,…g_n, where each value represents the instantaneous[2] value of $g(t)$ at a set of times t_1, t_2,…t_n. We shall only consider the typical situation when samples are made at a regular interval denoted by τ, or equivalently at a constant r samples/s ($r = 1/\tau$) (Fig. 4). We can denote mathematically the process of sampling as the process of multiplying the original function $g(t)$ by an infinitely long series or 'comb' of delta functions, III(t, τ). Thus, the new sampled function,, will be given by:

$$\tilde{g}(t) = g(t).\mathrm{III}(t, \tau).$$

Since the FT of a comb is also a comb, the FT of $\tilde{g}(t)$ is given by the convolution:

$$\tilde{G}(\omega) = \frac{2\pi}{\tau} G(\omega) \otimes \mathrm{III}\left(\omega, \frac{2\pi}{\tau}\right)$$

and obviously, the FT of the sampled function is not the same as the FT of the unsampled function. However, if $G(\omega)$ has no energy at frequencies greater than $\omega = \pi/\tau$, then we can recover $g(t)$ exactly by multiplying $\tilde{G}(\omega)$ by a low-pass filter that only passes frequencies below π/τ. This is the famous Sampling Theorem, which is described in detail by Bracewell (1986) and many other texts. It shows that by sampling a function at a rate that is at least twice the bandwidth of the function, then all the information in the original function is preserved and can be recovered. If, however, the sampling rate is less than this minimum, then distortion will occur (called aliasing).

[2] In reality sampling is not instantaneous but reflects a weighted average over some time due to the sampling function. This is equivalent to convolving $g(t)$ with the sampling function before instantaneous sampling. Thus, the sampled function will not be identical to $g(t)$.

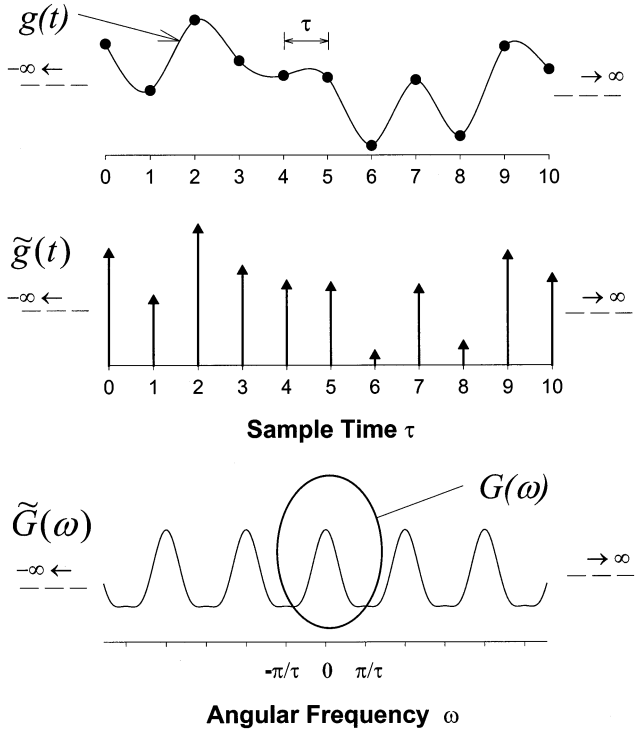

Fig. 4. Illustration of sampling. Top panel shows original infinitely long function, $g(t)$ which is sampled at discrete regular time intervals separated by τ. Middle panel shows the sampled function, $\tilde{g}(t)$, which can be considered to be $g(t)$ multiplied by an infinitely long comb of delta functions. The bottom panel shows $\tilde{G}(\omega)$ which is the Fourier transform of the sampled function $\tilde{g}(t)$. $\tilde{G}(\omega)$ is periodic where each cycle is $G(\omega)$, which is the true Fourier transform (circled) of the original function $g(t)$. $G(\omega)$ can be recovered from $\tilde{G}(\omega)$ by low-pass filtering, which then allows $g(t)$ to be recovered by the inverse Fourier transform. This is only accurate if $G(\omega)$ is bandwidth limited with no energy above twice the sampling rate (π/τ).

4.2. Sampling finite data sets

Thus, the digitization process $g(t) \Rightarrow \tilde{g}(t)$ can be reversed $g(t) \Leftarrow \tilde{g}(t)$ provided $g(t)$ is bandwidth-limited. In real experiments, our observation period cannot last forever, so we are constrained to finite data sets, which always have infinite bandwidth. Similarly, if the function is a transient, then it also has infinite bandwidth, and we would need to sample at an infinite rate to avoid distortion. This is impossible and we lose the reversible relationship $g(t) \Leftarrow \tilde{g}(t)$. This is no mystery because obviously we cannot resolve the start and end within a sampling period.

4.3. Discrete Fourier transform (DFT)

We assume that the time function is sampled over a finite time at $\tau = 0, 1, 2 \ldots N-1$, where τ now represents counts of intervals rather than time itself, and N is the total number of samples. We denote the DFT frequency by υ (to avoid confusion with continuous frequency f or ω). The discrete Fourier transform is usually defined as:

$$G(\upsilon) = \sum_{\tau=0}^{N-1} g(\tau)\, \mathrm{e}^{-i2\pi\upsilon\tau/N} = \sum_{\tau=0}^{N-1} g(\tau)\cos(2\pi\upsilon\tau/N) - i\sum_{\tau=0}^{N-1} g(\tau)\sin(2\pi\upsilon\tau/N). \quad (25a)$$

Since the cosines and sines are orthogonal, it can be shown that the original sampled function can be precisely recovered by the inverse DFT (DFT^{-1}):

$$g(\tau) = \frac{1}{N}\sum_{\upsilon=0}^{N-1} G(\upsilon)\, \mathrm{e}^{+i2\pi\upsilon\tau/N} = \frac{1}{N}\sum_{\upsilon=0}^{N-1} G(\upsilon)\ \cos(2\pi\upsilon\tau/N) + \frac{i}{N}\sum_{\upsilon=0}^{N-1} G(\upsilon)\sin(2\pi\upsilon\tau/N); \quad (25b)$$

(note that some authors scale Eq. (25a) by $1/N$ rather than Eq. (25b)). Thus, we can swap between $g(\tau)$ and $G(\upsilon)$ without loss or gain of information, and they constitute 'discrete Fourier pairs', parallel to the continuous case. Indeed, as in the continuous case there are operational relationships between discrete Fourier pairs similar to Table 1, although we do not show them here (Bracewell, 1986).

The fast Fourier transform (FFT) is a numerical method for finding the DFT. It is fast because it capitalizes on redundancies in computing the DFT when there are 2^n data samples. The technique was first described by Cooley and Tukey (1965) to compute the FS and it became an important technique in an era of relatively slow computational speed. Its origins as a DFT seemed to have become blurred and instead it has become virtually synonymous with the term Fourier transform, being now the most available technique for estimating Fourier transforms. However, the FFT is simply the DFT with 2^n points, and we will not dwell on it.

4.4. Discrete versus continuous Fourier transform

We now encounter a source of much confusion. Namely, our sampled function is not only finite in time (N time samples) but also finite in the frequency domain (N frequency samples). As stated above, a function cannot be both time- and bandwidth-limited, so is the DFT in error? The simple answer is that the DFT is not the FT of the unsampled function; only the sampled function $g(\tau)$ can be recovered by DFT^{-1}.

To reconcile the difference between the DFT and the FT, we ask what unsampled time function, $c(t)$, has a continuous FT, $C(\omega)$ that is identical to the DFT $G(\upsilon)$. Clearly the answer must be given by the inverse continuous Fourier transform of the DFT, i.e.: $c(t) = FT^{-1}\{G(\upsilon)\}$. At this point we recognize that the DFT is in reality simply a finite sum of exponential sinusoids at frequencies that are integral multiples of $1/N$. If we return to the exponential Fourier series and compare Eq. (25a) to Eq. (10), we see that the DFT (or FFT) is simply computing the coefficients of the Fourier series of the infinitely long periodic function $g(t)$, where each cycle is given by the $g(\tau)$—the sampled function is literally wrapped-around end-on-end forever.

4.5. Wrap-around

The effect of wrap-around is best illustrated by example. Consider the problem of finding the modulus of the FT of a step function, $u(t)$ (we have chosen the modulus rather than the energy spectrum simply because it is easier to display graphically). Our data set consists of 128 samples in which there is a step of unit amplitude midway (insert in Fig. 5A). Now we know analytically that the true FT of a step is $U(\omega) = 1/i\omega$ (Eq. (17)), and that the modulus spectrum $|U(\omega)| = 1/\omega$ must be independent of the time origin. Computing the FFT (using Matlab) reveals an oscillatory function in which even frequencies are zero, and odd frequencies are proportional to $1/\omega$ (Fig. 5A); reminiscent of, but not the same as $|U(\omega)|$. If we now shift the step forward in time so that it occurs at sample $n = 96$ (insert Fig. 5B), $|U(\omega)|$ should remain unchanged, but we see that the FFT has changed (Fig. 5B). How do we explain this quirkiness? In fact the FFT computations are not in error, rather it is our interpretation that is at fault. The FFT in Fig. 5A is the true FT of an infinitely long periodic function where each cycle lasts 128 samples given by the data set. In this case the function would be an infinitely long square wave with a 50% duty cycle (i.e. equal on and off periods), as shown in Fig. 5C. Shifting the step changes the duty cycle to 25% Fig. 5C), which has a different FS. It can be seen that the FFTs in Fig. 5A and 5B correspond to the modulus of the coefficients in the FS in Eqs. (5) and (6), respectively.

Many errors occur in Fourier analysis because of the failure to recognize that the FFT is the FT of the periodic function given by infinitely wrapping-around the data set—and not the FT of the function in the data window (we shall encounter an example of this error in the next section). Far fewer mistakes would be made if the FFT were more appropriately called the fast Fourier series!

4.6. Approximating the continuous FT

The DFT/FFT is not the only way to estimate the FT with digitized data. We can view the FT as simply a numerical integration problem, where the integrand is the product of the sampled $g(t)$ and a sampled sine/cosine. There are many numerical techniques for find-

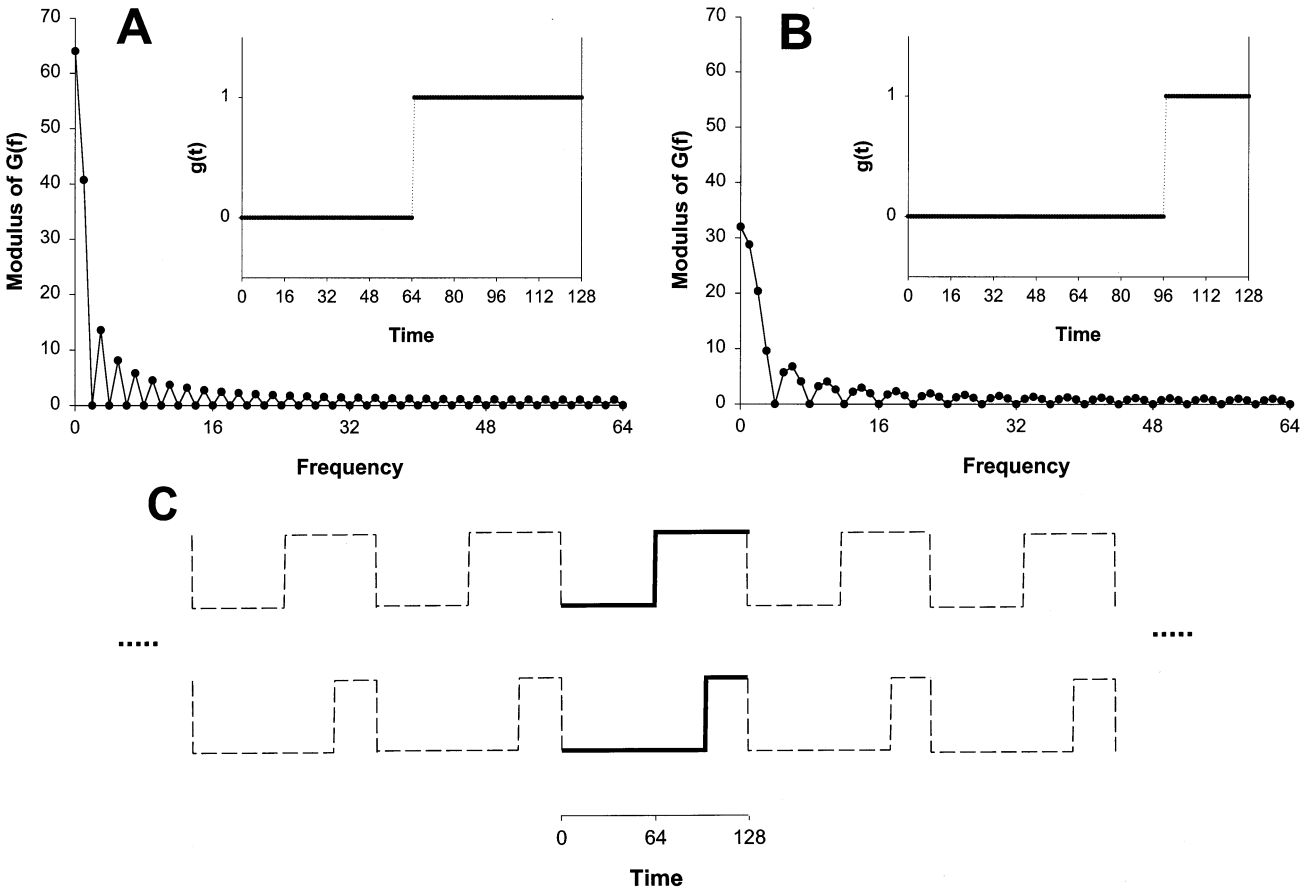

Fig. 5. Illustration of wrap-around in the FFT: (A) the modulus of the FFT of a step function midway in the data sequence (insert); (B) modulus of FFT of step function moved forward in time (insert). Note difference from (A). (C) The FFT computes the Fourier series of an infinitely long periodic function where each cycle is given by the data sequence (solid line) wrapped-around (dashed line).

ing the approximate integral of a function sampled at regular intervals (Press et al., 1992). The main issue is to avoid truncation errors. These can be greatly reduced if we have a priori information about how $g(t)$ behaves outside the sampling interval. In the next section we shall consider the case where $g(t)$ is known to be a constant outside the sampling interval. Even though the data are sampled, the transform should not be confused with a DFT, and we shall call it a continuous FT (CFT), albeit an approximation to the true FT.

5. Analysis of a transient: the saccade

In this section we examine the problem of finding the Fourier energy spectrum of a saccadic eye movement. A saccade is a very fast eye movement lasting some tens of milliseconds in which the eye is moved from one position to another. The saccade is a good example of a 'transient', especially since the start and end position are necessarily different.

Although the saccadic pathway is heavily parallel-distributed, it is often modeled as a lumped quasi-linear system. The transient movement is the output of the muscles behaving as a linear filter, called the 'ocular plant', which is driven by a pulsatile neural signal input, called the 'pulse'. However, there is no consensus on the shape of the saccadic pulse or the transfer function of the ocular plant. There is, nevertheless, some agreement on the output of the system since the eye movement can be recorded with specialised recording equipment. Fig. 6 shows a typical saccade recorded by an infra-red eye-tracker and sampled at $r = 1$ kHz ($\tau = 1$ ms).

5.1. An analytical model

We do not know what the FT of a saccade should be, so we introduce a model that is close to observed saccades in both the temporal domain and (as we shall see later) also in the frequency domain. It has the virtue of having a precisely known FT, so that we can test the performance of different numerical FT techniques. The fact that the model is not a perfect replica of a biological saccade is not important for our purposes.

The model is that the position of the saccade $s(t)$ at time, t, is given by a half of a cosine cycle (Fig. 7A):

$$s(t) = \begin{cases} 0 & t < 0 \\ \frac{1}{2}[A - A\cos(\pi t/T)] & 0 \le t \le T \\ A & t > T \end{cases} \quad (26)$$

which has a velocity profile (Fig. 7B) given by:

$$v(t) = \pi \begin{cases} \frac{\pi A}{2T}\sin(\pi t/T) & 0 \le t \le T \\ 0 & 0 > t > T \end{cases} \quad (27)$$

where A is the amplitude and T is the duration of the saccade. This symmetrical velocity profile approximates real saccades with an amplitude below about 20° (Collewijn et al., 1988), and was originally proposed by Yarbus (1967). The peak (maximum) velocity is $v_m = \pi A/2T$, and the ratio of peak velocity to mean velocity (amplitude/duration) is a constant $\pi/2(\sim 1.57)$, which is similar to the empirically observed value of about 1.65 (Becker, 1989) (and slightly better than the 1.50 for a parabolic velocity profile).

It can be shown that the energy spectrum of the positional profile is:

$$|S(\omega)|^2 = \left[\frac{\pi^2 A^2 \cos(\omega T/2)}{\omega(\pi^2 - \omega^2 T^2)}\right]^2, \quad (28)$$

which is shown by the solid line for $A = 10°$ and $T = 50$ ms in Fig. 8. For low frequencies $|S(\omega)|^2$ falls off linearly on a log–log plot with a slope of -2, which is to be expected since it represents the overall step component, where $|S(\omega)|^2 = A^2/\omega^2$ (shown by the dashed line in Fig. 8A). The fine structure at higher frequencies has maxima and sharp minima, as expected for a transient. It is this fine structure that conveys information about the trajectory shape. The energy falls off with a slope of -6 on a log–log plot (dotted line in Fig. 8A). This is to be expected because differentiating the position signal three times will lead to impulses at the beginning and end of the movement (Fig. 7D). The minima occur at non-harmonically related frequencies given by $f = 3/2T$, $5/2T$, $7/2T$, etc. which we shall denote as M1, M2, M3, etc. and can be best seen on a log–linear plot (Fig. 8B). Given this spectrum with its high-frequency structure, we now examine various numerical techniques to find it.

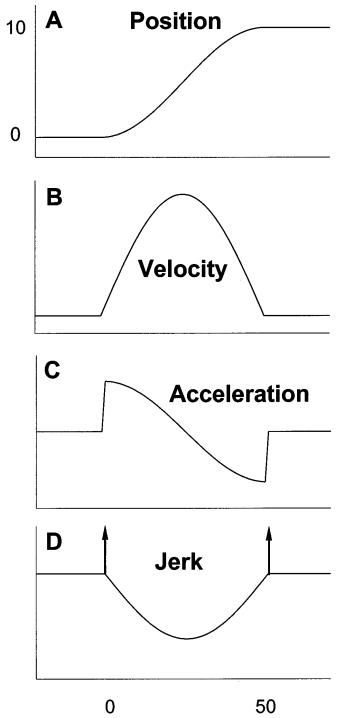

Fig. 7. A simple model of a saccade: (A) position of trajectory given by half cosine cycle (Eq. (26)); (B) velocity given by first derivative (Eq. (27)); (C) acceleration (2nd derivative); (D) jerk (3rd derivative). Note that jerk is discontinuous with impulses at the start and end of the movement, thus indicating third-order discontinuities (see text).

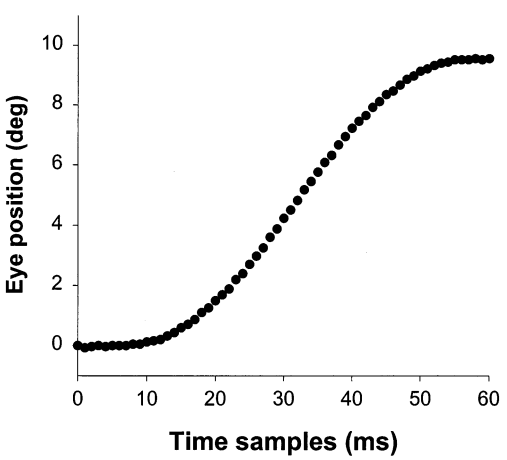

Fig. 6. Actual saccade of amplitude 9.5° and duration 50 ms in response to the onset of a visual target 10° to the right of fixation point. Saccade recorded with an infra-red eye tracker (Skalar) sampled at 1 kHz.

5.2. The standard approach to the FFT

First, consider the simple FFT. Throughout this section we will assume a sampling rate of 1000 samples/s and the data window of 128 samples, and use the Matlab (version 4) FFT routine without any scaling of the output. From our model, we will take a saccade of nominal amplitude of 10° and duration of 50 ms, which will have energy minima at 30, 50, 70 Hz etc. As can be seen from Fig. 9B, the FFT bears little resemblance to the true FT in Fig. 8B. Instead of minima at 30, 50, 70

Hz, there is a minimum at about 16 Hz. This is due to wrap-around, which will yield a quasi-rectangle wave with minima at even multiples of the reciprocal of the data sequence, which in this case is $2 \times 1/0.128 = 15.6$ Hz. One can also see that the frequency resolution is poor, being $1/0.128 = 7.8$ Hz.

5.3. An erroneous alternative

In an attempt to overcome truncation effects, van Opstal et al. (1985) added to the time signal a reflected version of the saccade to create a composite signal, $c(t)$, in which the signal returned to its starting position before the end of the data sequence (Fig. 9C). The implicit assumption was that the FT of this composite signal $C(\omega)$ would replicate the FT of the single saccade $S(\omega)$. However, this is neither true nor even approximate. This can be seen by recognizing that $c(t)$ can be considered to be the sum of a shifted saccade profile and a delayed version of the reflected saccade profile: $c(t) = s(t + \tau/2) + s(\tau/2 - t) - A$, where τ is the time difference between the peak velocities of the saccade and its reflection (Fig. 9C). Making use of the shift and scaling Fourier pair relationships in Table 2, the FT of the composite signal is $C(\omega) = S(\omega)\, e^{-i\omega\tau/2} + S(-\omega)\, e^{+i\omega\tau/2} - A\delta(\omega)$. To simplify (although this is not necessary), saccade position profiles tend to be anti-symmetric functions (i.e. symmetrical velocity profiles), so that $S(\omega) = -S(-\omega)$. Ignoring the delta function, the energy spectra of van Opstal et al.'s composite signal is $|C(\omega)|^2 = 4|S(\omega)|^2 \sin(\omega\tau)$, $(\omega \neq 0)$,and clearly $|C(\omega)|^2 \neq |S(\omega)|^2$. Not surprisingly therefore, the FFT of this composite function (Fig. 9D) is dominated by $\sin(\omega\tau)$ as seen by the rapid oscillations in the FFT, which is an artifact caused by the separation of the two 'saccades'.

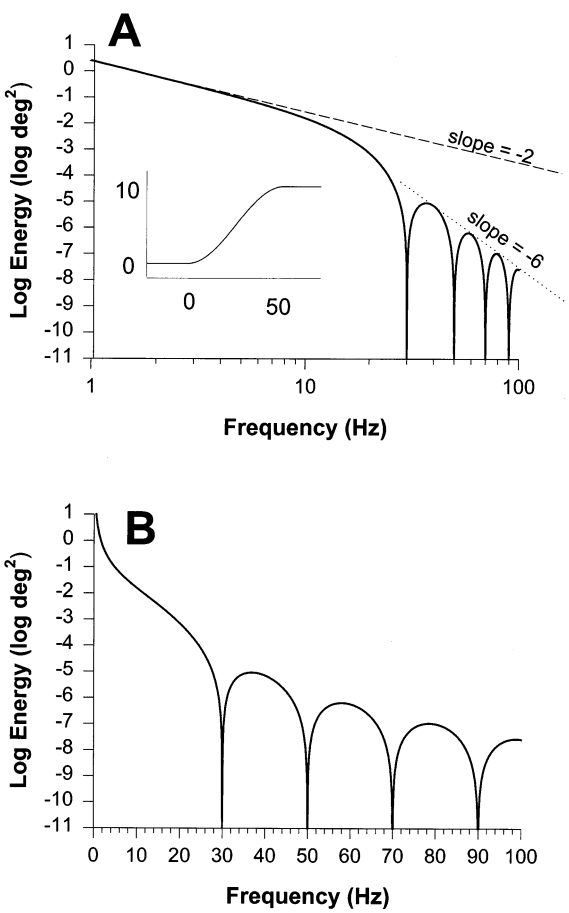

Fig. 8. The energy spectrum of the Fourier transform of the cosine saccade model found analytically (Eq. (28)). (A) Plotted in log–log coordinates for a 10° movement lasting 50 ms (insert). For low frequencies energy decreases linearly with a slope of -2 as expected from a step in position (dashed line). For higher frequencies the slope increases to -6 (dotted line) as expected for a trajectory with discontinuities first appearing in the third derivative (Fig. 7D). (B) Plotted in log–linear coordinates to show sharp minima occurring at the non-harmonic frequencies 30, 50, 70, and 90 Hz.

5.4. The FFT with a window

To overcome the sudden truncation effect caused by wrap-around, some investigators multiply the data set by a smooth function, called a 'window', to equalise the first and last points in the data sequence without an abrupt change. This inevitably introduces distortion, but mostly at low frequencies of the same order as the reciprocal of the duration of the data sequence. For those data sequences that are about filled by the transient, the window will, however, cause distortion at frequencies of interest. For example, when we multiply the data sequence by a cosine window (Fig. 9E), there is clearly distortion of the saccade, and the FFT (Fig. 9F) does not resemble the true FT (Fig. 8B).

5.5. Approximating the CFT with numerical integration

Instead of using a DFT/FFT, Harris et al. (1990) approximated the continuous FT by numerical integration. They added a step function of the appropriate amplitude starting at time D (Fig. 10A): $s(t) = \hat{s}(t) + Au(t - D)$. The FT of the saccade can be found by finding the FT of the data set and adding a correction factor given by the FT of the step, which is known analytically from the shift theorem applied to the FT of a step:

$$FT\{s(t)\} = FT\{\hat{s}(t)\} + FT\{Au(t - D)\} = FT\{\hat{s}(t)\} + \frac{A\, e^{-i\omega D}}{i\omega}. \quad (29)$$

where

$$FT\{\hat{s}(t)\} = \int_0^D s(t)\, e^{-i\omega t} dt.$$

Even though the data are sampled, $FT\{\hat{s}(t)\}$ is not a DFT and does not depend on the sampling theorem,

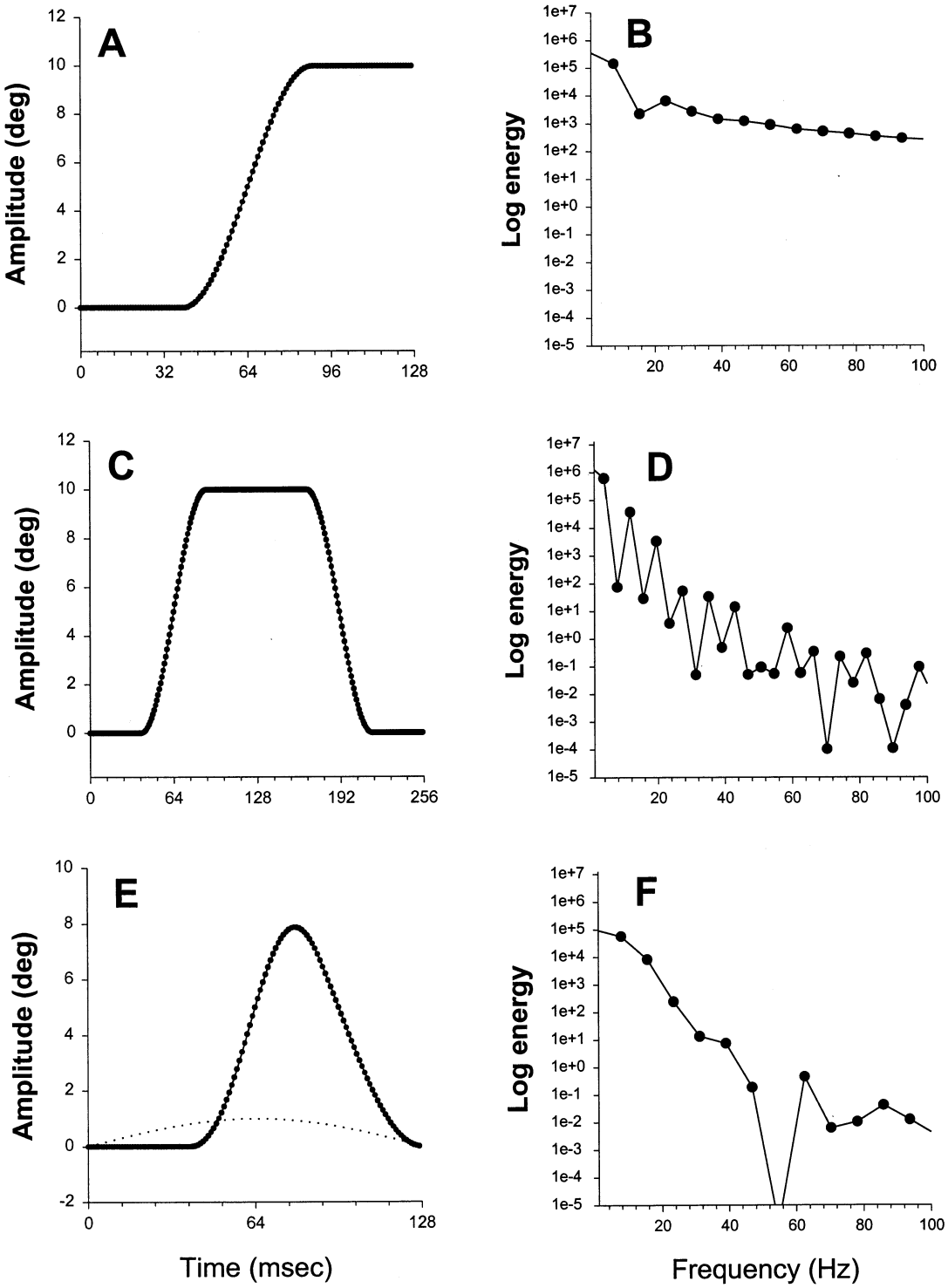

Fig. 9. Incorrect applications of the FFT on a model saccade lasting 50 samples at 1000 samples/s: (A) in the 'standard' approach the data are padded to power of two data points ($N = 128$) with sharp truncation at end of sequence; (B) a log–linear plot of the energy spectrum of the FFT of (A) showing little resemblance to the true FT in Fig. 8B; (C) a technique described by van Opstal et al. (1985) in which a reflected version of the saccade is added to eliminate the truncation effect (see text); (D) energy spectrum of the FFT of (C) showing artifact oscillations due to the separation of the two saccades; (E) as in (A) but multiplied by a cosine window (dotted curve) to smoothly eliminate the truncation; and (F) energy spectrum of FFT of (E) showing an artifact minimum.

but requires numerical integration to approximate the CFT. The final result depends, therefore, on the accuracy of the numerical integration. Using the model, Fig. 10B shows the results of three different numerical integration techniques: the standard extended trapezoidal rule (Eq. (30a)), which was used by Harris et al. (1990),

a composite Simpson/trapezoidal rule (Eq. (30b)) and a cubic polynomial rule (Eq. (30c)) described by Press et al. (1992)

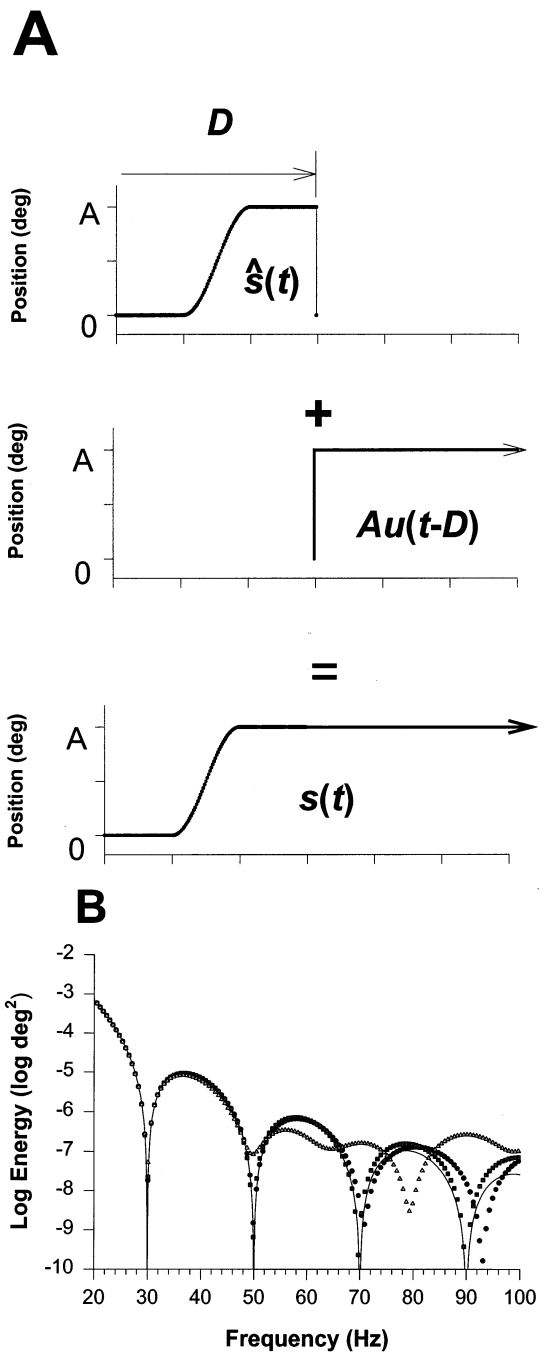

Fig. 10. Approximation to the continuous Fourier transform of the saccade model using numerical integration: (A) the FT is computed using numerical integration over the measured sample of length D, and corrected with the analytic FT of a shifted step (middle panel) to yield an estimate of the FT of the infinite function (bottom panel) (Eq. (29)); (B) FT showing log energy against linear frequency. Solid line shows analytic transform as plotted in Fig. 8B. Symbols show three different numerical techniques of integration: triangles, trapezoidal integration; circles, a composite Simpson/trapezoidal rule; squares, cubic polynomial rule (Eqs. (30a), (30b) and (30c)).

$$\int_{t_0}^{t_N} f(t)\mathrm{d}t \approx \Delta\left[\frac{1}{2}f_1 + f_2 + f_3 + f_4 + \ldots + f_{N-3} + f_{N-2} + f_{N-1} + \frac{1}{2}f_N\right] \quad (30a)$$

$$\int_{t_0}^{t_N} f(t)\mathrm{d}t \approx \Delta\left[\frac{5}{12}f_1 + \frac{13}{12}f_2 + f_3 + f_4 + \ldots + f_{N-3} + f_{N-2} + \frac{13}{12}f_{N-1} + \frac{5}{12}f_N\right] \quad (30b)$$

$$\int_{t_0}^{t_N} f(t)\mathrm{d}t \approx \Delta\left[\frac{3}{8}f_1 + \frac{7}{6}f_2 + \frac{23}{24}f_3 + f_4 + \ldots + f_{N-3} + \frac{23}{24}f_{N-2} + \frac{7}{6}f_{N-1} + \frac{3}{8}f_N\right] \quad (30c)$$

All three fit the true FT very well for low frequencies up to, and including M1. However, the pure trapezoidal rule became progressively inaccurate with higher frequencies. The composite rule found M2 accurately but failed M3 and M4. The cubic polynomial rule was clearly the best, finding M3 accurately and M4 also quite well. Presumably higher order techniques would be more accurate, but we have not tried any.

5.6. The FFT with padding and windowing

To use the FFT, the above suggests that we should 'pad' the data set to some large length (to an integer power of 2 for the FFT), where the initial point is extended back in time and/or the last point is extended into the future (Fig. 11A). However, by itself this does not improve the situation because there is still a truncation, and as we might have expected, the FFT now oscillates at a much higher frequency determined by the reciprocal of the new data length (Fig. 11B).

Applying a window to the padded data should eliminate the oscillations at the expense of increased distortion at frequencies of the reciprocal of the data length. If the padded data length is very long, window distortion should only effect very low frequencies, which are of little interest in this context. As shown in Fig. 11C, the application of a cosine window profoundly changes the FFT, which now (at last!) is similar to the theoretical FT of the model (Fig. 11D).

Padding and windowing can produce good approximations to FTs, but the frequency resolution is limited by the number of points. In our example with a sampling rate of 1 kHz, finding a minimum in the FFT to

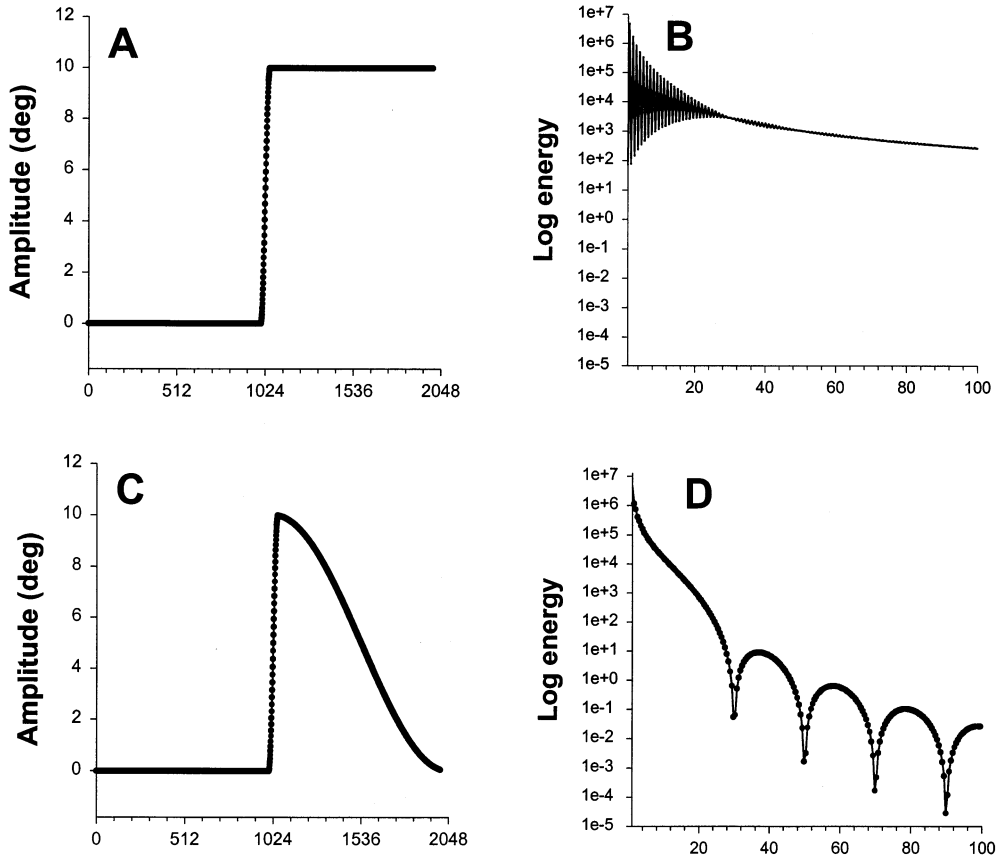

Fig. 11. FFT with padding and window on a model saccade lasting 50 samples at 1000 samples/s: (A) data sequence is padded out to a length of 2048 points; (B) Energy spectrum of (A) showing high frequency artifact oscillations determined by the reciprocal of the high data length; (C) padded as in (A) but multiplied by a cosine window; and (D) energy spectrum of (C) showing close resemblance to the true FT in Fig. 8B with minima at 30, 50, 70, and 90 Hz.

a resolution of about 0.1 Hz requires padding to a total data length of 8192 samples (to the nearest power of 2), yet the original data length may have been only 100 samples or less!

5.7. The FFT of the derivative (velocity)

The FT of $g(t)$ and the FT of the derivative $g'(t)$ are trivially related by the multiplication/division by $i\omega$ (Eq. (14)). If we take the derivative of the saccade, the initial and final velocities will both be zero, so that in the FFT truncation effects should not occur. If we were genuinely interested in the FT of the position profile we could then divide the FT by $i\omega$.

Fig. 12A shows the derivative of the saccade model (velocity profile) computed by taking the running difference of the position samples and padding to 128 points. As expected the frequency resolution of the FFT is coarse but does reveal minima at approximately the correct frequencies (Fig. 12B). If we now pad to 1024 points, the FFT looks very much like the true FT. Greater resolution can be achieved by padding to a greater length.

Some caution is needed when computing derivatives from sampled data because it is easy to introduce artifacts into the spectrum. The running difference between two samples introduces a minimum in the energy spectrum at multiples of half the reciprocal of the separation of the samples. In our example this is not a problem because at 1 kHz sampling rate the first artifact minimum would be at 500 Hz which is not of interest to us. However, other digital filters may not be so innocuous.

5.8. The FT of real saccades

We now consider the energy spectrum of a real saccade to illustrate the potential value of Fourier analysis of behavioural transients. The saccade was made by a human subject viewing a small red laser spot that jumped abruptly from primary position 10° to the right, and was recorded by an infra-red eye tracker. The saccade (Fig. 6) had an amplitude of 9.5°, and a duration of 50 ms, that is $1/T = 20$ Hz. The energy spectrum was computed using the FFT of the position signal padded to 4096 points with a cosine window (Fig. 13).

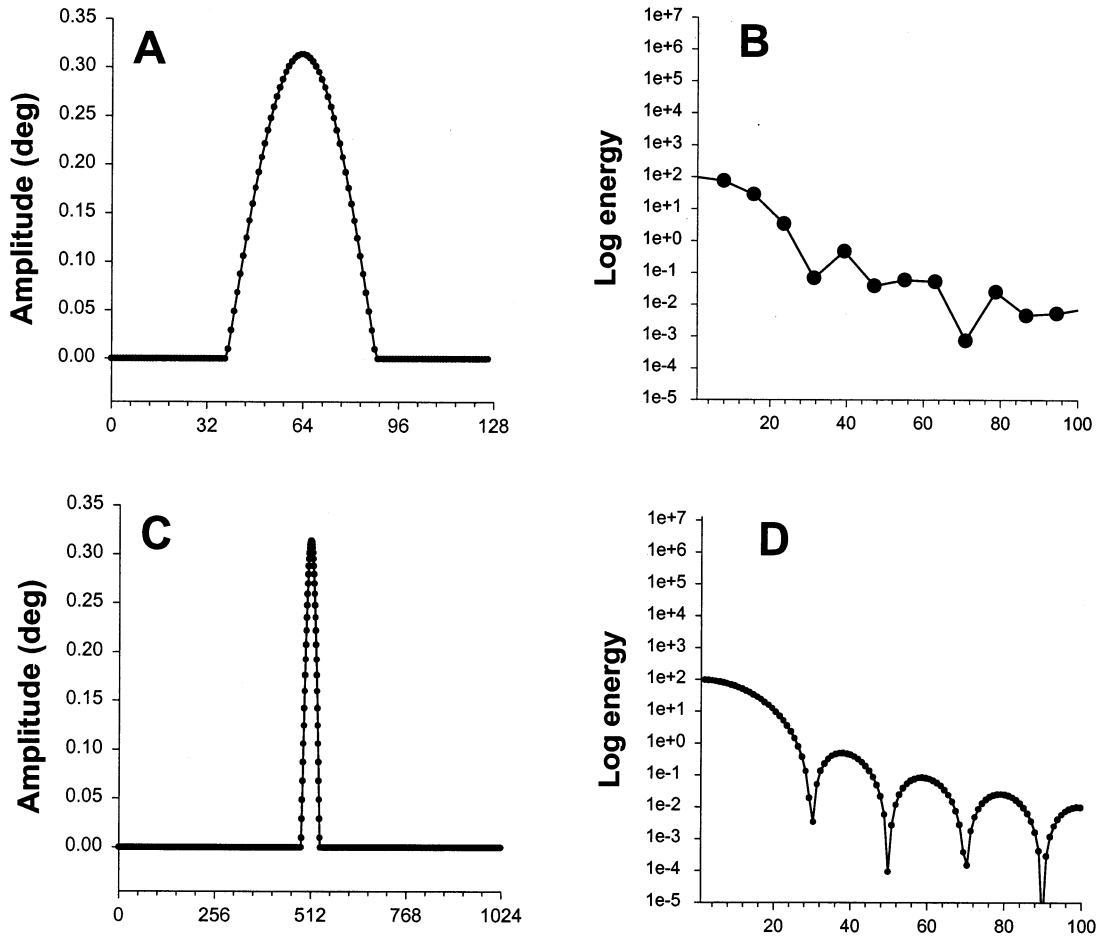

Fig. 12. FFT of derivative (velocity profile) of a model saccade lasting 50 samples at 1000 samples/s: (A) data sample padded to 128 points and derivative computed by the running difference of consecutive samples; (B) energy spectrum of (A) showing crude resemblance to the true energy spectrum in Fig. 8B; (C) data sample padded to 1024 points and then derivative computed by taking the running difference of consecutive points; (D) energy spectrum of (C) showing minima at 30, 50, 70 and 90 Hz as expected from the true FT in Fig. 8B. Note that energy spectrum of position profile can be obtained by dividing this energy spectrum by ω^2 ($\omega > 0$).

The energy spectrum exhibits a linear fall-off on a log–log plot with a slope of -2 for low frequencies, as expected from the overall step in the movement. At higher frequencies fine structure appears with minima occurring at non-harmonic frequencies 31, 49, and 71 Hz, which are clearly close to the pattern of $f = (m + 1/2)/T$ (rather than m/T). Although these minima do not uniquely identify the shape of the transient trajectory of the saccade, they do allow us to place some interesting constraints that any model of saccade generation would need to adhere to.

Let us assume that the system is linear. Our first deduction is that regardless of the transfer function of the eye muscles, the driving saccadic pulse must be a transient with the same minima, and hence cannot be any function that would have minima at, say, m/T. This immediately rules out the possibility that saccadic pulse is a rectangle (as well as all other symmetric or anti-symmetric profiles with odd-ordered discontinuities). This is significant because the saccade has often been modeled with a rectangular pulse, partly for convenience, but also for theoretical reasons since a rectangular pulse may be optimal (with certain assumptions) for reaching the final position in the minimum time. Thus, the FT of a saccade has important implications for optimal control, and is discussed elsewhere (Harris, 1998).

We also assume that the saccade is anti-symmetric, that is, the velocity profile is symmetric, which is plausible for brief saccades. Since the eye muscles are known to be viscous, the transfer function of the muscles must have a ω term and cannot be just a function of even powers of ω (i.e. 1, ω^2, ω^4, etc.). Therefore the driving signal cannot be symmetric (another reason for ruling out a rectangle pulse).

The order of the start discontinuity of the saccade is given by the order of the start discontinuity of the driving pulse plus the overall order of the plant. Since the saccade position discontinuities are odd-ordered, this is sufficient to rule out various models. For exam-

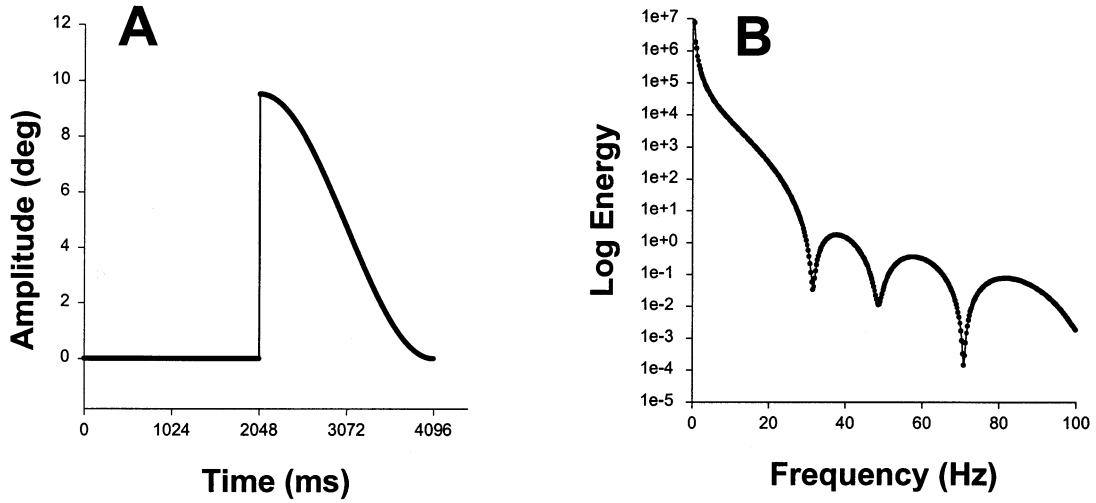

Fig. 13. Fourier transform of actual saccade: (A) saccade, from Fig. 6, padded to 4096 points with a cosine window; and (B) energy of FFT of (A) showing the first three minima at 31, 49, and 71 Hz.

ple, a step change (first-order) in the onset of the driving signal and a transfer function with two poles and one zero (1st-order) is even-ordered and incorrect since it will have minima at $f = m/T$ rather than $(m + 1/2)/T$.

The FT of a saccade also has descriptive value. Saccade trajectories are smooth and rather featureless. Typically they are described by their duration and peak velocity (Bahill et al., 1975), which are determined by only three points in the trajectory. In contrast, the energy spectrum of the FT of the trajectory depends on the shape of the whole trajectory and usually has sharp minima that can be easily found. Unlike peak velocity, the frequencies of these minima are independent of any amplitude scale (calibration) factor. Moreover, since the minima are found in the energy spectrum, they are independent of the time-origin and it is not necessary to find the start and end of the saccade, only to ensure that the whole trajectory is embraced by the data sequence. (In practice it is best to keep the data sequence as close to the duration of the saccade as feasible since longer data sets only add unwanted noise.) Thus, the FT may provide useful alternative descriptors of saccades, provided of course, that the FT is calculated correctly. In principle this approach could be applied to other smooth transients, such as arm movements.

6. Discussion

We have focused on the Fourier analysis of transients. Because transients are brief and not usually periodic, their FT's do not have the same intuitive appeal as those of long periodic signals. Nevertheless, their FTs do have features that are potentially interesting to the neuroscientist provided artifacts due to numerical algorithms can be avoided. A transient, defined here as a signal that has a discontinuous start and end, has a FT that oscillates at high frequencies with zeroes, or minima in the energy spectra. The frequencies of these zeroes reflect the shape of the transient but are particularly sensitive to the duration of the transient at high frequencies. Although the zeroes do not uniquely identify the profile of the transient, they are essentially invariant to linear transformation, and consequently provide constraints on linear models of the transient.

The digitization of data is the means to modern analysis but it is not an end in itself because at some time we will be required to make some description or inference of the actual continuous underlying biological signal. The mapping of a continuous biological signal by sampling into a discrete signal may seem innocuous, but the DFT (or the FFT; its most common implementation) yields a FT that is not so readily mapped back to the continuous FT. We have illustrated this by applying various Fourier analyses to a transient that has a similar profile to an actual biological behaviour, but has a precisely known theoretical Fourier transform. This provided differing results, where some approaches yielded FTs that were clearly incorrect, whereas others including the FFT with padding and windowing, the CFT with numerical integration, and the FFT of the derivative provided good approximations to the true FT. This clearly illustrates the need for care in applying Fourier analysis to biological data, since simply invoking a standard FFT may yield completely erroneous transforms. In particular, with the FFT, transient data will usually need preconditioning (padding, windowing, differentiation) which may not be available (or easily implementable) in 'off-the-shelf' software or hardware.

It should be recognized that the FFT is not synonymous with the FT, but actually provides an estimate of the Fourier *series* of a periodic function in which the data sequence represents exactly one cycle. This wrap-around can cause considerable confusion, especially in transients where the amplitude of the last point in the data sequence differs from the first point, because this provides an unwanted step in the periodic signal. With padding and windowing or differentiation, the FT can be recovered with the FFT technique, but one may wonder whether there is anything to be gained over simple numerical integration. The original reason for the FFT was to overcome the slowness of computer processing in the 1960s and 1970s. Today's machines are now so fast that speed is of little concern, except when the amount of data to be analysed is vast, or time is very critical as in real-time processing.

Whichever approach is taken, and whether or not the analysis uses off-the-shelf or in-house software, we cannot emphasize too strongly the need to apply the Fourier analysis first to functions with known Fourier transforms to check that the analysis yields sensible results.

Acknowledgements

I thank Dr Mary Faldon and Mark Harwood for their help in prepaing this manuscript. This work has been supported by The Medical Research Council, grant G9316292N, and the charities Iris Fund, Help Child to See and Child Health Research Appeal Trust.

References

Bahill AT, Clark MR, Stark L. The main sequence, a tool for studying human eye movements. Math Biosci 1975;24:191–204.

Becker W. Metrics. In: Wurtz RH, Goldberg ME, editors. The Neurobiology of Saccadic Eye Movements. Amsterdam: Elsevier, 1989:13–67.

Bracewell RN. The Fourier Transform and its Application. New York: McGraw-Hill, 1986.

Champeney DC. Fourier Transforms and their Physical Applications. London: Academic Press, 1973.

Collewijn H, Erkelens CJ, Steinman RM. Binocular coordination of human horizontal saccadic eye movements. J Physiol (London) 1988;404:157–82.

Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier Series. Math Comput 1965;19:297–301.

Harris CM. On the optimal control of behaviour: a stochastic perspective. J Neurosci Methods 1998;83:73–88.

Harris CM, Wallman J, Scudder C. Fourier analysis of primate saccades. J Neurophysiol 1990;63:877–86.

Hogan N. An organizing principle for a class of voluntary movements. J Neurosci 1984;4:2745–54.

Percival DB, Walden AT. Spectral Analysis for Physical Applications. Cambridge: Cambridge University Press, 1993.

Press WH, Teukolsky SA, Vetterling WT, Flannary BP. Numerical recipes, 2nd. New York: Cambridge University Press, 1992.

Slepian D. Some comments on Fourier analysis, uncertainty, and modeling. SIAM Rev 1983;25:379–393.

van Opstal AJ, van Gisbergen JAM, Eggermont JJ. Reconstruction of neural control signals for saccades based on an inverse method. Vis Res 1985;25:789–901.

Yarbus AL. Eye Movements and Vision. New York: Plenum, 1967.