Harmonic and musical wavelets

By David E. Newland

Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, U.K.

The concept of a harmonic wavelet is generalized to describe a family of mixed wavelets with the structure

$$w_{m,n}(x) = \{\exp(in2\pi x) - \exp(im2\pi x)\}/i(n-m)2\pi x.$$

It is shown that this family provides a complete set of orthogonal basis functions for signal analysis. By choosing the (real) numbers m and n (not necessarily integers) appropriately, wavelets whose frequency content ascends according to the musical scale can be generated. These musical wavelets provide greater frequency discrimination than is possible with harmonic wavelets whose frequency interval is always an octave. An example of the wavelet analysis of music illustrates possible applications.

1. Introduction

Harmonic wavelets (Newland 1993*b*) are functions from the orthogonal family

$$w(2^j x-k) = \{\exp\{i4\pi(2^j x-k)\} - \exp\{i2\pi(2^j x-k)\}\}/i2\pi(2^j x-k), \tag{1}$$

where integers j and k define the scale (or compression) of the wavelet and its position on the x-axis. The Fourier transform of $w(2^j x-k)$ is identically zero except in an octave band of positive frequencies from $2^j(2\pi)$ to $2^{j+1}(2\pi)$ within which its modulus is constant and of magnitude $1/2^j(2\pi)$. It has been shown (Newland 1993*b*) that this family provides a complete set of basis functions for expanding an arbitrary function $f(x)$ which is square-integrable on the real line x.

The concept can be extended by defining a harmonic wavelet whose Fourier transform is confined to a frequency band which is not necessarily an octave. Suppose that the Fourier transform $W_{m,n}(\omega)$ is identically zero outside the band $m(2\pi)$ to $n(2\pi)$ (where m and n are real and positive but not necessarily integer) and within this band has magnitude $1/(n-m)(2\pi)$ (see figure 1*a*). If

$$W_{m,n}(\omega) = \begin{cases} 1/(n-m)(2\pi) & \text{for} \quad m(2\pi) \leqslant \omega < n(2\pi), \\ 0 & \text{elsewhere,} \end{cases} \tag{2}$$

the corresponding wavelet function is

$$w_{m,n}(x) = \{\exp(in2\pi x) - \exp(im2\pi x)\}/i2\pi(n-m)x. \tag{3}$$

To allow for translation of the wavelets by step $k/(n-m)$, where we shall show later that k must be an integer, (3) becomes

$$w_{m,n}\left(x-\frac{k}{n-m}\right) = \left\{\exp\left(in2\pi\left(x-\frac{k}{n-m}\right)\right) - \exp\left(im2\pi\left(x-\frac{k}{n-m}\right)\right)\right\} \Big/ i2\pi(n-m)\left(x-\frac{k}{n-m}\right), \tag{4}$$

Proc. R. Soc. Lond. A (1994) **444**, 605–620
Printed in Great Britain

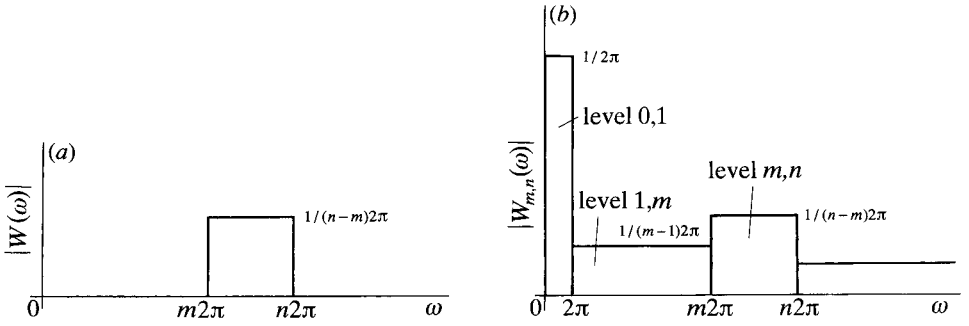

Figure 1. Modulus of the Fourier transform of general harmonic wavelets (*a*) for level m, n only, (*b*) for adjacent levels covering the full frequency range.

where now

$$W_{m,n}(\omega) = \begin{cases} \{1/(n-m)\,2\pi\} \exp\{-\mathrm{i}\omega k/(n-m)\} & \text{for} \quad m(2\pi) \leqslant \omega < n(2\pi), \\ 0 \quad \text{elsewhere.} \end{cases} \tag{5}$$

It can be seen that (4) reduces to (1) when $m = 2^j$ and $n = 2^{j+1}$. We can no longer talk about wavelet *level* j and instead we shall use the terminology *level* m, n to denote a wavelet in the frequency band $m2\pi$ to $n2\pi$ where $n > m$. To form a complete set of wavelets, adjacent wavelet levels must have Fourier transforms whose frequency bands touch each other, so that all values of ω along the line 0 to infinity are included (see figure 1*b*).

In this paper we begin by examining the properties of the general harmonic wavelet $w_{m,n}(x)$ defined by (4) to prove that the family of such wavelets forms a complete, orthogonal set of basis functions for signal analysis.

2. General harmonic wavelets

Because of their band-limited structure in the frequency domain, with no overlapping bands (see figure 1*b*), it is easy to prove that wavelets in different frequency bands (i.e. wavelets of different levels) are orthogonal to each other. The proof depends on the results from Fourier analysis theory that

$$\int_{-\infty}^{\infty} w(x)\, v(x)\, \mathrm{d}x = 2\pi \int_{-\infty}^{\infty} W(\omega)\, V(-\omega)\, \mathrm{d}\omega, \tag{6}$$

$$\int_{-\infty}^{\infty} w(x)\, \bar{v}(x)\, \mathrm{d}x = 2\pi \int_{-\infty}^{\infty} W(\omega)\, \bar{V}(\omega)\, \mathrm{d}\omega, \tag{7}$$

and is explained in detail in the earlier paper (Newland 1993*b*).

In the case of wavelets within the same frequency band (wavelets of the same level), we need to prove that

$$\int_{-\infty}^{\infty} w_{m,n}\left(x - \frac{k_1}{n-m}\right) w_{m,n}\left(x - \frac{k_2}{n-m}\right) \mathrm{d}x = 0 \tag{8}$$

and
$$\int_{-\infty}^{\infty} w_{m,n}\left(x-\frac{k_1}{n-m}\right)\bar{w}_{m,n}\left(x-\frac{k_2}{n-m}\right)\mathrm{d}x = 0, \quad k_1 \neq k_2. \tag{9}$$

On account of (6), we see at once that (8) is zero whatever the values of k_1 and k_2. Since both wavelets are in the same wavelet level m,n, if $W_{m,n}(\omega)$ is in the non-zero frequency band $m2\pi \leqslant \omega < n2\pi$, $W_{m,n}(-\omega)$ must be zero at the same frequency. This is because all the Fourier transforms of harmonic wavelets are identically zero for $\omega < 0$ (see Newland 1993*b*). We shall now show that (9) is also true provided that k_1 and k_2 are two different integers.

On using (7) and substituting for $W_{m,n}(\omega)$ from (5), we find that the left-hand side of (9) is

$$\begin{aligned}
&\int_{-\infty}^{\infty} w_{m,n}\left(x-\frac{k_1}{n-m}\right)\bar{w}_{m,n}\left(x-\frac{k_2}{n-m}\right)\mathrm{d}x \\
&\quad = \frac{(1/2\pi)}{(n-m)^2}\int_{m\,2\pi}^{n\,2\pi}\exp\left\{\frac{-\mathrm{i}\omega(k_1-k_2)}{(n-m)}\right\}\mathrm{d}\omega \\
&\quad = \left(\frac{1}{2\pi}\right)\left\{\frac{-1}{\mathrm{i}(n-m)(k_1-k_2)}\right\}\left\{\exp\left(\frac{-\mathrm{i}2\pi n(k_1-k_2)}{(n-m)}\right)-\exp\left(\frac{-\mathrm{i}2\pi m(k_1-k_2)}{(n-m)}\right)\right\} \\
&\quad = \left(\frac{1}{2\pi}\right)\left\{\frac{-1}{\mathrm{i}(n-m)(k_1-k_2)}\right\}\exp\left(\frac{-\mathrm{i}2\pi m(k_1-k_2)}{(n-m)}\right)\{\exp(-\mathrm{i}2\pi(k_1-k_2))-1\}.
\end{aligned} \tag{10}$$

This is zero when k_1-k_2 is any integer number (not zero). Hence any two translated wavelets at level m,n satisfy (9). For the case when $k_1 = k_2$, by taking the limit when $k_1-k_2 = \epsilon$ where $\epsilon \to 0$, we find that

$$\int_{-\infty}^{\infty}\left|w_{m,n}\left(x-\frac{k}{n-m}\right)\right|^2\mathrm{d}x = 1/(n-m). \tag{11}$$

The conclusion about orthogonality is that any two general harmonic wavelets are orthogonal if they are in different wavelet levels. When they are wavelets in the same level m,n, they are still orthogonal if they are translated with respect to each other by any interval $k/(n-m)$, where k is an integer.

Because of these results, the general wavelet expansion formula (36) of Newland (1993*b*)

$$f(x) = \sum_{k=-\infty}^{\infty}\{a_{\phi,k}\,\phi(x-k)+\tilde{a}_{\phi,k}\,\bar{\phi}(x-k)\}+\sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty}\{a_{j,k}\,w(2^j x-k)+\tilde{a}_{j,k}\,\bar{w}(2^j x-k)\} \tag{12}$$

becomes

$$\begin{aligned}
f(x) = &\sum_{k=-\infty}^{\infty}\{a_{\phi,k}\,\phi(x-k)+\tilde{a}_{\phi,k}\,\bar{\phi}(x-k)\} \\
&+\sum_{m,n}\sum_{k=-\infty}^{\infty}\left\{a_{m,n,k}\,w_{m,n}\left(x-\frac{k}{n-m}\right)+\tilde{a}_{m,n,k}\,\bar{w}_{m,n}\left(x-\frac{k}{n-m}\right)\right\},
\end{aligned} \tag{13}$$

where it is understood that the summation over m,n is to take account of all the pairs

of values of m and n needed to completely cover the real line $0 \leqslant \omega \leqslant \infty$ according to figure 1*b*. The complex wavelet coefficients are defined by

$$a_{m,n,k} = (n-m)\int_{-\infty}^{\infty} f(x)\,\bar{w}_{m,n}\left(x-\frac{k}{n-m}\right)\mathrm{d}x, \tag{14}$$

$$\tilde{a}_{m,n,k} = (n-m)\int_{-\infty}^{\infty} f(x)\,w_{m,n}\left(x-\frac{k}{n-m}\right)\mathrm{d}x, \tag{15}$$

and, as in the previous paper,

$$a_{\phi,k} = \int_{-\infty}^{\infty} f(x)\,\bar{\phi}(x-k)\,\mathrm{d}x, \tag{16}$$

$$\tilde{a}_{\phi,k} = \int_{-\infty}^{\infty} f(x)\,\phi(x-k)\,\mathrm{d}x, \tag{17}$$

where

$$\phi(x) = \{\exp(\mathrm{i}2\pi x)-1\}/\mathrm{i}2\pi x. \tag{18}$$

An heuristic proof that the family of wavelets (4) provides a complete set of basis functions for expanding an arbitrary function $f(x)$ is given in the Appendix. It is based on the idea that any arbitrary function can be decomposed into harmonic functions by Fourier analysis. Because it can be shown that any harmonic function (of arbitrary but constant frequency, phase and amplitude) can be modelled exactly by a train of such wavelets (see the Appendix), it follows that any arbitrary function can be represented exactly by a general harmonic wavelet series. Alternatively, a formal proof of the proposition can be made by showing that

$$\sum_{k=-\infty}^{\infty}\{|a_{\phi,k}|^2+|\tilde{a}_{\phi,k}|^2\}+\sum_{m,n}\sum_{k=-\infty}^{\infty}1/(n-m)\{|a_{m,n,k}|^2+|\tilde{a}_{m,n,k}|^2\} = \int_{-\infty}^{\infty}|f(x)|^2\,\mathrm{d}x \tag{19}$$

without using the expansion formula (13). This can be done by following through the calculation in Newland (1993*b*, Appendix C), with appropriate changes, and leads to the results that

$$\sum_{k=-\infty}^{\infty}|a_{m,n,k}|^2 = 2\pi(n-m)\int_{m\,2\pi}^{n\,2\pi}|F(\omega)|^2\,\mathrm{d}\omega, \tag{20}$$

$$\sum_{k=-\infty}^{\infty}|\tilde{a}_{m,n,k}|^2 = 2\pi(n-m)\int_{m\,2\pi}^{n\,2\pi}|F(-\omega)|^2\,\mathrm{d}\omega, \tag{21}$$

$$\sum_{k=-\infty}^{\infty}|a_{\phi,k}|^2 = 2\pi\int_{0}^{2\pi}|F(\omega)|^2\,\mathrm{d}\omega, \tag{22}$$

$$\sum_{k=-\infty}^{\infty}|\tilde{a}_{\phi,k}|^2 = 2\pi\int_{0}^{2\pi}|F(-\omega)|^2\,\mathrm{d}\omega, \tag{23}$$

where $F(\omega)$ is the Fourier transform of $f(\omega)$ which is assumed to be square-integrable on x. Because

$$2\pi\int_{-\infty}^{\infty}|F(\omega)|^2\,\mathrm{d}\omega = \int_{-\infty}^{\infty}|f(x)|^2\,\mathrm{d}x \tag{24}$$

the result (19), which is an expression of Parseval's theorem, follows immediately.

The terminology 'scaling function' for the function $\phi(x)$ defined by (18) has been

carried over from the earlier literature, but in the general analysis used here it is no longer necessary. Instead $a_{\phi,k}$ can be replaced by $a_{0,1,k}$, and similarly for $\tilde{a}_{\phi,k}$, so that the expansion formula (13) simplifies to

$$f(x) = \sum_{m,n} \sum_{k=-\infty}^{\infty} \left\{ a_{m,n,k}\, w_{m,n}\left(x - \frac{k}{n-m}\right) + \tilde{a}_{m,n,k}\, \bar{w}_{m,n}\left(x - \frac{k}{n-m}\right) \right\}, \tag{25}$$

where the summation over the pairs m, n must begin with a pair for which $m = 0$ and continue with touching (but not overlapping) pairs, for example $0, m$; m, n; n, p; p, q, etc., where $0 < m < n < p < q$ until sufficiently compressed wavelets have been included to ensure that all the fine detail of $f(x)$ has been accurately represented.

3. Musical wavelets

Because the values for the pair of numbers m, n in (4) can be chosen arbitrarily, there is considerable scope for assembling a family of wavelets whose properties are specially appropriate for a particular problem. In signal compression applications (see, for example, the commentary in Strang (1993)), the wavelet expansion (25) is truncated and the values of $a_{m,n,k}$ and $\tilde{a}_{m,n,k}$ are approximated for higher-order terms (i.e. for m, n large). In such cases it may be desirable to combine octave blocks to achieve maximum localization of detailed features of a signal before approximating the wavelet coefficients. Alternatively, in signal analysis problems, when it is necessary to study the fine structure of a signal by examining how its frequency characteristics change with time, the author has found that the octave bandwidth of harmonic wavelets (and the corresponding octave steps of orthogonal dilation wavelets) may be too wide to give adequate frequency discrimination. For such problems it is desirable to subdivide the octave and this can be done conveniently for the general harmonic wavelet described in this paper. By analogy with the musical scale, the author has chosen twelve steps per octave corresponding to the frequency steps of 'equal temperament' (see, for example, Jeans 1937). For example, taking the frequency of middle c to be 256 Hz and that of c′ in the octave above to be 512 Hz†, table 1 gives the frequencies of the equal temperament musical scale from c to c′. Each note has a frequency that is $2^{1/12}$ times that of the note below. A *musical wavelet* is defined as a wavelet whose Fourier transform has the band-limited form of equation (5) where $m = 2^{r/12}$ and $n = 2^{(r+1)/12}$ with r being a real integer.

Figure 2 shows the real and imaginary parts of such a musical wavelet for the case when $r = 96$ so that $m = 256$ and $n = 271.2$. At the scale drawn in figure 2, the real and imaginary parts of the wavelet are almost indistinguishable from each other except that the real part is an even function and has value unity at $x = 0$ while the imaginary part is an odd function which is zero at $x = 0$. When the fine structure is omitted by plotting the magnitude of the wavelet, the picture in figure 3 emerges. Figure 3*a* is for the same wavelet as figure 2, plotted from equation (4) with $m = 256$, $n = 271.2$, $k = 0$. The horizontal scale of x is extended from ± 0.2 in figure 2 to ± 0.5 in figure 3. Figure 3*b* shows $|w_{m,n,k}|$ plotted for $m = 256$, $n = 271.2$ and $k = 0$ (solid line), $k = 1$ (dot dashed line), $k = 2$ (dashed line). Incrementing k steps the wavelet along the x-axis but otherwise leaves it unchanged.

† For notes tuned to concert pitch, it is customary to tune a to 440 Hz, in which case middle c is 261.6 rather than 256 Hz. In the analysis here, it is convenient to put middle c at 256 Hz for two reasons. First, it falls into the octave sequence 1, 2, 4, 8, 16, etc., which occurs in the discrete wavelet transform. Second, the musical wavelet ascribed to middle c covers the frequency band 256 to 271.2 Hz and therefore straddles 261.6 Hz.

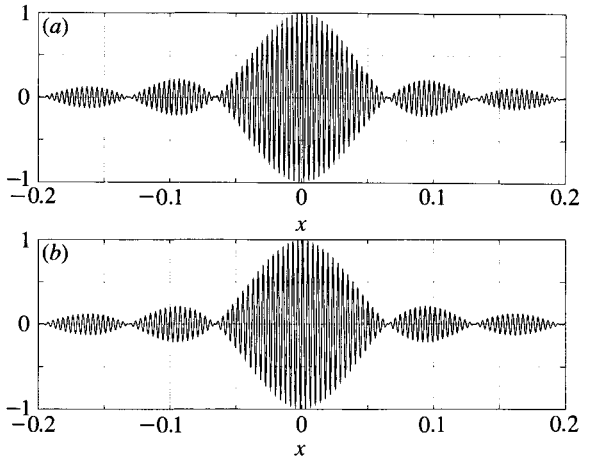

Figure 2. (*a*) Real part and (*b*) imaginary part of the harmonic wavelet for which $m = 256$, $n = 271.2$ (covering a semitone on the musical scale).

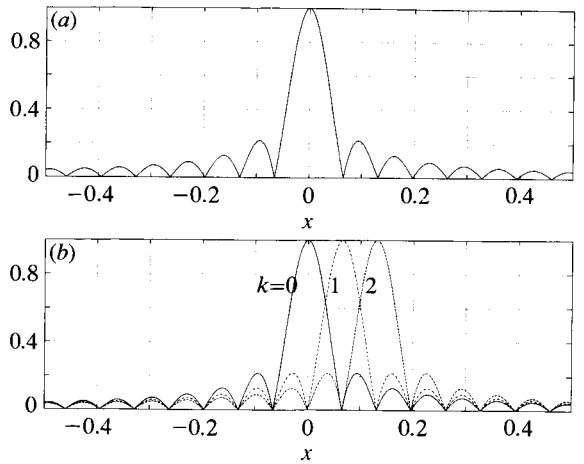

Figure 3. (*a*) Magnitude of the harmonic wavelet for which $m = 256$, $n = 271.2$; (*b*) three of the same wavelets as (*a*), each translated one step with respect to the next.

Table 1. *Fundamental frequencies of notes in the octave starting at middle* c *on the musical scale*

note	frequency/Hz	note	frequency/Hz
c	256	f#	$256(2^{6/12}) = 362.04$
c#	$256(2^{1/12}) = 271.22$	g	$256(2^{7/12}) = 383.57$
d	$256(2^{2/12}) = 287.35$	g#	$256(2^{8/12}) = 406.37$
d#	$256(2^{3/12}) = 304.44$	a	$256(2^{9/12}) = 430.54$
e	$256(2^{4/12}) = 322.54$	a#	$256(2^{10/12}) = 456.14$
f	$256(2^{5/12}) = 341.72$	b	$256(2^{11/12}) = 483.26$
		c′	$256(2) = 512$

To illustrate the collection of musical wavelets that together cover an octave band of frequencies, figure 4 shows the real parts of these plotted out; the names of corresponding musical notes are shown alongside. The relation between a wavelet that encompasses a band of frequencies and a continuous note of a single frequency

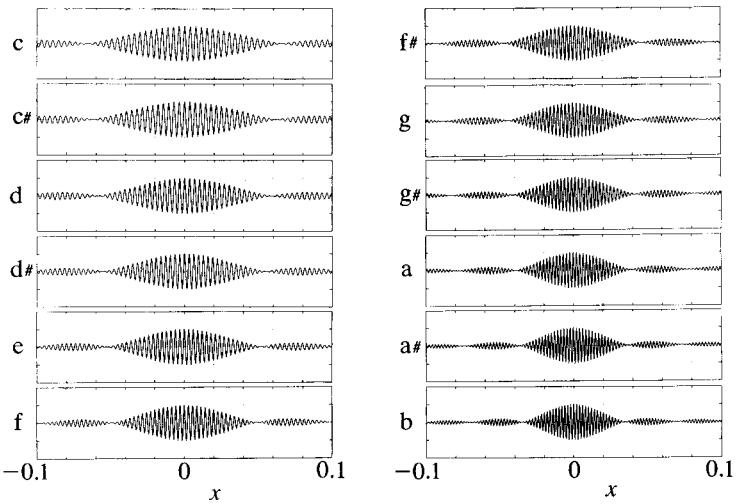

Figure 4. Real parts of 12 harmonic wavelets each of which has a frequency spread of one semitone; they are identified by the names of corresponding musical notes.

is of course only approximate. In practice, the notes of a musical instrument consist of a fundamental frequency and overlying harmonics (which for some instruments are more powerful than the fundamental). An example is given later.

For the case of a signal produced by summing a continuous succession of the same musical wavelets all of the same amplitude and each translated one step with respect to the next, we show in the Appendix that the resulting signal is always a pure tone. The frequency of this tone lies within the semitone frequency band of the wavelet. Its precise value depends on the phasing of the wavelets. By increasing the phase change from one wavelet in the train to the next, keeping a constant increment in phase from wavelet to wavelet, the frequency is progressively increased from $m2\pi$ to $n2\pi$ (see the Appendix).

4. Discrete algorithm

Musical wavelet analysis can be carried out by the discrete algorithm described in my previous paper (Newland 1993*b*). With reference to fig. 6 of that paper, the calculation is the same except that, instead of computing the inverse fast Fourier transform (IFFT) of octave blocks of the frequency coefficients F, the frequency coefficients are divided into smaller groups to correspond (approximately) to semitone blocks, with 12 semitone blocks per octave.

It is necessary to make integer approximations for the limits of each semitone frequency band so as to determine the extent of the semitone blocks in the discrete case. For example for the octave considered in table 1, the corresponding assignment of frequency coefficients to blocks is as listed in table 2. The error on frequency of the lower limit of each semitone block is shown in the right-hand column.

Subject to this different partitioning of the frequency coefficients before completing the second set of discrete transformations in figs 6 and 7 of the earlier paper, the discrete musical wavelet transform is the same as the discrete harmonic wavelet transform described previously. It retains the property of being a parallel process of calculation rather than the series calculation that is necessary when dilation wavelets are used (Mallat 1989; Newland 1993*c*). A diagrammatic drawing to illustrate the discrete musical wavelet transform is given in figure 5. This represents part of the

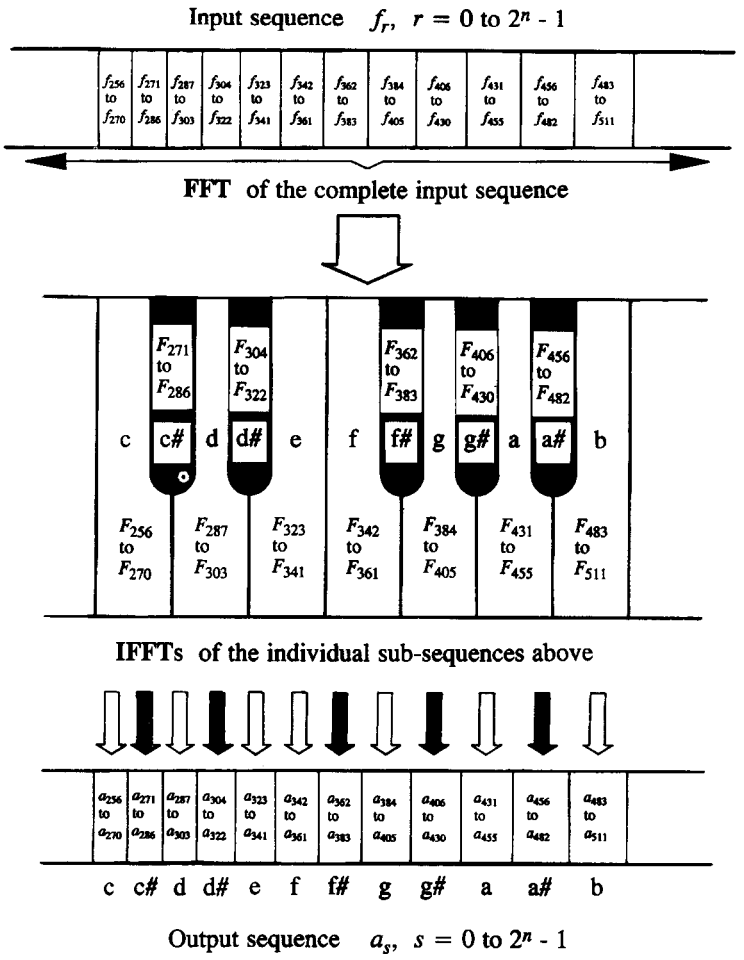

Figure 5. Discrete algorithm to compute the musical wavelet transform, based on the strategy shown in figs 6 and 7 of Newland (1993*b*).

Table 2. *Errors involved in taking integer approximations for the semitone blocks in the octave 256 to 512 Hz*

note	frequency/Hz	integer block	error on lower limit (%)
c	256	256–270	0
c#	271.22	271–286	0.08
d	287.35	287–303	0.12
d#	304.44	304–322	0.14
e	322.54	323–341	−0.14
f	341.72	342–361	−0.08
f#	362.04	362–383	0.01
g	383.57	384–405	−0.11
g#	406.37	406–430	0.09
a	430.54	431–455	−0.11
a#	456.14	456–482	0.03
b	483.26	483–511	0.05
c′	512		

left-hand side of the diagrams in figs 6 and 7 of the previous paper (Newland 1993*b*). To subdivide an octave, there must be a sufficient number of discrete elements in that octave. If a sequence $f_r, r = 0$ to $2^n - 1$, is generated by sampling at equally-spaced intervals over a record length of unity, the Fourier coefficients F_m generated by the FFT have frequencies $0, 2\pi, 4\pi, 6\pi$, etc. Therefore the frequency of coefficients doubles as m doubles in the series $m = 1, 2, 4, 8, 16$, etc. In the musical wavelet algorithm used by the author, the first octave to be divided into semitones is the octave which begins at $m = 32$. This octave has 32 elements and each semitone has a minimum of two elements. For Fourier coefficients F_m with $m < 32$, the elements are transformed in octaves rather than in semitones.

Although the frequency boundaries of discrete musical wavelets must be approximated, the discrete musical wavelet transform remains an exact transform. By reversing the calculation, beginning with the output sequence, the original input sequence is regained exactly (subject to rounding errors within the computer). This is because, according to (4) and (5), the frequency limits m, n can be chosen arbitrarily without prejudicing the orthogonality properties of the mixed wavelet family that results. Therefore, although the discrete implementation of the musical wavelet transform can only model musical wavelets approximately, it still yields an output sequence from which the input sequence can be regained exactly by the inverse transform. Since the main aim of the musical wavelet transform is to achieve greater frequency discrimination than is possible with the octave band structure of the harmonic wavelet transform, exactly how this is done is not important. Achieving the necessary frequency discrimination by subdividing each octave into blocks which approximate semitone blocks (rather than modelling the semitones exactly) is a satisfactory way of meeting the aim.

5. An application

Figure 6 shows the analysis of a short length of 'music'. A sequence of six pure tones at frequencies of 256, 322.5, 256, 383.6, 322.5 and 512 Hz, all of unit amplitude, has been generated artificially. The sequence length is $N = 2^{16} = 65536$ and the sampling time interval $\Delta = 6.1 \times 10^{-5}$ s, so that the record lasts for $N\Delta = 4$ s. The first five notes are crotchets, each lasting for 1/2 s, the last note is a dotted minim lasting for 3/2 s, and there is cosine smoothing at the junction between each note, the duration of the smoothing being 0.04 s in each case. This provides an artificial sequence of pure tones representing the notes c e c g e c′ which is the beginning of the tune Darwall's 148th (usually played for the hymn 'Ye holy angels bright').

Figure 6*b* is the musical wavelet map of this signal. This is a map plotted according to the rules described in my previous publications (Newland 1993*a*, *c*, 1994*a*, *b*) and summarized in the companion paper (Newland 1993*b*). On a base of frequency against time, the map shows contours of equal height for a surface that describes the distribution of a signal's mean-square value over frequency and time. The volume under the surface is devised to be proportional to the signal's mean-square value. In the diagram shown here, only a portion of the complete map is drawn, omitting the bottom of the map where there is insufficient discrimination to divide the octaves into semitones. Therefore semitones 1 to 12 fall in the octave defined by $m = 32$, $n = 63$, semitones 13 to 24 fall in the octave defined by $m = 64$, $n = 127$ and so on. The top row of the map is not shown on figure 6*b*. This is because the (constant) height of the top row of a mean-square wavelet surface depends only on the Nyquist

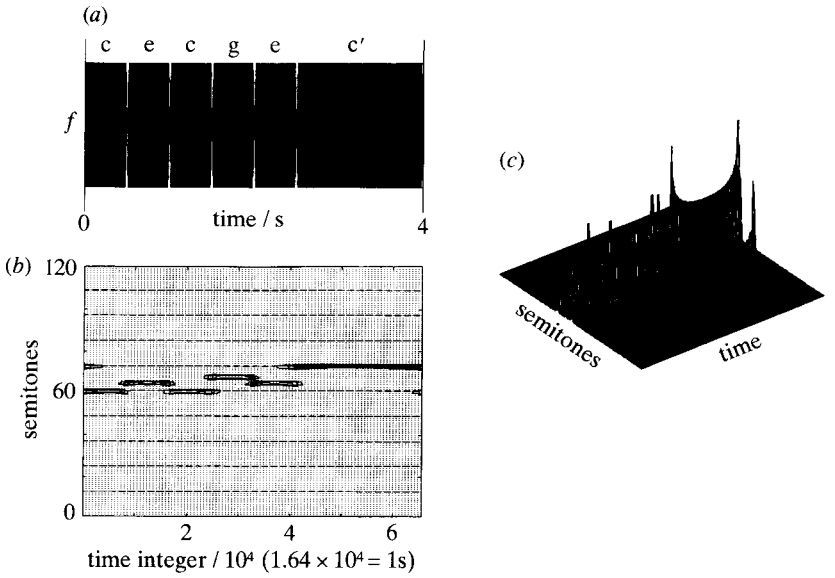

Figure 6. Signal consisting of six pure tones which form the melody line of the tune Darwall's 148th: (*a*) synthetic signal; (*b*) musical wavelet map of this signal; (*c*) mesh diagram of the map. Six contour levels are plotted in (*b*) at 0.1, 0.2, 0.4, 0.6, 0.8 and 0.95 of the maximum level on the map.

term $a_{N/2}$ (Newland 1993*b*). Provided that the sampling interval is small enough, so that the whole frequency spectrum of a signal can be represented fully, this term will always be zero to a close approximation (as it is in this example).

On figure 6*b*, the dashed lines are octave markers and the dotted lines designate semitones. Six contour lines are drawn, with their heights chosen to be 0.1, 0.2, 0.4, 0.6, 0.8 and 0.95 of the height of the highest peak on the map. Although each note is of constant amplitude, it is evident that the wavelet map does not represent a sequence of peaks of constant height. Figure 6*c* is a three-dimensional mesh diagram of the underlying mean-square surface and this shows how each note is represented by a peak at its beginning and ending and a saddle in between. This shape arises because, in general, the discrete wavelet expansion of a harmonic signal gives wavelet coefficients of different amplitudes if either (i) the harmonic does not have a discrete frequency (a multiple of 2π), or (ii) it does not have constant amplitude. These properties are examined in the Appendix. The upshot is that local peaks occur when a note does not have one of the preferred discrete frequencies or when it is turned on or off. They arise because of the process of cancellation that takes place between individual terms in a wavelet series expansion. This is also why, in figure 6*c*, there are low ridges between notes. These ridges do not show in figure 6*b* because their height is less than the height of the lowest contour on the map.

To identify frequencies on the semitone axis of a musical wavelet map computed by the discrete transform, for a record length of 1 s ($0 \leqslant x < 1$), these increase in the series $m, m+1, m+2, m+3, \ldots$ (in Hz). The baseline of the map is for a semitone with the integers $m = 32$, $m+1 = 33$, so that this line includes these two discrete frequencies measured in Hz. The next semitone should begin at $32 \times 2^{(1/12)} = 33.9$ and so its first discrete frequency is 34 Hz. It covers 34 and 35 Hz. The first octave marker is at $m = 64$ and the semitone at this marker covers $m = 64$, $m+1 = 65$,

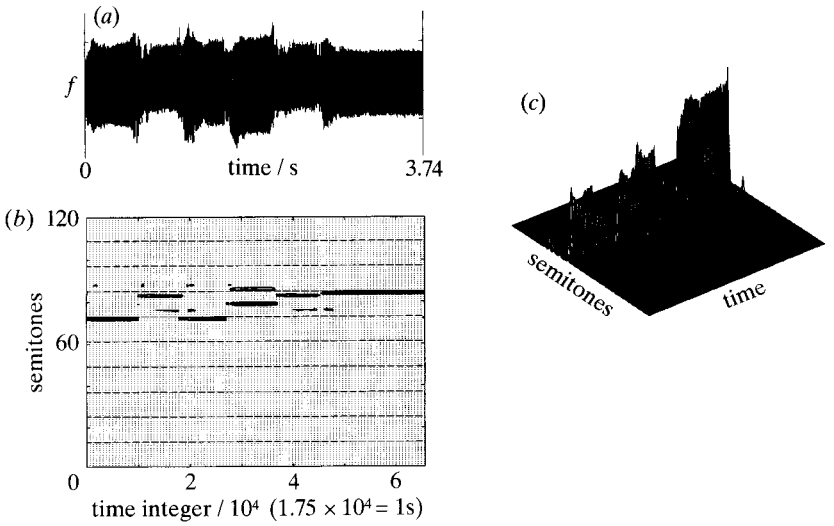

Figure 7. The first six notes of the melody line of the tune Darwall's 148th, played on the oboe stop of a pipe organ: (*a*) recorded signal; (*b*) musical wavelet map of this signal; (*c*) mesh diagram of the map. Six contour lines are plotted in (*b*) at 0.1, 0.2, 0.4, 0.6, 0.8 and 0.95 of the maximum level on the map.

$m+2 = 66$ and $m+3 = 67$, so that it includes the discrete frequencies 64, 65, 66 and 67 Hz. The semitone at the fifth octave marker above the baseline covers 1024 to 1084 Hz (since $1024 \times 2^{(1/12)} = 1084.9$). When, as in the example shown in figure 6, the record length is 4 s, all the frequencies are divided by 4, so that the semitone at octave marker 5 (semitone 61) then covers the frequencies 256, 256.25, 256.5, ..., 271 Hz.

Figure 7 shows the results of a similar analysis but carried out now for a digital recording of the same sequence of six notes, played on the oboe stop of a pipe organ. The signal analysed is shown in figure 7*a* and is represented by a sequence of length $N = 2^{16} = 65536$ and sampling interval $\Delta = 5.7 \times 10^{-5}$ s, so that $N\Delta = 3.74$ s. Because this is slightly different from the 4 s record length in figure 6, the frequencies of the semitone markers are slightly different. For example, the octave marker at semitone 61 has a lowest frequency of $1024/3.74 = 273.8$ Hz (256 Hz in figure 6) and that at semitone 73 has a lowest frequency of $2048/3.74 = 547.6$ (512 Hz in figure 6).

By comparing the wavelet map in figure 6 with that in figure 7, we see that the patterns are similar but approximately an octave higher in figure 7. This is because the oboe notes are rich in harmonics with the second and higher harmonics more pronounced than the fundamental. A typical frequency spectrum for middle c is shown in a previous paper (Newland 1993*a*), where it can be seen that the strength of the second harmonic is an order of magnitude greater than the strength of the fundamental. When middle c is played, the wavelet map registers maximum energy (maximum contribution to the mean-square of the recorded signal) at the frequency of the second harmonic, which is note c′. The second and third harmonics of notes e and g are prominent, as is the second harmonic of c′.

When contours are plotted according to a logarithmic scale, a more detailed picture emerges, as shown in figure 8. Now the contributions of the fundamental tones of each note show. There is a close similarity between the map in figure 6 (for a synthetic signal) and the corresponding portion of the map in figure 8 (which is for

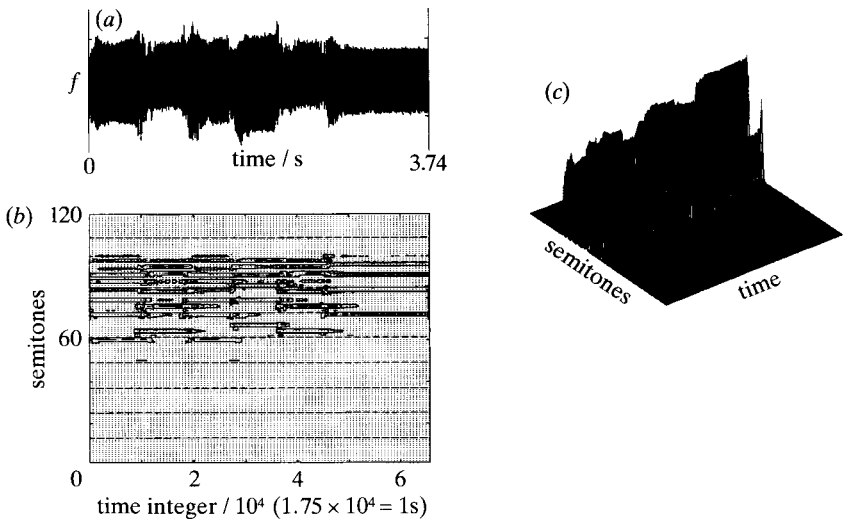

Figure 8. The same diagrams as in figure 7 except that in (*b*) the contours are plotted at 0.001, 0.01, 0.1, and 0.95 of the maximum level on the map and that in (*c*) the vertical axis is plotted to a log scale instead of a linear scale. In (*c*) the graph is not plotted for values less than 0.001 of the maximum.

a real signal). The second and third harmonics can be distinguished clearly on figure 8*b*. Higher harmonics than these begin to lose their identity along the time axis because the map cannot distinguish between different harmonics that have similar frequencies. For example, the fifth harmonic of c has a frequency of 5(256) = 1280 Hz which is almost the same frequency as that of the fourth harmonic of e which is 4(322.5) = 1290 Hz. This merging together of the higher harmonics of notes is characteristic of instruments like the oboe (or its pipe organ imitation) which have rich tonal quality with each note sounding many harmonics simultaneously.

A more detailed study of the application of wavelet maps is given in a separate paper (Newland 1994*b*, part II). This considers the relation of wavelet maps to sonograms, which are constructed by computing many short-time Fourier transforms of windowed segments of the input data. Both wavelet maps and sonograms show similar information; their main difference is that the wavelet map is computed by a single transformation of the complete data sequence (whereas the sonogram requires many separate transforms of different segments of the input data) and is an optimum display in the sense that there is no redundancy: the minimum number of data points is used.

6. Conclusion

The extension of the concept of harmonic wavelets described in this paper allows different families of mixed wavelets to be devised. To improve their frequency discrimination, one strategy is to subdivide octaves into semitones to match the musical scale. This allows the composition of sounds to be presented on a map which has close similarities to the musical stave. Figures 6, 7 and 8 show an example. Conversely, if octaves are combined instead of divided, the localization of detailed features of a signal in the physical domain can be improved.

A disadvantage of musical wavelets (as for harmonic wavelets generally) is that they are not compact in the x-domain and decay in proportion only to $1/x$. A

procedure for overcoming this disadvantage will be explored in a forthcoming paper (Newland 1994*c*). It involves combining harmonic wavelets in a way which preserves their orthogonal properties while improving their localization. A major advantage for all harmonic wavelets is that they can be computed by an efficient parallel algorithm rather than by the series algorithm necessary for the dilation wavelet transform. This saving in computation time may turn out to be the over-riding factor in determining the extent to which harmonic wavelets and their derivatives, musical wavelets, find general application.

Appendix. Properties of a train of general harmonic wavelets

(*a*) *Continuous functions*

Consider a train of wavelets at level m, n, each of which has unit amplitude and each of which is displaced one place with respect to the next. If the train is described by the real function $f(x)$, then, from (25),

$$f(x) = \sum_{k=-\infty}^{\infty} \left\{ a_k w_{m,n}\left(x - \frac{k}{n-m}\right) + \bar{a}_k \bar{w}_{m,n}\left(x - \frac{k}{n-m}\right)\right\}, \tag{A 1}$$

where a_k replaces $a_{m,n,k}$ and, from (14) and (15), $\tilde{a}_k = \bar{a}_k$ since $f(x)$ is real. By taking Fourier transforms and substituting from (5),

$$F(\omega) = \begin{cases} \displaystyle\sum_{k=-\infty}^{\infty} \left\{ \frac{a_k}{(n-m)\,2\pi} \exp\left\{\frac{-\mathrm{i}\,\omega k}{(n-m)}\right\} + \frac{\bar{a}_k}{(n-m)\,2\pi} \exp\left\{\frac{-\mathrm{i}\omega k}{(n-m)}\right\}\right\} \\ \qquad\qquad \text{for} \quad m2\pi \leqslant \omega < n2\pi \\ 0 \quad \text{elsewhere.} \end{cases} \tag{A 2}$$

The minus sign in the exponent on the right-hand side of (A 2) arises because if $W(\omega)$ is the Fourier transform of $w(x)$, then $\bar{W}(-\omega)$ is the Fourier transform of $\bar{w}(x)$.

The unit amplitude of the wavelets, a_k, may be constant or it may rotate in the complex plane. For reasons that become apparent below, we shall put

$$a_k = \exp(\mathrm{i}\theta k) \tag{A 3}$$

so that there is progressive rotation by equal increments θ when the integer k advances from one wavelet to the next. In that case, (A 2) becomes

$$F(\omega) = \begin{cases} \displaystyle\sum_{k=-\infty}^{\infty} \left\{ \frac{1}{(n-m)\,2\pi} \exp\left\{-\mathrm{i}\left(\frac{\omega}{n-m} - \theta\right)k\right\} + \frac{1}{(n-m)\,2\pi} \exp\left\{-\mathrm{i}\left(\frac{\omega}{n-m} + \theta\right)k\right\}\right\} \\ \qquad\qquad \text{for} \quad m2\pi \leqslant \omega < n2\pi \\ 0 \quad \text{elsewhere.} \end{cases} \tag{A 4}$$

According to Poisson's summation formula,

$$\sum_{k=-\infty}^{\infty} \exp\left\{-\mathrm{i}\left(\frac{\omega}{n-m} - \theta\right)k\right\} = 2\pi \sum_{j=-\infty}^{\infty} \delta\left(\frac{\omega}{n-m} - \theta - 2\pi j\right) \tag{A 5}$$

and so, substituting in the right-hand side of (A 4),

$$F(\omega) = \begin{cases} 1/(n-m) \displaystyle\sum_{j=-\infty}^{\infty} \left\{\delta\left(\frac{\omega}{n-m} - \theta - 2\pi j\right) + \delta\left(\frac{\omega}{n-m} + \theta + 2\pi j\right)\right\} \\ \qquad\qquad \text{for} \quad m2\pi \leqslant \omega < n2\pi \\ 0 \quad \text{elsewhere.} \end{cases} \tag{A 6}$$

It follows that $F(\omega)$ is zero everywhere except where $\omega = \pm 2\pi(n-m)(j+\theta/2\pi)$ subject to $m2\pi \leqslant \omega < n2\pi$. Therefore the values of $j+\theta/2\pi$ for which $F(\omega)$ is non-zero are given by

$$m \leqslant (n-m)\,|j+\theta/2\pi| < n \tag{A 7}$$

which may be written as

$$\frac{m}{n-m} \leqslant |j+\theta/2\pi| < 1+\frac{m}{n-m}. \tag{A 8}$$

In the case of musical wavelets

$$n = m2^{(1/12)}, \tag{A 9}$$

so that (A 8) gives

$$16.82 \leqslant |j+\theta/2\pi| < 1+16.82. \tag{A 10}$$

One solution is $j = 17$ with $\theta/2\pi$ lying in the range $-0.18 \leqslant \theta/2\pi < 0.82$ (so that angle θ can have any value in a 2π range). There is a corresponding solution with $j = -17$. If $|j|$ has a different value, θ also differs (by a multiple of 2π) because $|j+\theta/2\pi|$ must remain the same.

For θ restricted to a 2π range, there is only one value of $|j|$ satisfying (A 8). From (A 6), the Fourier transform $F(\omega)$ is then

$$F(\omega) = 1/(n-m)\left\{\delta\left(\frac{\omega}{n-m}-\theta-2\pi j\right)+\delta\left(\frac{\omega}{n-m}+\theta+2\pi j\right)\right\} \tag{A 11}$$

where $j+\theta/2\pi$ satisfies (A 7). Therefore, by using the inverse transform, the train of wavelets is given by

$$f(x) = \int_{-\infty}^{\infty} F(\omega)\exp(\mathrm{i}\omega x)\,\mathrm{d}\omega = 2\cos\{(n-m)(\theta+2\pi j)x\}. \tag{A 12}$$

We conclude that a train of general harmonic wavelets of constant amplitude always generates a pure tone. The frequency of this tone can be seen by combining (A 12) with the limiting equation (A 7). The lowest frequency achievable is $2\pi m$ (corresponding to the lower bound in (A 7)) and the highest $2\pi n$ (the upper bound), so that any tone in the frequency band of the wavelet can be modelled exactly by a train of harmonic wavelets of constant amplitude. The phase of the tone can be varied by adding a constant phase angle to a_k in (A 3).

Since a harmonic wave of arbitrary but constant frequency, phase and amplitude can be modelled exactly, and since any arbitrary function can be decomposed into harmonic functions by Fourier analysis, it follows that the general harmonic wavelets described here will provide a complete set of basis functions for signal analysis, thereby confirming the conclusion reached earlier in the paper by proving the validity of equation (19).

(*b*) *Discrete functions*

In the discrete transform illustrated in figure 5, continuous functions are assumed to be wrapped round the unit interval covered by the transform. The circular harmonic wavelet is

$$w_{c,m,n}\left(x-\frac{k}{n-m}\right) = \sum_{j=-\infty}^{\infty} w_{m,n}\left(x-j-\frac{k}{n-m}\right), \tag{A 13}$$

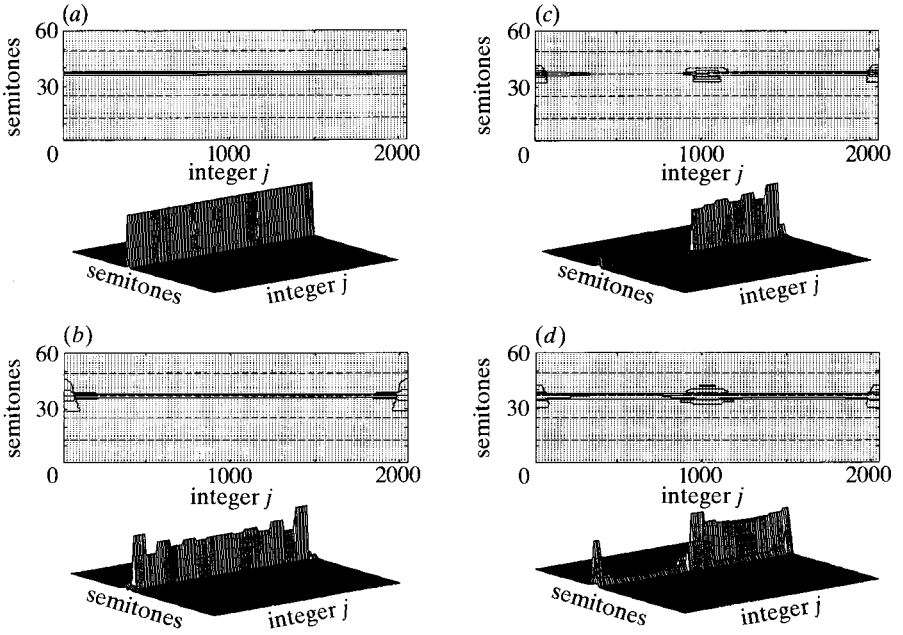

Figure 9. Discrete musical wavelet representation of the pure tone $f_j = \sin(2\pi jr/N)$, $N = 2048$; (*a*) for $r = 263$, $j = 1$ to N; (*b*) for $r = 263.5$, $j = 1$ to N; (*c*) for $r = 263$, $j = N/2+1$ to N ($f_j = 0$ for $j = 1$ to $N/2$); and (*d*) for $r = 256$, $j = N/2+1$ to N ($f_j = 0$ for $j = 1$ to $N/2$). Map contours log spaced; mesh scale linear.

where m and n must now be integers. If $W_{m,n}(\omega)$ is the Fourier transform of

$$w_{m,n}\left(x-\frac{k}{n-m}\right)$$

then $\exp(-\mathrm{i}\omega j)\, W_{m,n}(\omega)$ is the Fourier transform of

$$w_{m,n}\left(x-j-\frac{k}{n-m}\right).$$

Hence, from (5), by using the inverse transform,

$$w_{c,m,n}\left(x-\frac{k}{n-m}\right) = \sum_{j=-\infty}^{\infty} \int_{m\,2\pi}^{n\,2\pi} 1/\{(n-m)\,2\pi\} \exp\left\{-\mathrm{i}\omega\left(j+\frac{k}{n-m}-x\right)\right\} \mathrm{d}\omega\Big\}. \quad \text{(A 14)}$$

The summation over j gives

$$\sum_{j=-\infty}^{\infty} \exp(-\mathrm{i}\omega j) = 2\pi \sum_{r=-\infty}^{\infty} \delta(\omega - 2\pi r), \quad \text{(A 15)}$$

where ω is restricted to the range $m2\pi \leqslant \omega < n2\pi$. Let r, for which

$$m \leqslant r < n, \quad \text{(A 16)}$$

have the integer solutions $r = m$ to $n-1$, so that, combining (A 14) and (A 15),

$$w_{c,m,n}\left(x-\frac{k}{n-m}\right) = 1/(n-m) \sum_{r=m}^{n-1} \exp\left\{\mathrm{i}2\pi r\left(x-\frac{k}{n-m}\right)\right\}. \quad \text{(A 17)}$$

The conclusion is that a circular musical wavelet is the sum of a finite series of discrete harmonics whose frequencies are $2\pi m$, $2\pi(m+1)$, $2\pi(m+2), \ldots, 2\pi(n-1)$. It follows that a train of constant-amplitude circular harmonic wavelets can model a pure tone provided that it has one of these discrete frequencies. Tones with frequencies between the discrete values can still be modelled faithfully by the discrete transform but successive amplitudes in the wavelet expansion are then no longer the same. This is illustrated in figure 9, which shows musical wavelet maps and corresponding mesh diagrams for the analysis of a pure tone.

In figure 9*a*, the tone has frequency 263 which, because it is a discrete frequency, can be modelled exactly by a train of wavelets of constant amplitude. Its wavelet map has the simple form shown with a peak of constant height at semitone 37, which includes frequencies 256 to 270 (see table 2). Figure 9*b* is similar except that the tone's frequency is changed to the non-integer 263.5. Now the peak is still at semitone 37, but it is of irregular height. There is a discontinuity when the ends of the record are wrapped round, and this shows at each side of the wavelet map. In figure 9*c*, the original tone with a discrete frequency of 263 is turned on half way along the record and off at the end of the record (there is no signal for the first half of the record). This gives a peak on the wavelet map with an irregular profile and introduces local hills where the tone is turned on and off. In figure 9*d*, the same picture is repeated for a tone whose discrete frequency is 256. Local peaks at switching on and off are now more pronounced and there is a significant ridge along the map where there is zero signal. This is because 256 is the lowest discrete frequency in semitone 37 and the process of cancellation in the wavelet expansion involves terms from semitone 36 as well as 37. By examining figure 9*c*, *d* closely, it can be seen that the main ridge in (*d*) covers two semitones while that in (*c*) covers only one semitone.

An analogy can be drawn with the phenomenon of 'leakage' in discrete Fourier analysis: a discrete tone leaks amplitude into sidebands if the frequency of this tone is not an integer multiple of the fundamental discrete frequency. In our case the position is more complicated, with leakage occurring both (i) between wavelets at different positions and (ii) between wavelets in different frequency bands.

References

Jeans, Sir James 1937 *Science and music*. Cambridge University Press.

Mallat, S. 1989 A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Analysis Machine Intell.* **11**, 674–693.

Newland, D. E. 1993*a* Wavelet analysis of vibration. In *Proc. Structural Dynamics and Vibration Symp.*, ASME Energy-Sources Technology Conference, Houston PD-vol. 52, pp. 1–12. ASME.

Newland, D. E. 1993*b* Harmonic wavelet analysis. *Proc. R. Soc. Lond.* A **443**, 203–225.

Newland, D. E. 1993*c* *Random vibrations, spectral and wavelet analysis*, 3rd edn. Harlow: Longman and New York: John Wiley.

Newland, D. E. 1994*a* Some properties of discrete wavelet maps. *Probabilistic Engng Mech.* **9**. (Issue to mark the 65th birthday of Professor T. K. Caughey.) (In the press.)

Newland, D. E. 1994*b* Wavelet analysis of vibration – Part I: Theory, Part II: Wavelet maps. *J. Vib. Acoust.* **116**. (In the press.)

Newland, D. E. 1994*c* Harmonic wavelet families. (At present Camb. Univ. Engng Dept Report CUED/C-MECH/TR.60.)

Strang, G. 1993 Wavelet transforms versus Fourier transforms. *Bull. Am. Math. Soc.* **28**, 288–305.

Received 3 June 1993; accepted 5 October 1993