J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*

HILBERT TRANSFORMS

The Hilbert transform (*HT*) of a one-dimensional function $x(q)$ of q is another function of q, obtained by a filtering process that has special properties. It is of special interest in modulation theory and signal processing, in which the variable q is a time variable, but it is also of great significance in circuit theory where the variable of interest is frequency. Because of the Fourier transform relationship between the time and frequency domains, properties of the HT in one domain hold also in the other domain. While the formal properties are the same, they signify different important properties of the signals or circuit variables. The development of this topic is substantially mature.

Definition

The Hilbert transform of a function $x(t)$, denoted by $\hat{x}(t)$, is defined by

$$\begin{aligned}\hat{x}(t) &\triangleq -\int_{-\infty}^{\infty}\frac{x(\lambda)}{\pi(\lambda-t)}\,\mathrm{d}\lambda = \int_{-\infty}^{\infty}\frac{x(\lambda)}{\pi(t-\lambda)}\,\mathrm{d}\lambda \\ &= \int_{-\infty}^{\infty}\frac{x(t-\lambda)}{\pi\lambda}\,\mathrm{d}\lambda\end{aligned} \tag{1}$$

We note that this expression is just the convolution of $x(t)$, and $h(t), = 1/\pi t$, Fig. 1(a). By definition, the Cauchy principal value of the integral is taken; this choice is seldom of practical significance in real systems because physical constraints that restrict the integrand where we might expect $1/(t-\lambda)$ to become infinite. This aspect is referred to again later in connection with realization and approximation aspects.

The HT is a one-dimensional function, but variants have been applied to obtain some of its properties in two or more dimensions.

Salient Properties

The most significant properties of Hilbert transforms for electrical and electronic engineering are as follows:

(1) The transfer function of the filtering effected by the HT has a constant magnitude at all frequencies, and a phase of $+\pi/2$ at negative frequencies and $-\pi/2$ at positive frequencies [Fig. 1(b)].
(2) The signal $x(t) + j\hat{x}(t)$ with real $x(t)$ contains only components at positive frequencies in the frequency domain.
(3) The frequency function $X(f) - j\hat{x}(f)$ with real $X(f)$ corresponds to a time function that is zero for time $t \leq 0$, namely a causal time function.

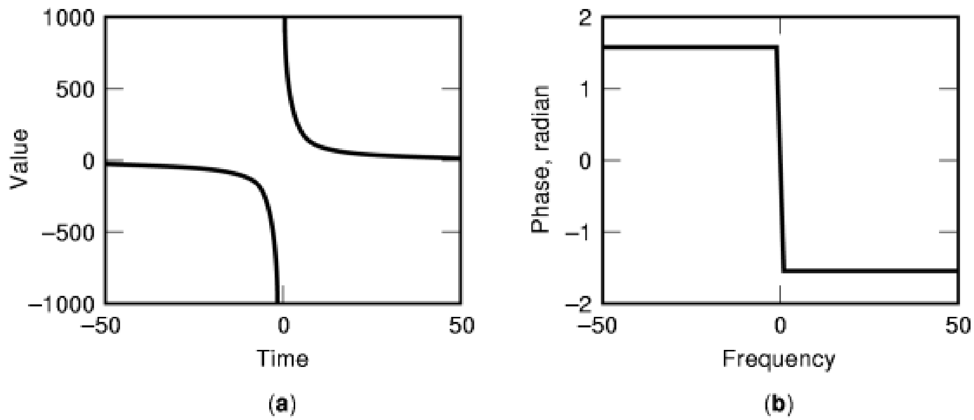

Fig. 1. (a) HT impulse response and (b) phase of its Fourier transform.

(4) For minimum phase signals there is an HT relationship between their log-envelopes and their phases.

In a later section we refer to more details of some of the properties.

Brief List of Applications

More discussion of selected applications may be found later in the article; the following is a list of its applications in electrical and electronic engineering:

(1) The "analytic signal" representation of practical signals in terms of complex values corresponding to the complex phasor method of representing sinusoids as introduced by Ref. 12
(2) In modulation theory—in single sideband modulation, envelope-compatible single sideband, and single-sideband frequency modulation (1,10,25)
(3) In analysis and application of causality to synthesis of circuits (2,16)
(4) In filter design (9)
(5) In influence functions in sheet-electrostatics, such as for acoustic surface-wave devices
(6) In pattern matching, for establishing additional dimensions (6)
(7) In image processing in connection with edge detection
(8) In speech signal processing for manipulating the bandwidth by dividing the instantaneous frequency (4)

Hilbert Transformer

A Hilbert transformer is an implementation for generating an output signal that is the Hilbert transform of an input signal. It may operate by implementation of the convolution or by operations in the frequency domain by Fourier transformation. An implemented Hilbert transformer must be an approximation, as discussed later. As a consequence of the necessary approximations some of the mathematical elegance of the true Hilbert transform disappears.

Origins

The original introduction of the HT into electrical and electronic engineering is traceable to Bode in 1945 in connection with fundamental realizability conditions for electrical circuits (2), and to Gabor in 1946 in connection with the analytic signal (12). Gabor's approach opened up the complex signal approach to communication theory just as Steinmetz's use of complex numbers had revolutionized circuit theory. Both authors referenced the theory of the HT given by Titchmarsh in 1937 which is still widely referenced (24).

Properties

Table 1 lists the salient properties of the HT for electrical and electronic engineering; these follow from the definitions. More-extensive tables are given in Refs. 11,14, and 15. Later subsections discuss several specific properties or relationships that form the basis of important applications.

Fourier Transform. The Fourier transform of the Hilbert transform is a useful tool in manipulation. It is easiest to derive the relationship by considering the unit step $u(t) = -\frac{1}{2}$, for $t < 0$; $u(t) = \frac{1}{2}$, for $t \geq 0$. The Fourier transform $U(f)$ of $u(t)$ is given by

$$U(f) = \int_{-\infty}^{\infty} u(t)\exp(-j2\pi ft)\,\mathrm{d}t \tag{2}$$

$$= \lim_{a\to 0}\int_{-\infty}^{0} \frac{1}{2}\exp(at)\exp(-j2\pi ft)\,\mathrm{d}t + \lim_{a\to 0}\int_{0}^{\infty} \frac{1}{2}\exp(-at)\exp(-j2\pi ft)\,\mathrm{d}t \tag{3}$$

$$= \frac{1}{j2\pi f} \tag{4}$$

Thus, writing "↔" for the relationship between a function and its Fourier transform

$$j2u(t) \leftrightarrow \frac{1}{\pi f}$$

and similarly, with $u(f)$ being the unit step in the frequency domain,

$$-j2u(f) \leftrightarrow \frac{1}{\pi t}$$

These characteristics in time and frequency behaviors are illustrated in Fig. 1. In particular, the magnitude of the Fourier transform is constant at unity and all positive frequencies experience $-\pi/2$ phase shift, and all negative frequencies experience $+\pi/2$

Hilbert Transform of a Constant. If we consider the HT of a constant, $x(t) = c$, namely a "dc" component, and use the Cauchy principal value of the integral, we have

$$\hat{c} = \int_{-\infty}^{\infty} \frac{c}{\pi\lambda}\,\mathrm{d}\lambda = 0 \tag{5}$$

Table 1. Salient Properties of the Hilbert Transform

	Function	Hilbert Transform or Property	Notes
a	$x(t)$	$\hat{x}(t) = (1/\pi)\int_{-\infty}^{\infty}[x(\lambda)/(t-\lambda)]d\lambda$	Notation; convolution with $1/\pi t$
b	$\hat{x}(t)$	$-x(t)$	HT of Hilbert transform
c	$ax(t) + by(t)$	$a\hat{x}(t) + b\hat{y}(t)$	Superposition, from linearity
d	$x(at), a > 0$	$\hat{x}(at)$	Scale independence
e	$x(-at), a > 0$	$-\hat{x}(-at)$	Scale independence
f	$\int_{-\infty}^{\infty} x(\lambda)h(t-\lambda)d\lambda$	$\int_{-\infty}^{\infty} \hat{x}(\lambda)h(t-\lambda)d\lambda$	Convolution, integration
g	$d^n[x(t)]/dt^n$	$d^n[\hat{x}(t)]/dt^n$	Derivatives
h	$\sin(at), a > 0$	$-\cos(at)$	Phase relationship
i	$\cos(at)$	$\sin(at)$	phase relationship
j	$\exp(j\omega t)$	$\exp\ j[\omega t - \mathrm{sgn}(\omega)\frac{\pi}{2}]$	
k	$\frac{a}{t^2+a^2}$	$\frac{t}{t^2+a^2}$	a is a constant
l	constant	0	
m	correlation $\int_{-\infty}^{\infty} x(t)y(t)dt$	$= \int_{-\infty}^{\infty} \hat{x}(t)\hat{y}(t)dt$	real x, y
n	energy $\int_{-\infty}^{\infty}[x(t)]^2dt$	$= \int_{-\infty}^{\infty}[\hat{x}(t)]^2dt$	energy or power
o	$= \int_{-\infty}^{\infty} x(t)\hat{x}(t)dt = 0$		x, $\hat{x}$ are orthogonal

which is consistent with the observation that equal and opposite contributions are made by positive-λ and negative-λ parts of the integral.

We note that if instead of the Cauchy principal value we take the integral along a path that avoids the singularity at $\lambda = 0$ in the last integral of Eq. (1) by passing around the singularity via a small semicircle such as the path γ in Fig. 2, with $p = 0$, then

$$\mathrm{HT}_{\gamma}(c) = \int_{-\infty}^{\infty} \frac{c}{\pi\lambda}\,\mathrm{d}\lambda = -jc \tag{6}$$

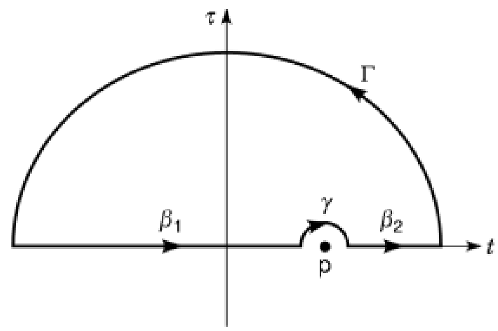

Fig. 2. Path of integration for $y(z) = y(t + j\tau)$ which is analytic for $\tau > 0$. The integration path along the real axis, $\tau = 0$, is $\beta_1 + \beta_2$.

where c is the constant value of $x(t)$.

The inversion of this HT is obtained by repetition of the HT as in the next section. Then we find the inverse $\mathrm{HT}_\gamma(-jc) = c$, as we might expect.

Iteration. The repetition of a Hilbert transform produces the original function, with its sign changed:

$$\mathrm{HT}[\hat{x}(t)] = -x(t)$$

This property is easily appreciated by noting that the HT in the frequency domain is $-ju(f)$ and $[-ju(f)]^2 = -1$. We note that this property is in accordance with the treatment of a constant by an ideal HT, but that for realizable approximate HT any constant in the original function $x(t)$ is lost.

Causality. A very fundamental appearance of the Hilbert transform arises in relationships between real and imaginary components of Fourier transforms of causal time-functions. A causal time-function $x(t)$ is one where $x(t) = 0$, $t < 0$ and $x(t) \neq 0$ for some $t \geq 0$, namely if it is regarded as the response of a system to an impulsive input it does not anticipate the input.

If the inverse Fourier transform (IFT) $x(t)$ of some frequency function $X(f)$ is causal, then the real and imaginary parts $R(f) = \mathrm{Re}(X(f))$ and $I(f) = \mathrm{Im}(X(f))$ of $X(f)$ have a special relationship as follows. Let us consider $q(t)$ to be the symmetric function, $q(t) = x(t) + x(-t)$, and then $x(t) = [\frac{1}{2} + u(t)]q(t)$, where $u(t)$ is the unit step. With $Q(f)$ written for the Fourier transform of $q(t)$ it follows that

$$X(f) = \int_{-\infty}^{\infty} \left[\frac{1}{2} + u(t)\right] q(t) e^{-j2\pi f t} \mathrm{d}t = \left[\frac{\delta(f)}{2} + U(f)\right] * Q(f)$$

where $*$ denotes the convolution, $\delta(f)$ is the Dirac δ function $\delta(f) = 0, f \neq 0$, and the Fourier transform of $u(t)$ is $U(f) = 1/(j2\pi f)$. Then

$$X(f) = \frac{1}{2}\left[Q(f) + \int_{-\infty}^{\infty} \frac{Q(\lambda)}{j\pi(f-\lambda)} \mathrm{d}\lambda\right] = \mathrm{Re}\,X(f) + j\,\mathrm{Im}\,X(f)$$

from which it is evident that

$$\mathrm{Im}(X(f) = -\mathrm{HT}[\mathrm{Re}(X(f)]$$

Furthermore

$$\mathrm{HT}[\mathrm{Im}\,X(f)] = \mathrm{HT}[-\mathrm{HT}[\mathrm{Re}\,X(f)]] = +\mathrm{Re}\,X(f)$$

Thus the real and imaginary parts of the Fourier transform of a causal function form a Hilbert transform pair.

Single-Sided Fourier Transform. If a complex signal $x(t)$ has a Fourier transform $X(f)$ that is nonzero for only positive frequencies, then $x(t) = \mathrm{Re}[x(t)] + j\mathrm{HT}\{\mathrm{Re}[x(t)]\}$; that is, the real and imaginary parts of $x(t)$ form a Hilbert transform pair. This property can be proved as for the causal time-function above. This property of causal systems, and some extensions, are sometimes referred to as the Kramers–Kronig relations (14).

Hilbert Transform of a Sinusoid. An important signal with a single-sided spectrum is the complex exponential, $c(t) = \exp(j2\pi f_1 t)$ whose Fourier transform is $C(f) = \delta(f - f_1)$. That is, it is causal in the sense that $C(f) = 0, f < 0$. The real and imaginary parts of the Fourier components of $c(t)$ are $\cos(2\pi f_1 t)$ and $\sin(2\pi f_1 t)$ respectively. Hence, from the causality in frequency, $\mathrm{HT}[\cos(2\pi f_1 t)] = \sin(2\pi f_1 t)$ and $\mathrm{HT}[\sin(2\pi f_1 t)] = -\cos(2\pi f_1 t)$.

Pre-Envelope, Envelope, and Instantaneous Frequency. The concepts of "envelope" and "instantaneous phase" of a signal are intuitively apparent for narrowband signals. Generalizing the concept of a complex function with magnitude A and total phase $\omega t + \phi$, namely $x(t) = A\,\exp(j\omega t + j\phi)$, to represent a sinusoid $s(t) = A\cos(\omega t + \phi)$, it is customary to represent a compound signal [e.g., Fig. 3(d)] as a sinusoid with time-varying real amplitude $A(t)$ and total phase $\Phi(t)$:

$$\begin{aligned} s(t) &= \sum_k A_k \cos(\omega_k t + \phi_k) = A(t)\cos[\Phi(t)] \\ &= \mathrm{Re}[A(t)e^{j\Phi(t)}] = \mathrm{Re}[y(t)] \end{aligned}$$

It is useful to define the imaginary part of $y(t)$ to be

$$\mathrm{Im}[y(t)] = \mathrm{Im}[A(t)e^{j\Phi(t)}] = \sum_k A_k \sin(\omega_k t + \phi_k) = \hat{s}(t)$$

where $\hat{s}(t)$ is the HT of $s(t)$, via Eq. (1). Dugundji, Ref. 10 has called the complex signal $y(t) = s(t) + j\hat{s}(t)$ the "pre-envelope" of $s(t)$, noting that the envelope $A(t)$ is given by

$$A(t) = [s^2(t) + \hat{s}^2(t)]^{1/2}. \tag{7}$$

Also the total phase $\Phi(t)$ is given by

$$\Phi(t) = \arctan\left[\frac{\hat{s}}{s}\right] \tag{8}$$

The instantaneous frequency $\omega_{\mathrm{inst}}(t)$ is defined to be $d\Phi/dt$ radian/s:

$$\omega_{\mathrm{inst}}(t) = \frac{d\Phi}{dt} = \frac{s(d\hat{s}/dt) - \hat{s}(ds/dt)}{s^2 + \hat{s}^2} \tag{9}$$

These $A(t)$ and $\omega_{\mathrm{inst}}(t)$ are the apparent envelope (bold line) and frequency of the oscillation in Fig. 3(d).

The special value of these definitions of envelope and phase or frequency is that they are consistent when the compound signal is shifted in frequency. That is, if each component of $s(t)$ is shifted in frequency by Δ,

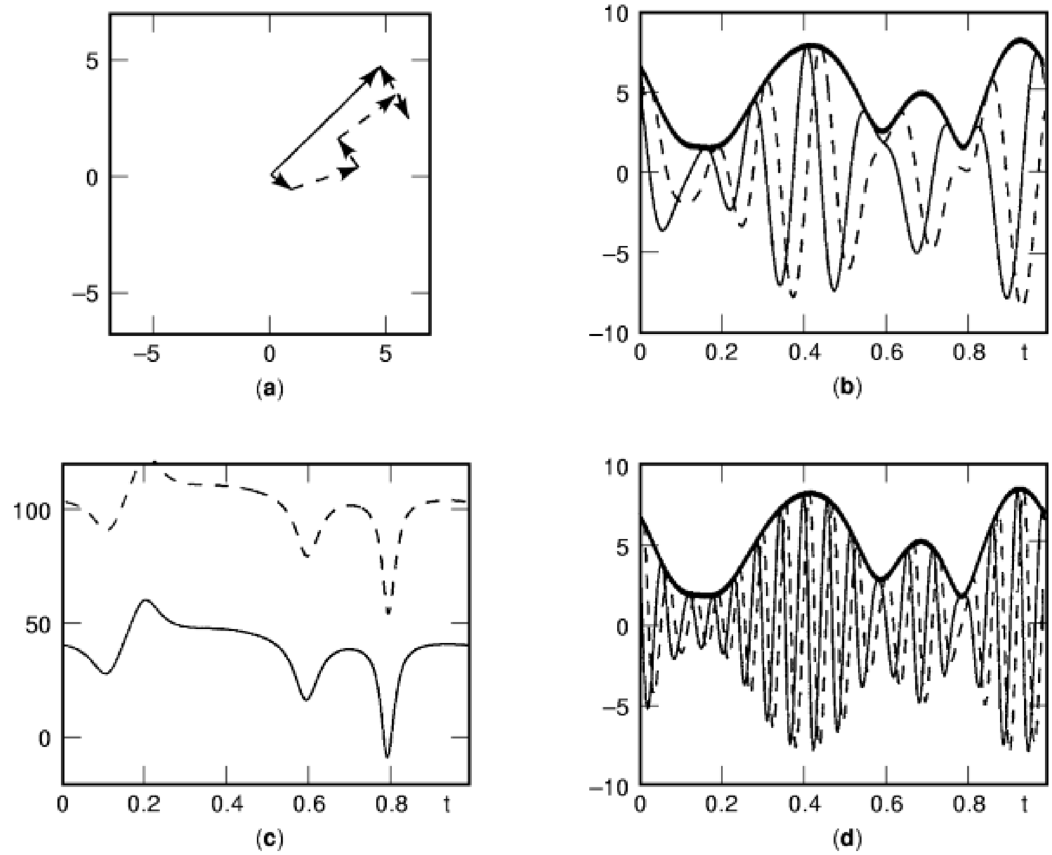

Fig. 3. Envelope and phase: $x(t) + j\hat{x}(t)$ represented by $\Sigma^6{}_{k=1}\, A_k \exp(j\theta_k)$ with arbitrary A_k and θ_k. (a) Sum of six phasors representing frequency components (*dashed*) of the signal $x(t)$ and the resultant phasor (*full line*). (b) Real signal $x(t)$ (*full line*), $\hat{x}(t)$ (*dashed*), and envelope (*heavy line*) for components at frequencies 4, 5, 6, 7, 8, and 9 Hz; (d) as for (b) but with each component shifted up in frequency by 10 Hz = 20π rad/s Hz; (c) instantaneous frequency, $d\phi/dt$, for the signal of (b) (*full line*) and of (d) (*dashed*), respectively.

forming

$$s_2(t) = \sum_k A_k \cos(\Delta t + \omega_k t + \phi_k)$$

$$= \mathrm{Re}\left[\exp(j\Delta t)\sum_k A_k \exp(j\omega_k t + j\phi_k)\right]$$

then the envelope of $s_2(t)$ is $A(t)$, the same as in Eq. (7) and the new instantaneous frequency of s_2 is $\omega_{\mathrm{inst},2}(t) = \omega_{\mathrm{inst}}(t) + \Delta$. The signals of Fig. 3(b) and 3(d) differ only in that each component has been shifted by the same amount. Fig. 3(c) shows that $\omega_{\mathrm{inst}}(t) = \omega_{\mathrm{inst},2}(t)$; these signals have a constant difference.

Analytic Signal. The term “analytic” signal is used more often than “pre-envelope” for the complex signal $y(t) = x(t) + j\hat{x}(t)$. This name is appropriate because the HT properties arise if any function $y(t)$ is analytic in the upper half complex time ($z = t + j\tau$) plane and if $y \to 0$ for $t \to \infty$, which it satisfies for finite-energy signals. Consider a function $y(z)$ in the complex z plane that is analytic (i.e., has no singularities) for $\tau \geq 0$ as shown in Fig. 2. We form the function that we are interested in integrating to get the HT, namely $y(t)/(t-p)$, and consider

$$\int_{\Gamma} \frac{y(z)}{p-z}\,dz + \int_{\gamma} \frac{y(z)}{p-z}\,dz + \int_{\beta_1,\beta_2} \frac{y(z)}{p-z}\,dz = 0$$

Letting $z = Re^{j\theta}$ on Γ, and $z = \lim_{r\to 0} re^{j\phi}$ on γ, we find that

$$j\pi y(\infty) - j\pi y(p) + \int_{-\infty}^{\infty} \frac{y(t)}{p-t}\,dt = 0$$

where the last term is the Cauchy principal value of the integral (on β_1 and β_2). By equating real and imaginary parts separately, we have

$$x(p) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\text{Im}[y(t)]}{p-t}\,dt + x(\infty) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\hat{x}(t)}{p-t}\,dt + x(\infty) \tag{10}$$

$$\hat{x}(p) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\text{Re}[y(t)]}{p-t}\,dt + \hat{x}(\infty) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(t)}{p-t}\,dt + \hat{x}(\infty) \tag{11}$$

Thus, with the exception of $x(\infty)$ and $\hat{x}(\infty)$, the analytic condition leads directly to the HT relationship between the real and imaginary parts of the complex signal $x + j\hat{x}$. Now $x(\infty)$ corresponds to any constant or dc term, which, as we have seen, is not passed via the HT. The $\hat{x}(\infty)$ term would correspond to a constant (which is zero) in the resultant HT of the original real signal.

Minimum Phase. Functions may be classified as being minimum phase in the time domain or in the frequency domain.

Minimum Phase in Time Domain. A function $y(t)$ is said to be minimum phase if $y(z) = y(t + j\tau)$, expressed as a polynomial function of z, has no zeros in the positive-τ half complex z plane (Fig. 2):

$$y(z) = \propto \prod_n (z - z_n) \tag{12}$$

Such a function has a magnitude that is unchanged by the conjugation of a zero (value of z_n). The term "minimum phase" is used to denote that function of a given magnitude [of the "common envelope set" (25)] that has the phase

$$\angle y(z) = \sum_n (\text{zero angles})$$

which is minimized if the zeros are confined to the lower-half z plane. An important result is that the spectrum of the minimum phase signal is bunched toward small values of frequency.

Minimum Phase in Frequency Domain. Similarly a Laplace transform $X(p)$ is said to be minimum phase if there are no zeros for $\sigma > 0$, that is, on the right-half of the complex-frequency $p = j2\pi f + \sigma$ plane. All passive impedance functions are of this class. Transfer of a zero from the left to the right half plane would not change the magnitude of the Fourier transform $X(f)$, but it would change the phase. Thus the concept of "common envelope" pertains in the frequency domain as well. The impedance functions of all passive circuits are minimum phase, as are those of unconditionally stable active circuits. The time function of a circuit with minimum phase frequency response tends to be bunched in time toward the beginning of the function.

With the function $x(t)$ (we consider only the time function here) expressed in amplitude and phase form

$$x(t) = x_R(t) + jx_I(t) = e^{\alpha(t)}e^{j\phi(t)}$$

where x_R and x_I are the real and imaginary parts of $x(t)$ and $\alpha(t) = \ln\sqrt{x_R^2 + x_I^2}$ is the "logarithmic envelope" of $x(t)$ and $\phi(t) = \arctan(x_I/x_R)$ is the "phase" of $x(t)$, then $\alpha(t)$ and $\phi(t)$ form a Hilbert transform pair.

The phase may be computed from

$$\phi(t) = \mathbf{HT}\{\ln[A(t)]\} \qquad (13)$$

with $A(t)$ being the envelope corresponding to Eq. 7. Figure 4 illustrates (a) a specified envelope, (b) the phase according to 13, and (c) the resulting minimum-phase signal. Part (e) of the figure shows the maximum-phase signal obtained by reversing all the phases, and (d) shows the zero-phase version, that is, with all the phases set to zero. The crosses on the plots show checking values obtained directly from processing the originally known signal which was a simple exponential.

Equation (13) is equivalent to

$$A(t) = -\mathbf{HT}\{\exp[\phi(t)]\} \qquad (14)$$

so that a signal with a minimum-phase spectrum may be produced with a specified envelope or a specified phase.

Scale Invariance. In pattern recognition tasks it is often necessary to ensure that the size or scale of the patterns to be recognized does not interfere with the recognition of similarity. Since the Hilbert transform of $s(At)$ (with A a positive constant) is $\hat{s}(At)$, the Hilbert transform of a signal is unchanged in form if the signal is changed in time scale. This is readily shown by substitution in the convolution integral, Eq. (1), and it may be appreciated easily by noting that the transfer function of the Hilbert transform is constant for all positive frequencies and also for all negative frequencies. Then change of frequency scale does not cause differential distortion.

Hilbert transform product theorem. This theorem relates HTs of multiplied signals $g(t)$ and $h(t)$ and their product. It states that if the Fourier transforms $G(f)$ and $H(f)$ are such that:

$$G(f) = 0, \quad f > f_0 \quad \text{and } H(f) = 0, \quad f < f_0$$

then

$$\mathbf{HT}[g(t)h(t)] = \mathbf{HT}[g(t)]h(t) = g(t)\mathbf{HT}[h(t)]$$

The theorem has been generalized for multidimensional signals (26).

Two-Dimensional Hilbert Transforms. It is possible to generalize the definition of the HT for two or more dimensions. If the original signal or image is the scalar function of N variables t_n, namely $x\{t_1, t_2, \ldots\}$,

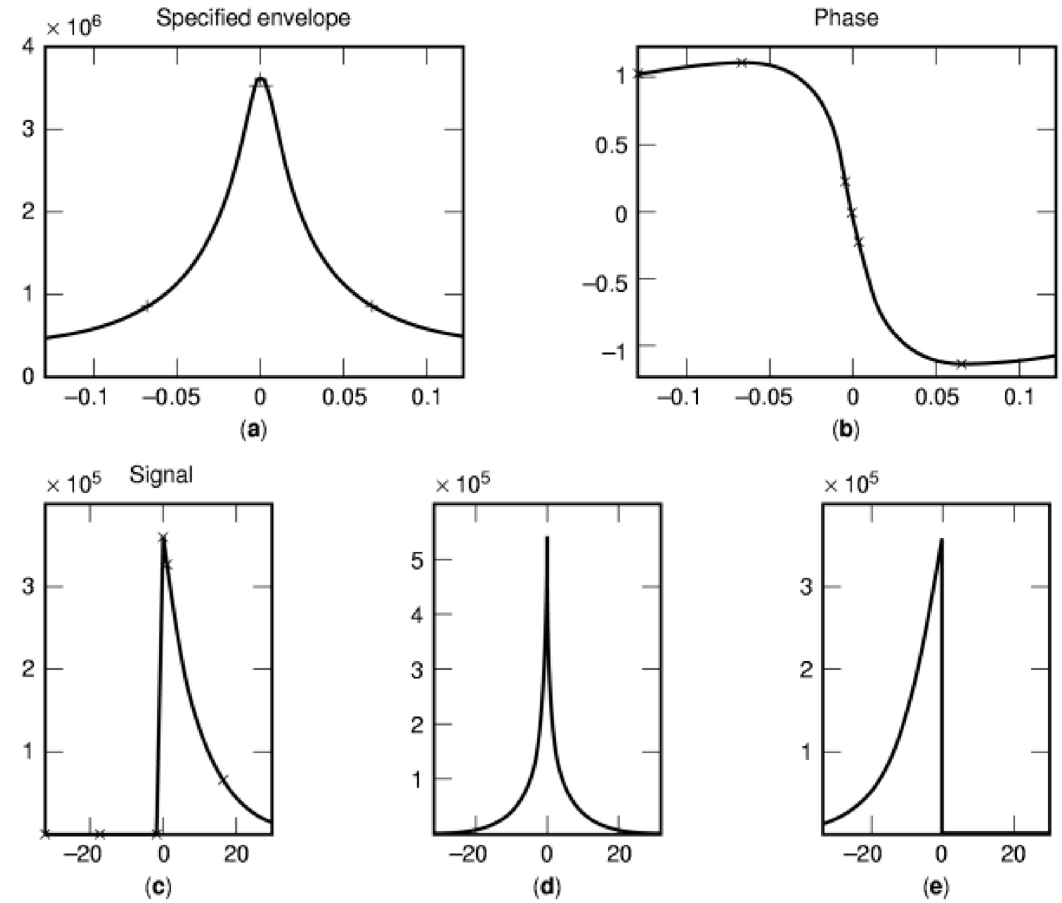

Fig. 4. Finding phase from envelope: (a) The envelope $A(t)$ of some function $x(t)$ of time; (b) phase computed from $\phi(t) = \text{HT}[\ln[A(t)]$; (c) Fourier transform of minimum-phase signal $A(t)\exp(j\phi)$; (d) Fourier transform of zero-phase signal $A(t)\exp(j\phi)$; (e) Fourier transform of maximum-phase signal $A(t)\exp(-j\phi)$.

then the N-dimensional HT can be defined as

$$\hat{x}(\boldsymbol{t}) = \left(\frac{1}{\pi}\right)^N \int \cdots \int \frac{x\{\lambda_1, \lambda_2, \ldots\}}{(t_1 - \lambda_1)(t_2 - \lambda_2)\ldots} d\lambda_1 d\lambda_2 \ldots$$

An example of the resultant point spread function of a two-dimensional HT is shown in Fig. 5. The Fourier transform of $\hat{x}(\boldsymbol{t})$ then becomes

$$j^N \text{sgn}(f_1)\text{sgn}(f_2)\ldots X(f_1, f_2, \ldots)$$

where $X(f_1, f_2, \ldots) \leftrightarrow x\{t_1, t_2, \ldots\}$.

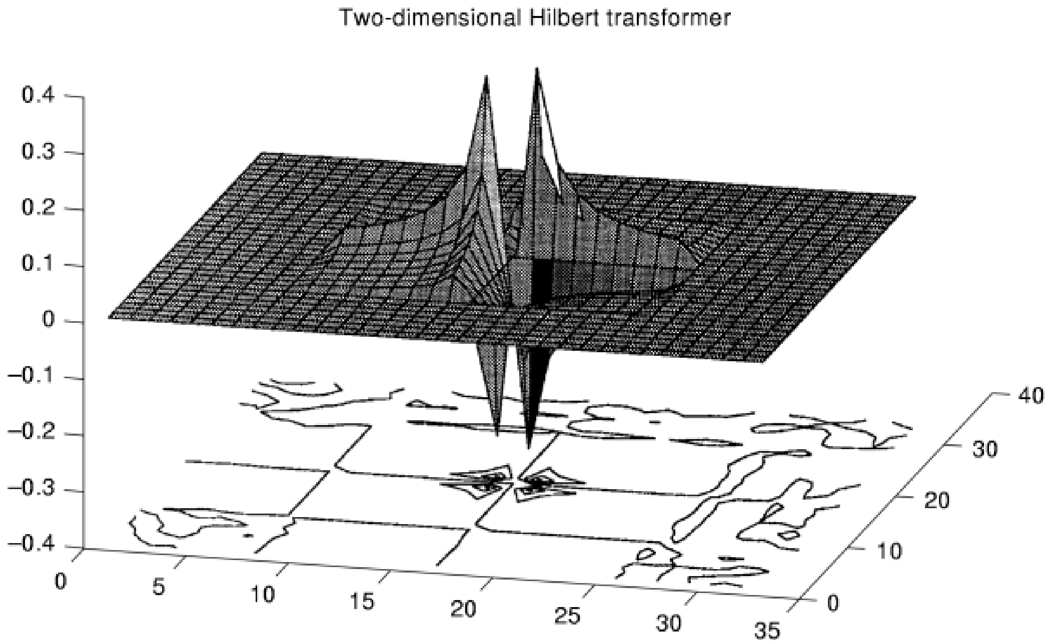

Fig. 5. Point spread function of a two-dimensional HT.

Other generalizations of the HT. The Reisz transform (27) is a singular integral operator that reduces to the HT when its parameter $n = 1$. The kernel of the Reisz transform is:

$$K_j(\mathbf{x}) = c_n\left(\frac{x_j}{|\mathbf{x}|^{n+1}}\right)$$

where $X = \{x_1, x_2, \ldots, x_j, \ldots, x_n\}$ and $C_n = \Gamma[(n + 1)/2]/\pi^{(n+1)/2}$, Γ being the gamma function. In particular, for $n = 1$ we have the HT kernel:

$$K_1(x) = \frac{x}{\pi\,|x|^2} = \frac{1}{\pi x}$$

Implementation of Hilbert Transformers

Approximations and Effects. As noted above, all implementations of Hilbert transformers involve some approximations and hence departures from the ideal. While there are many detailed aspects of the approximation problem, three major departures are fundamental. They all arise from the properties of the kernel $1/t$ of the integral (1) etc.

Causality. The time function $1/t$ begins at $t < 0$ as shown in Fig. 6(a). Thus the HT cannot be realized in a causal world and the practical response has to be zero for $t < 0$. The realized time response has to be approximated, with $h_r(t) = 0$, $t < T_1$; $h_r(t) = 1/\pi(t - \tau)$, $t \geq T_1$, involving a delay τ. Furthermore implementations involving computation of convolutions (in nonrecursive filters) with finite numbers of elements require that

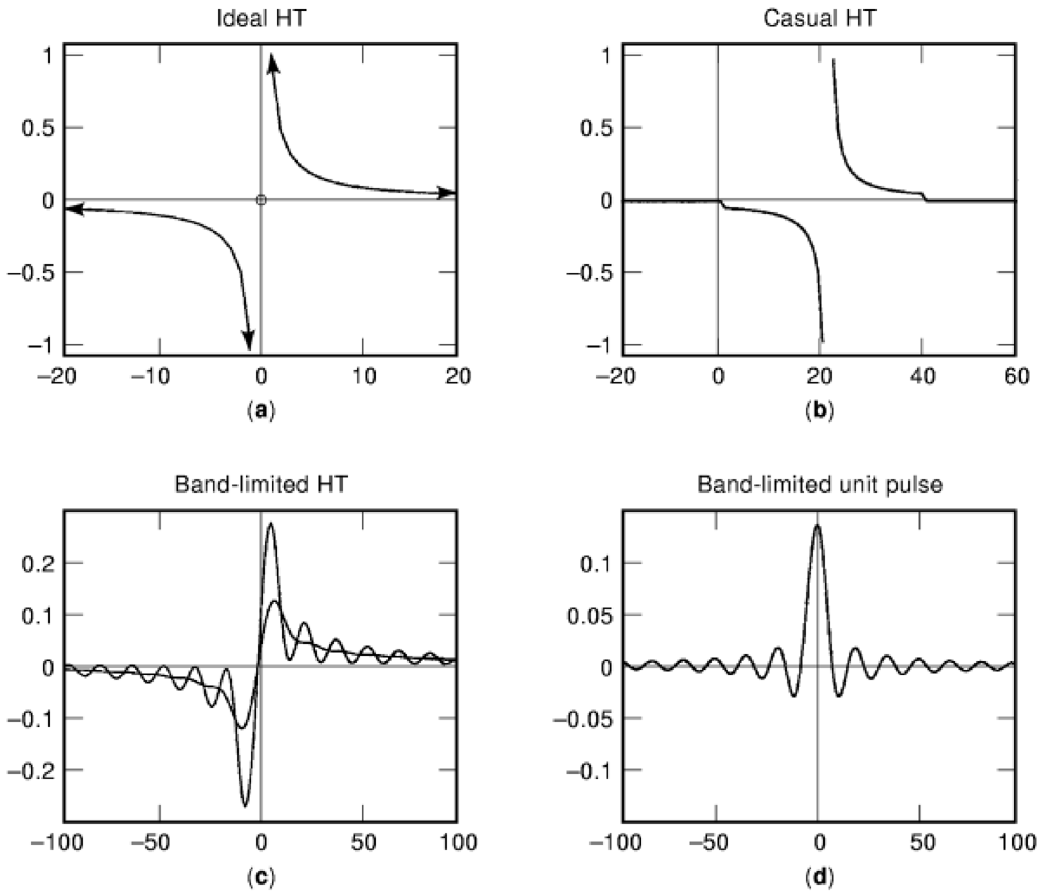

Fig. 6. Fundamental restrictions of practical Hilbert transformers: (a) Ideal HT begins at $-\infty$, continues to ∞, and has singularity at the origin; (b) causal HT must be zero before $t=0$; (c) band-limited HT is continuous through $t=0$ and has finite values. The rippling HT corresponds to hard-limiting in frequency; the smoother one was obtained by band-limiting with a Hanning window, (d) $\sin(t)/t$ form of band-limiting operator used to obtain the rippling version in (c).

$h_r(t)=0$ for t greater than some T_2. Typically we use symmetric formulations which result in no distortion of the phase, and set

$$h_r(t) = 0, \qquad t < 0 \tag{15}$$

$$= \frac{1}{\pi(t-\tau)}, \qquad 0 \le t \le 2\tau \tag{16}$$

$$= 0, \qquad t > 2\tau \tag{17}$$

as shown in Fig. 6(b).

Finite Magnitude. The function $1/t$ tends to $-\infty$ as $t \to 0-$ and to $+\infty$ as $t \to 0+$. In physical systems the values are limited. Fortunately we are normally also restricted to processing signals that are effectively band-limited, whether ideally (with sharp cutoffs) or with gradual cutoffs. Band-limiting effects a convolution

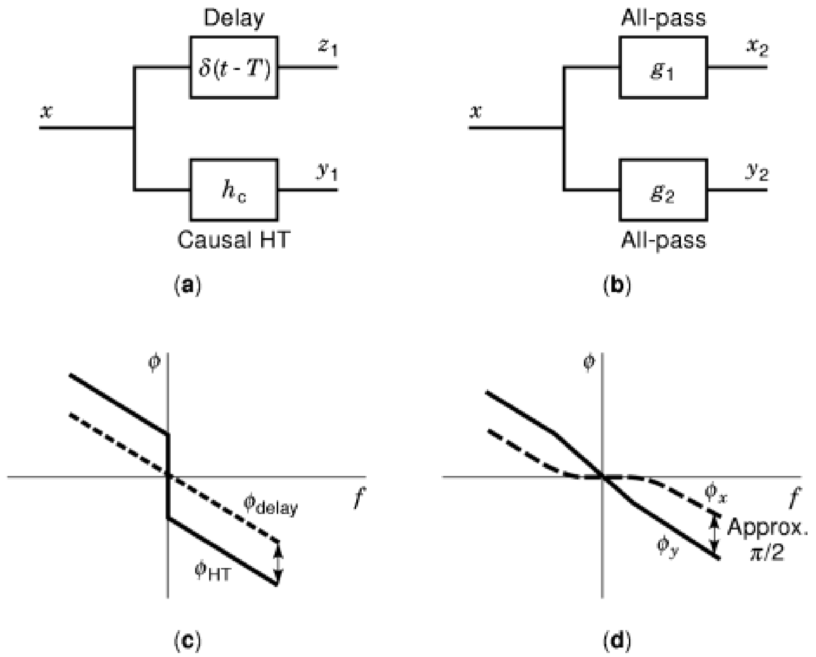

Fig. 7. Quadrature filters: (a) A causal approximate HT requires a delay. The desired quadrature phase relationship is obtained between a delayed version $x(t - T) = x_1(t)$ of the original signal $x(t)$ and $\mathrm{HT}_c(x(t)) = y_1(t)$. Phases of the relevant paths are shown in (c). (b) An equivalent quadrature phase relationship may obtain between two all-pass versions $x_2(t)$ and $y_2(t)$, as depicted in (d).

of the $1/t$ ideal HT with some smoothing function, and the effect is to make finite the transition slope at $t = 0$ and hence to restrict the negative and positive peaks at $t = 0-$ and at $t = 0+$, as Fig. 6(c) and 6(d) illustrate.

Quadrature Filters. Often it is not necessary to generate exactly the HT, but it is sufficient to have two signals that have an approximate HT relationship between them because of practical conveniences of synthesizing suitable all-pass phase-shifting circuits. Figure 7 illustrates the principles. Such systems have been used widely in connection with generation of single-sideband signals, as explained in the next section on applications.

Discrete Hilbert Transformers. Most HTs are implemented in digital signal processing systems, and they operate on sampled data. Thus the the Nyquist criterion is applicable; the sampling interval, T s, should be at most $1/(2B)$ where B is the highest frequency of the signal, unless there are special considerations associated with narrowband signals.

In direct convolution, the integral of Eq. (1) is approximated by the finite summation

$$y_n = \sum_{r=1}^{N} h_r x_{n-r} \qquad (18)$$

where the values of the unit pulse response of the Hilbert transformer represent the function $1/t$. An ideal HT band-limited to $B = \frac{1}{2}T$ Hz is readily found by inverse Fourier transforming the frequency function $\mathrm{H}(f) = +j$, $-B < f < 0; = -j$, $0 \leq f < B$, and its sample values are $h_n = [\cos(n\pi) - 1]/n$, which are just samples of $1/n$, with the even-n values set to zero. Of course a finite duration is needed, and the resultant errors in the frequency response may be controlled by the choice of a suitable window function. The overall duration in seconds of the

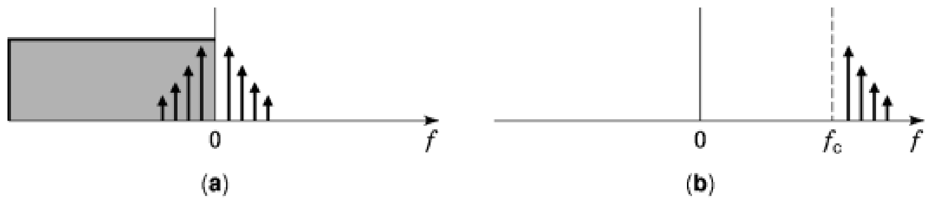

Fig. 8. Single-sideband generation: (a) Spectrum of analytic signal $z = x + j\hat{x}$ (b) shifted spectrum of $ze^{j2\pi f}c_t$.

response is controlled by the frequency resolution required, which is determined by the lowest frequency since the frequency response has to change from $-j$ to j. Letting the transition band be b Hz sets the duration to be of the order of 1/b s, corresponding to T/b samples. Several times this length may have to be used depending on the accuracy of the required response. If the samples of the response are truly skew-symmetric, the only errors in the transform are in the magnitude of the response and not of the phase. References 19 and 21 provide good discussions of the design principles and methods.

Transform Domain HT. Since the HT is a convolution then any method for effecting convolutions is potentially useful. The fast Fourier transform algorithm for implementing the discrete Fourier transform (*DFT*) is widely used for such convolutions. The HT may be specified conveniently in the frequency domain by $H(f) = -j\,\mathrm{sgn}(f)$, which is a multiplication of the relevant coefficients of the DFT. Because DFT filters form finite-duration convolutions, they are equivalent to the direct convolution methods as far as design fundamentals are concerned.

Recursive HT for Narrowband Signals. The duration of the response of a filter is associated with its frequency resolution. Recursive structures provide infinite impulse response (*IIR*) filters. A complex resonator is formed by the difference equation

$$y_n = ay_{n-1} + w_n$$

with $a = \exp(j\theta)$ which has the complex impulse response $\{1, a, a_2, \ldots\}$ whose real and imaginary parts are $\{1, \cos(\theta), \cos(2\theta), \ldots\}$ and $\{0, \sin(\theta), \sin(2\theta), \ldots\}$, respectively. Reference (3) uses the fact that such responses are approximate HTs of each other to form Hilbert transformers in speech processing. The resonators were combined with comb filters such that the complex exponentials were canceled after a finite time. The resultant filters have transfer functions which were shown to be equivalent to DFT filters, but with few computations if the relative bandwidths were small.

Applications

Modulation and Frequency Shifting. The HT provides a theoretical framework for many processes on modulated signals. Single sideband modulation is readily effected by shifting the frequency of an analytic signal: First the analytic signal is formed from the original message signal, $z(t) = x(t) + j\hat{x}(t)$. Then the frequency is shifted by multiplication by $\exp(2\pi f_c t)$, Fig. 8. Instead of the formal HT, if linear distortion of the signals is acceptable, as it is in the case of speech signals, the quadrature filter approach is often used.

Other modulation processes involving HTs have been aimed at conservation of radio spectrum bandwidth:

(1) Compatible Single Sideband 1 Refers to the generation of amplitude modulated signals that are compatible with the envelope detection used in the majority of broadcast-band receivers but that have the spectral

occupancy of single sideband signals. The principle invoked here is that if the envelope $A(t)$ is specified then the minimum-phase signal may be generated by setting the phase to $\Phi(t) = \mathrm{HT}(\log[A(t)]$, and it turns out that the resultant one-sided spectrum is more compact than that of a double-sideband signal. The complicated dependence of the performance on the filters used in radio receivers is a significant deterrent in these approaches.

(2) Single-Sideband FM 1 Has a similar motivation, and depends on the same relationship, used in reverse. A signal that has been frequency modulated so that it has the desired instantaneous frequency and phase has an envelope impressed on it such that the spectrum is more compact. The relationship $A(t) = -\exp(\mathrm{HT}[\Phi(t)]$ leads again to the minimum phase signal that has a single-sided and compacted spectrum.

Instantaneous Frequency Manipulations of Signals. There were many efforts to reduce the bandwidth occupied by communication signals such as speech. Typically manipulations of the instantaneous phase, or frequency dividing, were shown to have some compressing effect. The use of the HT was instrumental in the design and diagnosis of such systems, for example, Ref. 4.

Electric Circuits. The relationships between real and imaginary parts of a Fourier transform in the frequency domain have been widely recognized in electric circuit theory and design and constitute some of the earliest applications in electrical and electronic engineering (2), although Bode did not use the term "Hilbert transform" explicitly.

Filter and Transmission-Line Synthesis. The HT has been valuable in solving problems of finding a lossless electrical network or transmission line cascade that will provide a desired power spectrum into a known resistive load from a given current source (16,23). An example is the design of a cascade of acoustic tubes to approximate the shape of the human vocal tract for a given speech sound. The radiated power density into the air is a known function of frequency, $P(f)$, and the load is the acoustic radiation impedance of the lips. The power input to the configuration is $P_1(f) = |I_1(f)|^2 \mathrm{R}_1(f)$, where $I_1(f)$ is the assumed known input (air) current transform and $R_1(f)$ is the real part of the input impedance $Z_1(f) = R_1(f) + jX_1(f)$. Since the system is assumed to be lossless, then $P_1(f) = P(f)$, and hence we can find that $R_1(f) = P(f)/|I_1(f)|^2$. We need to find the imaginary part $X_1(f)$. Since the physical system is realizable the impulse response of $Z_1(f)$ is causal and then $X_1(f) = \hat{R}_1(f)$. This is the step that depends on the HT. The subsequent procedure depends on an algorithmic approach to undoing iteratively a system of echoes that are assumed to correspond to reflections from changes in cross section of the cascade of tubes.

Filter Design. The single-sided spectral property of an analytic signal is equivalent to high-pass filtering the real signal with a cutoff at zero frequency [Fig. 9(a) and 9(b)]. Reference 9 has used this property with frequency shifting or heterodyning to implement filters with adjustable cutoff frequencies. In this application of the HT a signal s is converted to an analytic signal s_A. The spectrum of s_A with components 1, 2, 3, and 4 is shifted by multiplication with a complex carrier $\exp(-j\omega t)$ so that the cutoff of zero frequency appears at the desired part of the translated signal spectrum, here between 2 and 3 [Fig. 9(c)]. Thus the components (1, 2) below this frequency are eliminated, giving s_B. The resultant complex signal is translated back to the original position by multiplying s_B by $\exp(+j\omega t)$, as in Fig. 9(d).

Pattern Recognition; Scale Invariance. The property of scale invariance is attractive in pattern recognition. References 6 and 8 are early expositions of these principles.

The simplest illustration is that a sine wave becomes a cosine wave under Hilbert transformation, and the frequency is of course retained as is the amplitude. Thus the plot of a sinusoidal signal and its HT is just a circle, as is well known; it matters not if the signal is at 1 Hz or 100 MHz. The scale-invariant effect is shown in Fig. 10 for two multifrequency signals that differ in their time scales.

A more interesting effect is seen in Fig. 11 for two multifrequency signals which differ in that their time-scales are varying slowly with time (5). Here, the slow time variation causes only a slight distortion of the relationship between the two patterns. If the drift of scale is maintained the patterns remain similar, but the

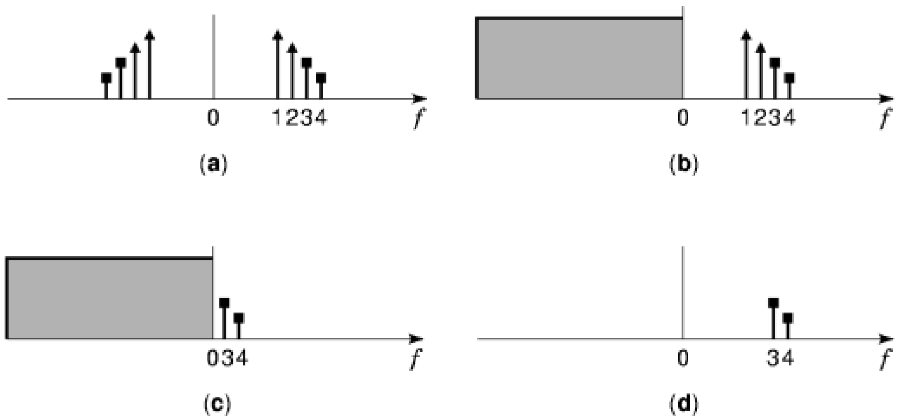

Fig. 9. Use of HT for filter design: (a) Spectrum of original signal s; (b) spectrum of analytic signal s_A derived from (a); (c) spectrum of $s_B = \mathrm{HT}[s_A \exp(-j\omega t)]$; (d) spectrum of $s_{HP} = \mathrm{Real}[s_B \exp(+j\omega t)]$.

frequency of sampling points may change over a wide range. Such ideas were developed for matching warped patterns in Ref. 18.

Image Processing and Edge Detection. Edge detection in images is based on some form of differentiation that produces a peak where the derivative of the intensity changes fast. For edges that have a given shape of intensity versus position, it might be desirable to ensure that the magnitude edge detection signal does not depend significantly on the scale of the edge, namely that edges whose transition regions are different should produce the same magnitude of output. The scale invariant property of the HT might appear to be useful here. Such reasoning appears to lie behind some attempts to use the HT in two dimensions as a filter whose output is a good indicator of edges and is independent of scale. However, we are of the opinion the two-dimensional HTs are not really appropriate in that they have characteristics dependent on the orientation of an edge (5).

However, direct application of the two-dimensional HT is not appropriate as it has characteristics that are dependent on the orientation of an edge. Instead, different but related filtering that derives from the Reisz transform (see above, under “Other generalizations of the HT”) have been used.

Computation of bias in spectral estimation. In spectral estimation a record of the signal of duration T s is selected and the spectral density is estimated by evaluating the squared magnitude of its Fourier transform. Provided T is large enough to avoid truncation of the autocorrelation of the signal then the bias in the spectral estimate $\tilde{W}(f)$ due to the finite value of T is found to be

$$W(f) - \tilde{W}(f) = \frac{1}{2\pi^2 T} \frac{\mathrm{d}}{\mathrm{d}f} \int_{-\infty}^{\infty} \frac{W(\zeta)}{f - \zeta} d\zeta$$

where $W(f)$ is the true spectrum and principal value of the integral is used. Stark et al (28) give corresponding two dimensional expressions relevant to the optical estimation of the spectral density in images.

HT relations in potential problems. Some important HT relationships exist between functions that occur in potential theory. These have been used in geophysics in the analysis of perturbations of the earth’s magnetic field. Such relations are general in electric and magnetic field theory, and a simple example illustrates. We introduce the ideas with reference to electrostatic potentials because these may be more familiar. The

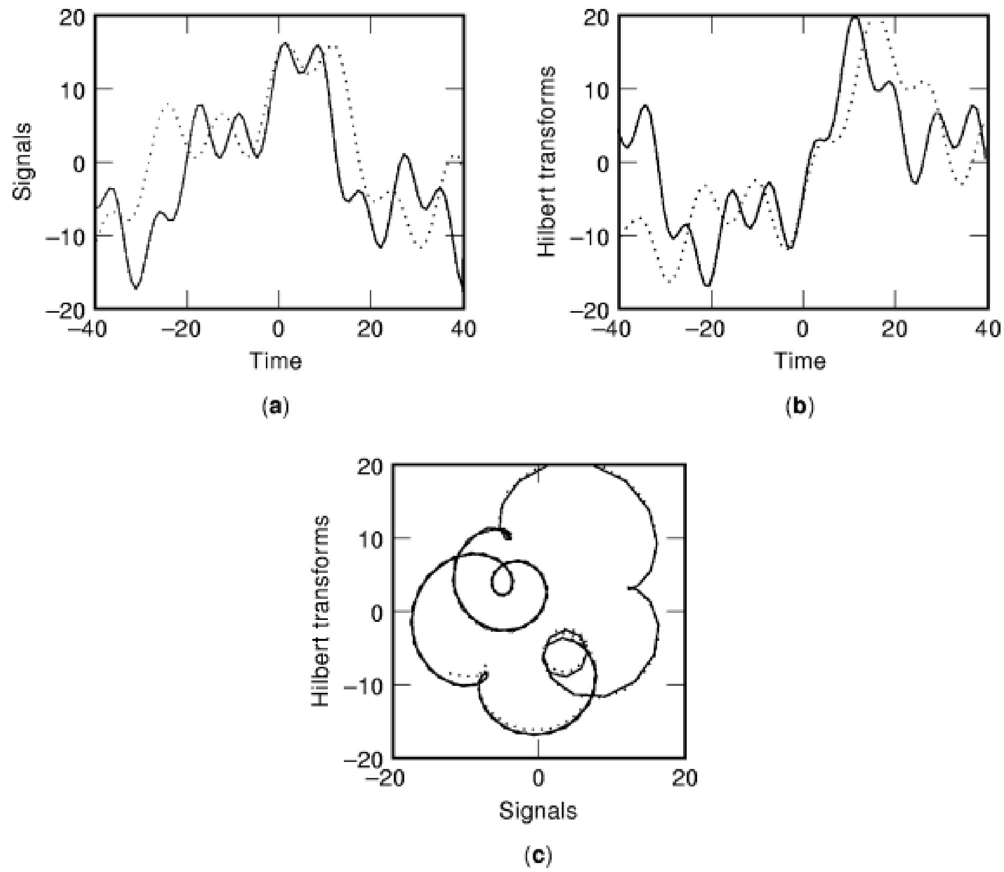

Fig. 10. (a) Signal $x(t)$, and time-scaled $x(kt)$ (*dotted*); (b) Hilbert transforms $\hat{x}(t)$ and $\hat{x}(kt)$; (c) plots of HT($x(t)$) against $x(t)$ and of $\hat{x}(kt)$ against $x(kt)$ (*dotted*).

potential $V(x, y, z)$ at the point $\{x, y, 0\}$ due to a charge q coulombs at the origin is

$$V = \frac{q}{\epsilon r} = \frac{q}{\epsilon}\frac{1}{\sqrt{x^2 + y^2 + 0^2}}$$

where ϵ is the permittivity. Thus we have

$$\frac{\partial V}{\partial x} = \frac{q}{\epsilon}\frac{x}{x^2 + y^2} \quad \text{and} \quad \frac{\partial V}{\partial y} = \frac{q}{\epsilon}\frac{y}{x^2 + y^2} \tag{2}$$

By comparing with the HT relations:

$$\mathrm{HT}\left[\frac{a}{t^2 + a^2}\right] = \frac{t}{t^2 + a^2} \quad \text{and} \quad \mathrm{HT}\left[\frac{t}{t^2 + a^2}\right] = \frac{-a}{t^2 + a^2}$$

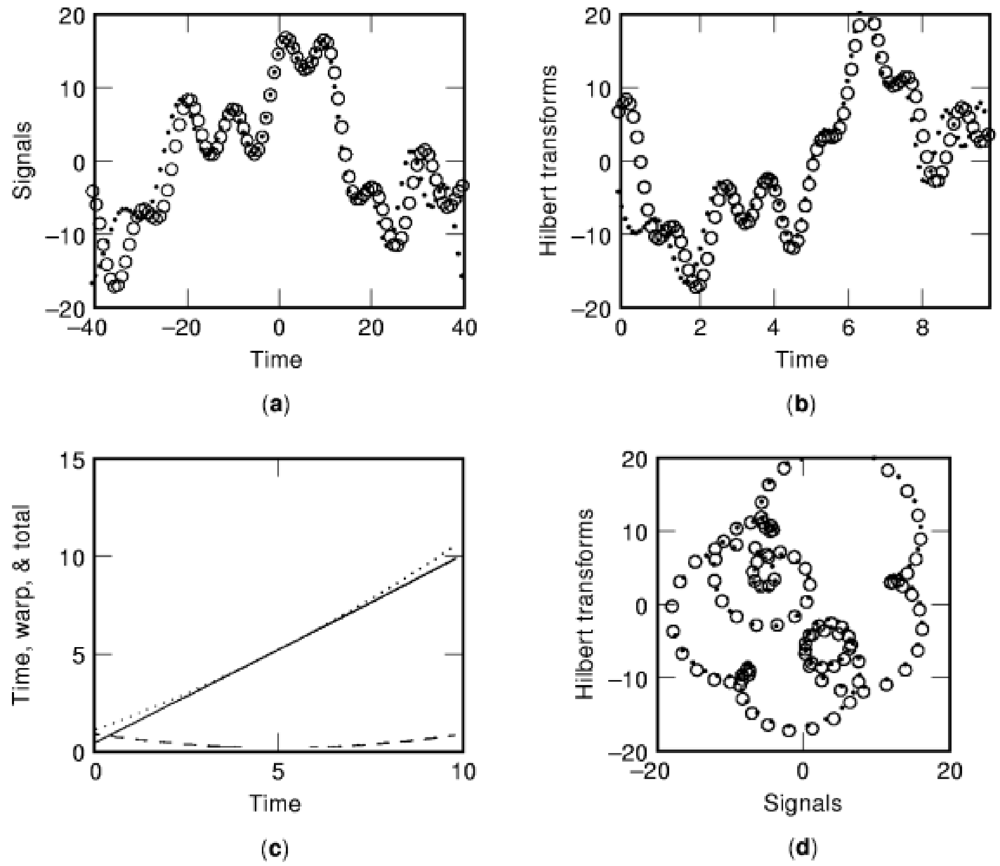

Fig. 11. (a) Signal $x(t)$, and time-warped $x[f(t)]$; (b) Hilbert transforms $\hat{x}(t)$ and $\hat{x}[(f(t)]$; (c) the time scales (*dotted* is the warped time vs. original time, and *dashed* is the distortion component, a quadratic term). (d) Plots of $\hat{x}(t)$ against $x(t)$ and of $\hat{x}[f(t)]$ against $x[f(t)]$.

where a is a constant, we see that for constant y:

$$\frac{\partial V}{\partial y} = \mathrm{HT}\left[\frac{\partial V}{\partial x}\right] \quad \text{and} \quad \frac{\partial V}{\partial x} = -\mathrm{HT}\left[\frac{\partial V}{\partial y}\right] \tag{3}$$

Nabighian, Ref. (29) derived from such relationships two dimensional methods for resolving the magnetic anomalies of the earth's magnetic field associated with magnetic ore bodies and he extended them in 3D in Ref (30).

If electric charges are arranged in rows on a flat surface, the influences of these charges may lead to Hilbert transform relationships. A practical example illustrates. Acoustic surface wave filters are devices in which a pattern of electric charges is applied to the surface of a piezoelectric plate and the resultant stress launches travelling surface waves [Fig. 12(a)]. These waves are sensed at another site on the plate by receiving electrodes on which potentials arise from the induced charges associated with the pattern of strain in the waves.

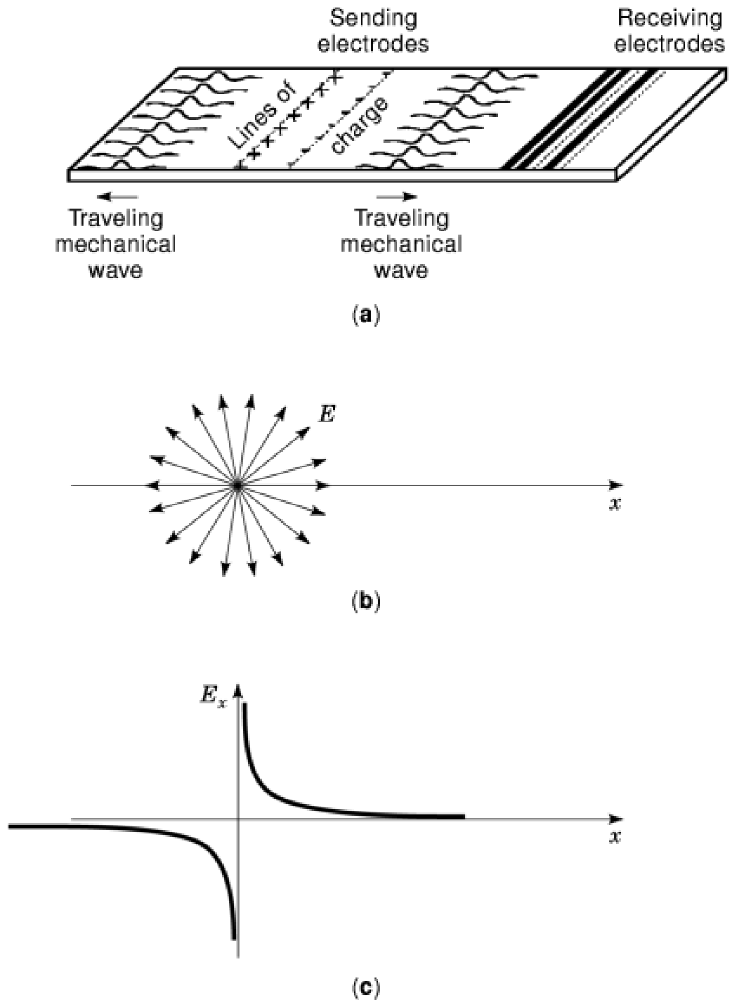

Fig. 12. Lines of surface charge: (a) Line of charge in a surface; (b) field at radius *r* from the line; (c) value of *x* component of field is proportional to 1/*x*.

The sending and the receiving electrodes are in the form of straight bars across the surface, and the waves ideally have straight wavefronts perpendicular to the direction of propagation. The influence of the sending electrodes is seen if we consider the electric field of a single line of positive charge [Fig. 12(b)]. The magnitude of the field is proportional to 1/*r*, where *r* is the radial distance from the line. The direction of the field is radially from the charge. Thus in the plane of the surface, the field along the surface, E_x, is constant/*r*. For a distribution of charge density $q(x)$ the net field at a position d, which is $d - x$ from the charge at x, is obtained by integrating over all regions of charge:

$$E_x(d) = \text{constant} \times \int_{x=-\infty}^{\infty} \frac{q(x)}{d-x}\, dx$$

and we recognize this expression as the Hilbert transform again. A similar discussion shows that the reception electrodes introduce another Hilbert transform. Hence the overall effect of the surface wave device is to include two convolutions by Hilbert transforms, which corresponds to a simple delta function.

Further Reading

References 7 and 20 present the theory of the HT in very readable form. References 19 and 21 are excellent expositions of the principles in discrete-time systems and design methods. For the mathematics, Refs. 11,13, and 24 are recommended; Refs. 14 and 15 are more extensive treatments oriented toward signal processing.

BIBLIOGRAPHY

1. E. Bedrosian The analytic signal representation of modulated waveforms, *Proc. IRE*, **50**:, 2071–2076, 1962.
2. H. W. Bode *Network Analysis and Feedback Amplifier Design*, New York: Van Nostrand, 1945.
3. R. E. Bogner Frequency Sampling Filters, Hilbert Transformers and Resonators, *Bell Syst. Tech. J.*, **48**: 501–510, 1969.
4. R. E. Bogner J. L. Flanagan Frequency multiplication of speech signals, *IEEE Trans. Audio Electroacoust.*, **17** (3): 202–208, 1969.
5. R. E. Bogner Scale independent filters, *Proc. Int. Symp. Signal Process. Its Appl.*, Gold Coast, Qld., 1996, pp. 495–498.
6. R. J. Bolton *Presentation and Pattern Recognition of Hilbert Transformed Electrocardiograms*, Rep. EE83/9, July 1983, Dept. Electr. Eng., Univ. of Queensland.
7. R. N. Bracewell *The Fourier Transform and Its Applications*, New York: McGraw-Hill, 1965.
8. M. S. Corrington *Hilbert Transforms*, Radio Corp. of America Company, Private EM-62.421-25, June 15, 1962.
9. G. D. Cain Hilbert transform relations for products, *Proc. IEEE*, **61**: 663–664, 1973.
10. J. Dugundji Envelopes and pre-envelopes of real waveforms, *IRE Trans. Inf. Theory*, **4**: 53–57, 1958.
11. A. Erdelyi, (ed.) *Tables of Integral Transforms*, Bateman Manuscript Project, vol. 2, pp. 239–262.
12. D. Gabor Theory of communication, *J. IEE*, **93**, part III (26): 429–441, 1946.
13. E. A. Guillemin *The Mathematics of Circuit Analysis*, New York: Wiley, 1949.
14. S. L. Hahn *Hilbert Transforms in Signal Processing*, Norwood, MA: Artech House, 1996.
15. S. L. Hahn Hilbert transforms, in A. D. Poularikas, (ed.), *The Transforms and Applications Handbook*, Boca Raton, FL: CRC Press and IEEE Press, 1996.
16. B. Kinariwala F. F. Kuo N.-K. Tsao *Linear Circuits and Computation*, New York: Wiley, 1973.
17. N. Maheswaran R. E. Bogner A new timescale warping algorithm for single dimensional and multidimensional speech parameter contours, *Proc. 1st Australian Speech Sci. Technol. Conf.*, Canberra, 1986.
18. N. Maheswaran B. R. Davis Analytical signal processing for pattern recognition, *IEEE Trans. Acoust. Speech Signal Process.*, **36**: 1645–1649, 1990.
19. A. V. Oppenheim R. W. Schafer *Discrete-Time Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
20. A. Papoulis *The Fourier Integral and Its Applications*, New York: McGraw-Hill, 1962.
21. L. R. Rabiner B. Gold *Theory and Application of Digital Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1975.
22. S. O. Rice Envelopes of narrow-band signals, *Proc. IEEE*, **70**: 692–699, 1982.
23. E. V. Stansfield R. E. Bogner Determination of vocal tract area function from transfer impedance, *Proc. IEEE*, **120**: 153–158, 1973.
24. E. C. Titchmarsh *Introduction to the Theory of Fourier Integrals*, Oxford UK: Oxford Univ. Press, 1937, pp. 119–151.
25. H. B. Voelcker Towards a unified theory of modulation, Part I: Phase-envelope relationships, *Proc. IEEE*, **54**: 340–353, 1966; Part II: Zero manipulation, **54**: 735–755, 1966.
26. Henry Stark An extension of the Hillbert Transform product theorem, *Proc. IEEE* **59**: 1359–1360, 1971.
27. E.M. Stein G. Weiss Introduction to Fourier analysis on euclidian spaces, Princeton, NJ: Princeton University Press, 1971, pp 223–224
28. H Stark W.R. Bennett M. Arm Design Considerations in Power Spectra Measurements by Diffraction of Coherent Light, *Applied Optics* **8**: 2165–2172, 1969.

29. M.N. Nabighian The analytic signal of two-dimensional magnetic bodies with polygonal cross-section: its properties and use for automated anomaly interpretation, *Geophysics* **37**: 507:517, 1972.
30. M.N. Nabighian Toward a three-dimensional automatic interpretation of potential field data via generalized Hilbert transforms: fundamental relations, *Geophysics* **49**: 780–786, 1984.

ROBERT E. BOGNER
The University of Adelaide