COMMUNICATION SCIENCES

AND

ENGINEERING

X. STATISTICAL COMMUNICATION THEORY*

Academic and Research Staff

Prof. Y. W. Lee
Prof. A. G. Bose
Prof. J. D. Bruce
Prof. A. V. Oppenheim
Prof. D. E. Nelsen
G. Gambardella

Graduate Students

T. Huang
D. J. Matthiessen
M. F. Medress
V. Nedzelnitsky
P. E. Perkins
E. M. Portner, Jr.
L. R. Poulo
A. E. Roland
S. Samsam-Bakhtiary
R. W. Schafer
J. E. Schindall
F. P. Tuhy
J. L. Walker
J. J. Wawzonek

A. WORK COMPLETED

1. EFFECTS OF NOISE ON CEPSTRAL ESTIMATION OF ECHOES

This study has been completed by E. M. Portner, Jr. In August 1967, he submitted the results to the Department of Electrical Engineering, M. I. T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

A. V. Oppenheim

B. SOME PROPERTIES OF THE CEPSTRUM

1. Introduction

The cepstrum has been defined[1] as the Fourier transform of the logarithm of the amplitude spectrum of a signal. It has been found that it provides a very effective technique in machine analysis of speech,[2] particularly as far as pitch extraction of voiced sounds is concerned.

An interpretation of the cepstrum in terms of a general formalism for separation of convolved signals has also been proposed,[3] and new applications of the same technique, again in the field of speech analysis, seem to be possible.[4]

Investigation of the mathematical properties of the cepstrum has not yet been carried out extensively, although some properties have already been shown for both continuous and discrete signals.[5,6]

This work was motivated by the desire both for a better understanding of the mathematical properties of the cepstrum, and for testing Oppenheim's speculation that the cepstrum might be used in a model accounting for pitch perception, at least at the level of psychoacoustics.

The unifying idea of these two points of view is that the cepstrum could synthetically express the temporal features of a periodical waveform (such as the period of a sine

*This work was supported in part by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DA 28-043-AMC-02536(E), and in part by the National Aeronautics and Space Administration (Grant NsG-496).

wave) and, in some cases, be more convenient or suitable for measurements than other functions (such as the simple inverse Fourier transform of the energy spectrum). One might observe that the signal itself is the best expression of its temporal features, but the point is that in many cases the actual waveform of the signal is not known or available. The signal might be added to noise, and we know that, in this case, the autocorrelation function provides a good means for detecting the period of the signal.[7] Suppose now that the only information available about the signal is its short time energy spectrum, $\Phi(t, \omega)$, weighted by some function of ω. [This could be the case physically (for example, compressed band speech transmission, model of the peripheral auditory system, etc.).]

We shall show that there is a large class of weighting functions to which the cepstrum is almost insensitive, whereas the simple inverse Fourier transform is not. Therefore, the cepstrum might, in this case, be a more convenient way of looking at the time properties of the signal.

In the following discussion, before showing this, the analytical properties of the cepstrum will be considered in some generality.

2. Analytical Computation of the Cepstrum

According to definition, the cepstrum can be written

$$c(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \log \Phi(t, \omega)\, e^{j\omega\tau}\, d\omega, \tag{1}$$

where $\Phi(t, \omega)$ is the square of the magnitude of the short-time spectrum of a signal $x(t)$.

We have decided to consider the short-time spectrum rather than the infinite-time spectrum because it is more general and appropriate for a physical interpretation and computation of the cepstrum. It can be defined[8, 9] as

$$F(t, \omega) = \int_{-\infty}^{t} x(\lambda)\, g(t-\lambda)\, e^{-j\omega\lambda}\, d\lambda, \tag{2}$$

where $x(\lambda)$ is the signal, and $g(\lambda)$ is a weighting function that, for the moment, can be considered as the impulse time response of a linear physical system, which satisfies the causality condition: $g(\lambda) = 0,\ \lambda < 0$.

$\Phi(t, \omega)$ in (1) is given by

$$\Phi(t, \omega) = |F(t, \omega)|^2. \tag{3}$$

Incidentally, we notice that the inverse Fourier transform of $\Phi(t, \omega)$, in ω, is the short-time autocorrelation function[9] $\phi(t, \tau)$.

$$\phi(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi(t, \omega)\, e^{j\omega t}\, d\omega. \tag{4}$$

Although $\Phi(t, \omega)$ in (1) has now been defined, Eq. 1 seems to need some more discussion and elaboration. In fact, the integral in (1) is not well defined from a mathematical point of view, the reason being that $\Phi(t, \omega)$ diverges as ω goes to infinity, since $\Phi(t, \omega)$ in general tends to zero as ω tends to infinity. We know, however, that the behavior of $\Phi(t, \omega)$ at infinity affects the behavior of its Fourier transform only around $\tau = 0$; if we decide, therefore, to ignore the cepstrum around $\tau = 0$ (that is, the limit as $\tau \to 0$) we may find a different expression for it which is better defined from a mathematical point of view. Such an expression exists, in fact, and has already been used.[5]

$$c(t, \tau) = -\frac{1}{2\pi j\tau} \int_{-\infty}^{+\infty} \frac{\Phi'(t, \omega)}{\Phi(t, \omega)} e^{j\omega\tau}\, d\omega; \quad \tau \neq 0. \tag{5}$$

To show its validity one can use the generalized Fourier integral[10,11] in order to handle the original integral (1) at very high frequencies; one can further use the theory of distributions and generalized limits.[12,13]

From now on we shall use (5) as the correct analytical expression of the cepstrum. This requires that integral (5) exist. The conditions under which it does exist relate to the class of functions that have a Fourier transform, but we shall not consider the problem in this generality.

We could ask, however, if integral (5) exists for the class of functions that we are interested in. And since, at present, our interest is confined to short-time spectra of periodic signals with a finite number of Fourier coefficients, the answer to this question seems to be affirmative. A sufficient requirement is, for example, that the Fourier transform of the weighting function is rational and has both poles and zeros in the left half-plane, which is in fact the case for the weighting functions that we are going to use.

3. Cepstrum and Autocorrelation of Single Sine Waves Weighted by a Fixed Function

The short-time energy spectrum $\Phi(t, \omega)$ of a single sine wave can be written as follows:

$$\Phi(\Psi, \omega) = \left| \int_0^{\infty} \cos\,(\omega_0 x + \Psi)\, w(x)\, e^{j\omega x}\, dx \right|^2. \tag{6}$$

A very reasonable choice for the weighting function seems to be an exponential such as $w(t) = e^{-\alpha t}$ $(t \geq 0)$, because of the resulting simplicity for both the analytical computation and the physical measurement of $\Phi(\psi, \omega)$. [By physical measurement of the energy spectrum we mean what is referred to in the literature as "simultaneous spectral analysis."[12]]As far as this is concerned, the choice of weighting function which we have made corresponds, for $\omega \gg \alpha$, to the use of symmetrical bandpass filters, all having the same equivalent lowpass filter, whose system function is $W(\omega) = 1/(\alpha+j\omega)$. The short-time energy spectrum, in fact, can be computed simply by taking the envelope of the output of such a set of bandpass filters, all in parallel, the signal being the input. The parameter ω in $\Phi(t,\omega)$, in this case, has to be interpreted as the center frequency of our bandpass filters.

Analytically, in case of a single sine wave and according to (6) we have

$$\Phi(\phi,\omega) = \frac{1}{\omega_0^2}\cdot\frac{\left(\frac{\omega}{\omega_0}\right)^2\cos^2\psi + \left(\frac{\alpha}{\omega_0}\cos\psi - \sin\psi\right)^2}{\left[\left(\frac{\omega}{\omega_0}+1\right)^2 + \left(\frac{\alpha}{\omega_0}\right)^2\right]\left[\left(\frac{\omega}{\omega_0}-1\right)^2 + \left(\frac{\alpha}{\omega_0}\right)^2\right]} \tag{7}$$

$$\phi(\psi,\tau) = \frac{1}{4\omega_0}\cdot\frac{e^{-\frac{\alpha}{\omega_0}\omega_0\tau}}{\frac{\alpha}{\omega_0}\left[1+\left(\frac{\alpha}{\omega_0}\right)^2\right]}\left\{\left[1 - 2\frac{\alpha}{\omega_0}\sin\psi\cos\psi + 2\left(\frac{\alpha}{\omega_0}\right)^2\cos^2\psi\right]\cdot\cos\omega_0\tau\right.$$
$$\left. - \left[\frac{\alpha}{\omega_0}\cos^2\psi + 2\left(\frac{\alpha}{\omega_0}\right)^2\sin\psi\cos\psi - \frac{\alpha}{\omega_0}\sin^2\psi\right]\sin\omega_0\tau\right\} \qquad \tau \geq 0 \tag{8}$$

$$c(\psi,\tau) = \frac{\omega_0}{\omega_0\tau}\left\{2\,e^{-\frac{\alpha}{\omega_0}\omega_0\tau}\cos\omega_0\tau - e^{-\left|\frac{\alpha}{\omega_0}-\tan\psi\right|\omega_0\tau}\right\} \qquad \tau \geq 0. \tag{9}$$

Expressions (3) and (4) have been plotted in Fig. X-1 for the value $\frac{\alpha}{\omega_0} = 0.25$. Only two different values of the phase ψ have been considered, namely $\psi = 0$ and $\psi = \frac{\pi}{2}$, which correspond to some sort of limiting case.

The abscissa of the first positive maximum for both $\phi(\psi,\tau)$ and $c(\psi,\tau)$ clearly corresponds to the period of the signal. We notice that the two different values of the phase for such an abscissa produce a shift that decreases (relatively to the period) as ω_0 increases. We might call this relative shift the "phase effect" on the cepstrum and autocorrelation of a single weighted sine wave.

Both $\phi(t,\tau)$ and $c(t,\tau)$ have been computed, and the phase effect has also been considered, in the cases of the sum of two and three sine waves of different frequencies. These computations show that both $\phi(t,\tau)$ and $c(t,\tau)$ preserve many of the qualitative features of the periodic signal. In particular, both $\phi(t,\tau)$ and $c(t,\tau)$ have a maximum corresponding to the period T of the signal, provided that T is of the order of magnitude

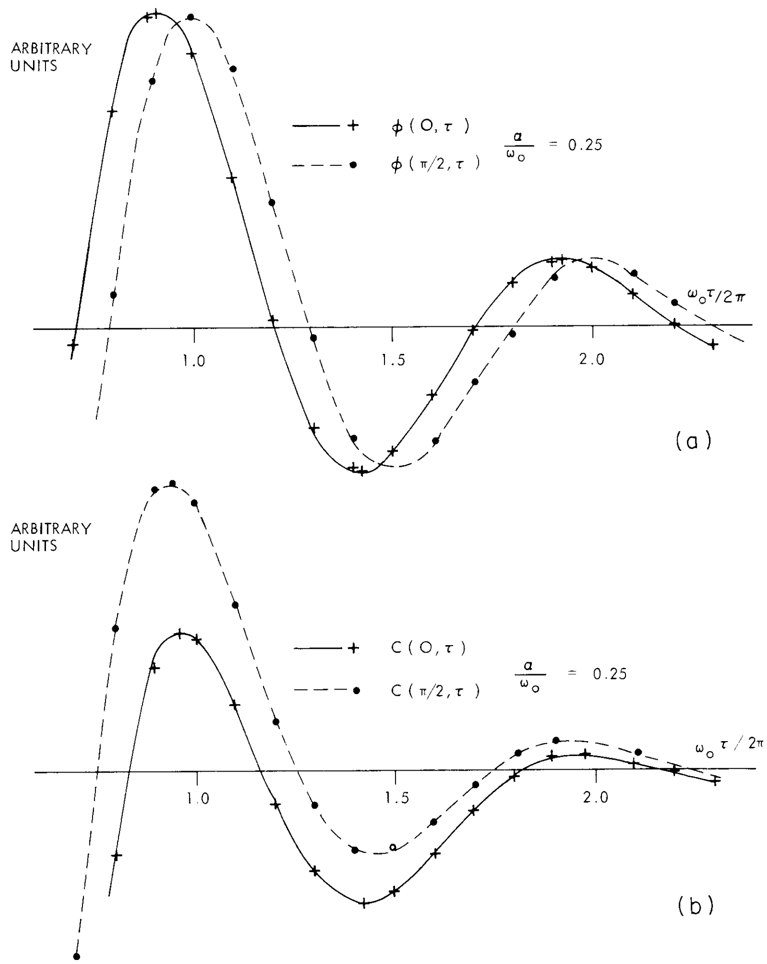

Fig. X-1. (a) Short-time autocorrelation function of a single sine wave (ω_0) with a fixed weighting function. Only two values of the phase ψ (which accounts for the parameter t) have been considered, $\psi = 0$ and $\psi = \frac{\pi}{2}$, corresponding to limiting cases. The initial maximum ($\tau = 0$) has been omitted for scale reasons.
(b) Cepstrum of a single sine wave calculated for the same conditions as in (a). Since the cepstrum tends to ∞ as τ approaches zero, this part of the graph has been omitted for scale reasons.

or less than the "length" of the weighting function. We already know that the autocorrelation function preserves some of the fundamental temporal features of a periodic signal. On the basis of the examples considered, the conjecture can be made that the cepstrum, too, reveals some of the main temporal features of a periodic signal.

4. Use of an ω-dependent Weighting Function

We shall now extend the computation of the short-time energy spectrum $\Phi(t, \omega)$ to the case of an ω-dependent weighting function $w(t, \omega)$ $(t > 0)$. This case is physically interesting because it corresponds to the evaluation of the short-time energy spectrum obtained by means of a set of bandpass filters that do not have the same system function (except for the center frequency), unlike the case of a fixed weighting function. For instance, if the weighting function $e^{-k\omega t}$ $(t>0)$ is used, a set of bandpass filters results, all having the same Q factor, instead of the same band, B, as in the case of the function $e^{-\alpha t}$ $(t>0)$ which was previously considered. The function $e^{-k\omega t}$ $(t>0)$ itself does not correspond to a lowpass filter, but formally it can be considered as such, and used for generating the bandpass filters. The parameter ω, in fact, corresponds physically to the center frequency of these bandpass filters. In some cases we call it ω_c, in order to avoid possible confusion.

If the system functions of the bandpass filters in the two cases ($e^{-\alpha t}$ and $e^{-k\omega t}$) are compared, we have

<u>Case a</u> $$H_{BP}(\omega) = \frac{1}{\alpha + j(\omega-\omega_c)}; \qquad \alpha \ll \omega_c, \quad \omega > 0$$

<u>Case b</u> $$H_{BP}(\omega) = \frac{1}{k\omega_c + j(\omega-\omega_c)}; \qquad k \ll 1, \quad \omega > 0.$$

In case a the band and the maximum amplitude do not depend upon ω_c; furthermore, if the input is a single sine wave, the total spectral energy $\int_{-\infty}^{+\infty} \Phi(t,\omega)\, d\omega$ is a constant, that is, it does not depend on the frequency ω_x of the input (provided that $\alpha \ll \omega_x$).

In case b we have $Q = \text{const.} \approx 1/2k$; the maximum amplitude of the system function is inversely proportional to the center frequency; and the total spectral energy, when the input is a sine wave of frequency ω_x, is inversely proportional to ω_x. Therefore, we can say that, in case a, a "correctly weighted" energy spectrum $[\Phi_a(t,\omega)]$ is generated, whereas, in case b, the energy spectrum $[\Phi_b(t,\omega)]$ clearly has an attenuation at the higher frequencies.

One can show, however, that by multiplying $\Phi_b(t,\omega)$ by a factor proportional to ω, a correctly weighted energy spectrum is restored, in the sense that the total energy (the input being a single sine wave) becomes a constant as in case a.

We assume therefore that $\Phi_b(t,\omega)$, "compensated" by the factor ω, is the correctly weighted energy spectrum of case b.

Let us, however, compute the cepstrum $c(t,\tau)$ and the simple Fourier transform of $\Phi_b(t,\omega)$ without any compensation. The simple Fourier transform of $\Phi_b(t,\omega)$ will now be called $\sigma(t,\tau)$, in order to distinguish it from the short-time autocorrelation function previously considered (4).

We shall show that, in the case of the cepstrum, there is no substantial difference when $\Phi_b(t,\omega)$ is or is not compensated by the factor ω, while in the case of $\sigma(t,\tau)$ there is.

The mathematical expressions for a single sine wave are the following (the phase ψ accounting for the parameter t):

$$\sigma(0, \tau) = \frac{e^{-k\omega_{0k}\tau}}{4k\omega_0}[\cos \omega_{0k}\tau - k \sin \omega_{0k}\tau] \qquad (\tau \geq 0) \tag{10}$$

$$\sigma\left(\frac{\pi}{2}, \tau\right) = \frac{e^{-k\omega_{0k}\tau}}{4k(1+k^2)\,\omega_0}[\cos \omega_{0k}\tau + k \sin \omega_{0k}\tau] \qquad (\tau \geq 0) \tag{11}$$

$$c(0, \tau) = \frac{1}{\tau}\left[-1 + 2\,e^{-k\omega_{0k}\tau} \cos \omega_{0k}\tau\right] \qquad (\tau > 0) \tag{12}$$

$$c\left(\frac{\pi}{2}, \tau\right) = \frac{1}{\tau}\left[2\,e^{-k\omega_{0k}\tau} \cos \omega_{0k}\tau\right] \qquad (\tau > 0), \tag{13}$$

where

$$\omega_{0k} = \frac{\omega_0}{1 + k^2}.$$

We do not show the graphs of these expressions since they are not substantially different from those of Fig. X-1 for a reasonable choice of k, such as 0.05. It can be shown, however, that in this case the relative shift of the maxima with the phase (which we have called the "phase effect") does not depend on ω_0.

In the case of two sine waves with the same amplitude,

$$x(t) = \cos(\omega_1 t + \psi) + \cos(q\omega_1 t + \psi + \gamma); \qquad q > 1,$$

the mathematical expressions, for $\psi = 0$, $\gamma = \pi$, are

$$\Phi_b(\omega) \equiv \frac{y^2}{[(y-1)^2+k^2][(y+1)^2+k^2][(y-q)^2+k^2q^2][(y+q)^2+k^2q^2]} \tag{14}$$

$$\frac{\Phi_b'(\omega)}{\Phi_b(\omega)} \equiv y - \frac{1}{y}\left\{\frac{y-1}{(y-1)^2+k^2} + \frac{y+1}{(y+1)^2+k^2} + \frac{y-q}{(y-q)^2+k^2q^2} + \frac{y+q}{(y+q)^2+k^2q^2}\right\}, \tag{15}$$

where

$$y = \frac{\omega}{\omega_{1k}}; \qquad \omega_{1k} = \frac{\omega_1}{1 + k^2}; \qquad \psi = 0, \qquad \gamma = \pi.$$

$$\{\sigma(\tau)\}_{\substack{\psi=0\\ \gamma=\pi}} = \frac{e^{-k\omega_{1k}\tau}}{\omega_1}\left\{\left[\frac{1}{4k}-\frac{c1}{D}\right]\cos\omega_{1k}\tau - \left[\frac{1}{4}-\frac{s1}{D}\right]\sin\omega_{1k}\tau\right\}$$

$$+\frac{e^{-kq\omega_{1k}\tau}}{\omega_1}\left\{\left[\frac{1}{4kq}-\frac{c2}{D}\right]\cos q\omega_{1k}\tau - \left[\frac{1}{4q}-\frac{s2}{D}\right]\cdot\sin q\omega_{1k}\tau\right\}; \quad \tau \geq 0, \tag{16}$$

where

$$c1 = 2k(q^2+1) + k(1-k^2)(q^2-1)$$

$$s1 = 2k^2(q^2+1) - (q^2-1)(1-k^2)$$

$$c2 = 2k(q^2+1) - k(1-k^2)(q^2-1)$$

$$s2 = 2k^2(q^2+1) + (q^2-1)(1-k^2)$$

$$D = (q^2-1)^2\ (1-k^2) + 4k^2(q^2+1)^2.$$

$$\{c(\tau)\}_{\substack{\psi=0\\ \gamma=\pi}} = \frac{1}{\tau}\left\{-1+2\,e^{-k\omega_{1k}}\cos\omega_{1k}\tau + 2\,e^{-kq\omega_{1k}}\cos q\omega_{1k}\tau\right\}; \quad \tau > 0. \tag{17}$$

These expressions, for the value $q = 1.5$, have been plotted in Figs. X-2 and X-3, while the signal has been plotted in Fig. X-4, for purposes of comparison.

We observe that the energy spectrum is attenuated at high frequencies, as expected, and that its Fourier transform $\sigma(\tau)$ has, in fact, lost its strict similarity to the signal waveform. The cepstrum, on the other hand, has a waveform very similar to that of the signal. This can easily be recognized if the signal is shifted along the time axis until one of its major maxima falls at the origin.

What we call "similarity" mainly consists in having the same number of peaks with approximately the same relative amplitudes for an interval of time (or τ) of the order of magnitude of the period. In computing the cepstrum, apparently, a correctly weighted energy spectrum is in some way automatically restored. We might understand, and at the same time generalize, this effect by looking at the manner in which the cepstrum would have changed if a correctly weighted energy spectrum $\Phi(t,\omega)$ had been used instead of $\Phi_b(t,\omega)$. In our case, $\Phi(t,\omega) \equiv \omega\cdot\Phi_b(t,\omega)$. Since

$$\frac{\Phi'(t,\omega)}{\Phi(t,\omega)} = \frac{1}{\omega} + \frac{\Phi_b'(t,\omega)}{\Phi_b(t,\omega)},$$

the "correct" cepstrum would differ from the cepstrum that we had computed only by a

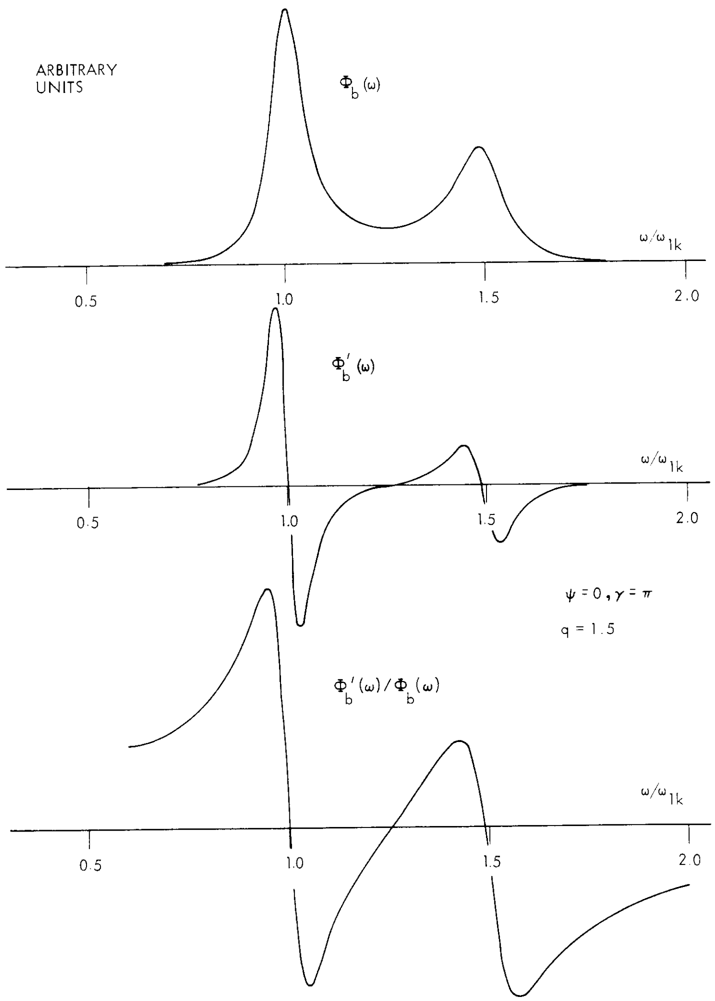

Fig. X-2. Short-time energy spectrum, $\Phi_b(\omega)$, of the sum of two sine waves with the same amplitude and frequency ratio 1.5. The weighting function is the ω-dependent function $e^{-0.05\omega t}$ $(t > 0)$. The phases ψ and γ account for both the phase between the two sine waves and their shift (t) with respect to the weighting function. Also shown is the effect of the transformation from $\Phi_b(\omega)$ to $\Phi_b'(\omega)/\Phi_b(\omega)$, which was used for the analytical computation of the cepstrum.

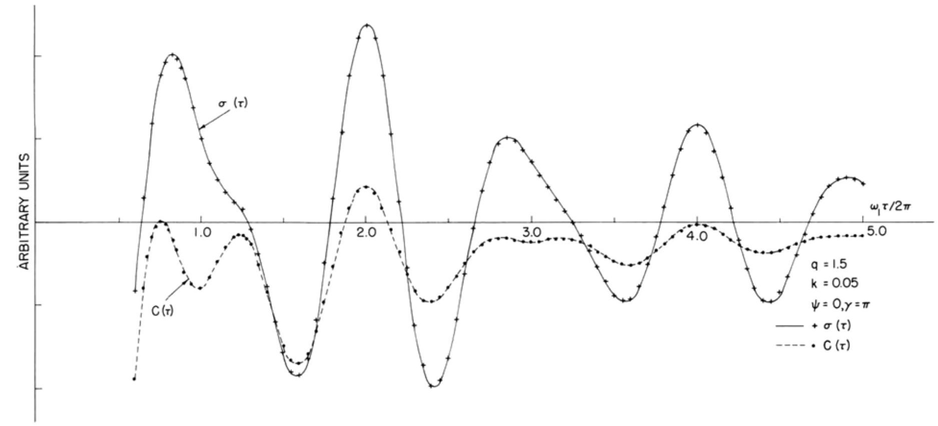

Fig. X-3. Cepstrum $c(\tau)$ and the function $\sigma(t)$ calculated when the signal is the sum of two sine waves and for the same conditions as in Fig. X-2. $\sigma(\tau)$ is the inverse Fourier transform of $\Phi_b(\omega)$. As τ approaches zero the functions $\sigma(\tau)$ and $c(\tau)$ tend to an absolute positive maximum and to ∞, respectively. This part has been omitted for scale reasons.

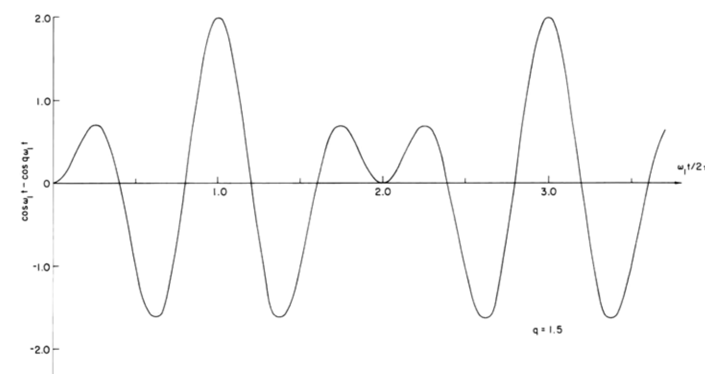

Fig. X-4. The signal (sum of two sine waves) to which Figs. X-2 and X-3 refer. It is shown for comparison with $\sigma(\tau)$ and $c(\tau)$ in Fig. X-3. The aim is to show the similarity between the cepstrum and the signal waveform after the latter has been properly shifted along the time axis.

term -0.5 to be added to the terms that multiply $1/\tau$ in (17). If we do this, the picture of $c(\tau)$ does not change in any substantial way.

This effect is far more general. In fact, if we have to weight (compensate) our energy spectrum by any function of the kind

$$A\, e^{p\omega^s} \cdot \omega^r \qquad A, p, s, r \qquad \text{real numbers,}$$

instead of by ω, we can easily see that the cepstrum would still not change in any substantial way, provided that $|r|$ and $|s|$ are of the order of magnitude of 1 or less. Obviously, $\sigma(\tau)$ does not have the same property.

This property of the cepstrum is physically very interesting because it greatly enlarges the class of bandpass filters by means of which the energy spectrum of a signal can be physically computed and the time properties of the signal directly (that is, via the cepstrum) recovered.

5. Conclusion

We have shown that (a) One can conjecture that the cepstrum is an expression that, similarly to the autocorrelation function, reveals some of the main temporal features of a periodic signal, and (b) The cepstrum has the original property of being almost insensitive to a wide class of transformations of the short-time energy spectrum of the signal, $\Phi(t, \omega)$.

We have also commented briefly on the physical significance and interest of these results.

Both statements a and b might lead to more definite and quantitative conclusions and, therefore, they implicitly indicate the direction of further work which seems to be worth doing.

I am deeply indebted to Professor Alan V. Oppenheim for helping me choose this problem and for his continued criticism and encouragement.

G. Gambardella

References

1. B. P. Bogert, M. J. R. Healy, and J. W. Tukey, Proceedings of the Symposium on Time Series Analysis, edited by M. Rosenblatt (John Wiley and Sons, Inc., New York, 1963), Chap. 15, pp. 203-243.
2. A. M. Noll, "Cepstrum Pitch Determination," J. Acoust. Soc. Am. 41, 293-303 (1967).
3. A. V. Oppenheim, "Nonlinear Filtering of Convolved Signals," Quarterly Progress Report No. 80, Research Laboratory of Electronics, M. I. T., January 15, 1966, pp. 168-175.
4. A. V. Oppenheim, communication to the 74th Meeting of the Acoustical Society of America, New York, April 1967.
5. R. W. Schafer, "A New Approach to Echo Removal," Quarterly Progress Report No. 84, Research Laboratory of Electronics, M. I. T., January 15, 1967, pp. 194-201.
6. A. V. Oppenheim and R. W. Schafer, "Discrete Analysis of Homomorphic Deconvolution," Quarterly Progress Report No. 85, Research Laboratory of Electronics, M. I. T., April 15, 1967, pp. 231-234.
7. Y. W. Lee, Statistical Theory of Communication (John Wiley and Sons, Inc., New York, 1960).
8. M. R. Schroeder and B. S. Atal, "Generalized Short-Time Power Spectra and Autocorrelation Functions," J. Acoust. Soc. Am. 34, 1678-1683 (1962).

9. J. L. Flanagan, Speech Analysis, Synthesis and Perception (Academic Press, Inc., New York, 1965).

10. E. R. Titchmarsh, Theory of Fourier Integrals (Oxford University Press, London, 1948).

11. N. Weiner, Fourier Integral (Cambridge University Press, London, 1932).

12. A. Papoulis, The Fourier Integral and Its Applications (McGraw-Hill Publishing Company, New York, 1962).

13. H. Breverman, Distributions, Complex Variables and Fourier Transforms (Addison-Wesley Publishing Company, Inc., South Reading, Mass., 1965.).