HOMOMORPHIC MODULATION SPECTRA

Les Atlas, Qin Li*, and Jeffrey Thompson***
Department of Electrical Engineering, University of Washington
Seattle, WA 98195-2500, USA
**Visiting at Fraunhofer Institute IIS-A
Erlangen, Germany

ABSTRACT

Physical evidence points to the importance of a concept called "modulation frequency." This dimension exists jointly with standard Fourier or acoustic frequency. Thus, akin to other time-varying analysis, we seek a two-dimensional representation, the "modulation spectrum," where the first dimension is the well-known acoustic frequency and the second dimension is modulation frequency. We describe some deficiencies in previous discussions of this concept, and then address those deficiencies via a homomorphic approach. We also reduce previous difficulties in homomorphic demultiplication by integrating this processing into modulation spectra and, in particular, show how assumption of analytic and relatively narrowband sub-bands allows more accurate and practical use of homomorphic demultiplication. Lastly, we show how an unambiguous demultiplication concept is only consistent with complex modulator envelopes. The assumption of complex envelopes is necessary for accurate modulation spectral analysis and filtering.

1. INTRODUCTION

Zadeh first proposed that a separate dimension of modulation frequency could supplant the standard concept of system function frequency analysis [1]. His proposed two-dimensional system function had two separate frequency dimensions—one for standard frequency and the other a transform of the time variation. This two-dimensional bi-frequency system function was only defined, but was not analyzed. Kailath followed up nine years later [2] with the first analysis of this joint system function. More recently, Gardner (e.g. [3,4]) greatly extended the concept of joint frequency analysis for cyclostationary systems. However, transforms that are used in compression and for many pattern recognition applications usually have a need for invertibility which cyclostationary analysis does not provide. Indeed, the notion of filtering or coding in modulation offers potential for new forms of noise removal, source separation, signal modification, and efficient coding.

2. BACKGROUND

For further progress to be made in the understanding and applications of modulation spectra, a well-defined foundation for the concept of modulation frequency analysis/synthesis needs to be established. By "acoustic frequency" we mean an exact or approximate conventional long-time or short-time Fourier decomposition of a signal. "Modulation frequency" is the dimension that this section will begin to strictly define.

The notion of modulation frequency is quite well understood for many signals that are generated synthetically. A simple case consists of an amplitude modulated fixed-frequency carrier

$$s(t) = m(t)\cos\omega_c t \qquad (1)$$

where the modulating signal $m(t)$ is real and non-negative and has an upper frequency band limit suitable for its perfect and easy recovery from $s(t)$. It is straightforward that the modulation frequency for this signal should be the Fourier transform of the modulating signal only

$$M(e^{j\omega}) = F\{m(t)\} = \int_{-\infty}^{\infty} m(t)e^{-j\omega t}\,dt \qquad (2)$$

But what is a two-dimensional transform of acoustic versus modulation frequency? Namely, how should this signal be represented as the two-dimensional distribution $P(\eta,\omega)$, where η is modulation frequency and ω is acoustic frequency? Moreover, what should happen for an arbitrary signal, say speech or audio?

2.1. Previous Definition

A modulation transform or modulation filtering can be performed via a two-dimensional transformation [5,6], which is computed through a three-step process. The process begins with a base transform which transforms input signal to a time-frequency representation (e.g. short-time Fourier Transform). Then a nonlinear detection operation (e.g. magnitude square or Hilbert envelope) is performed on each base transform sub-band to extract signal envelopes. A second transform is then performed across time within each sub-band. If the base transform is a Fourier transform the whole process yields an acoustic frequency versus modulation frequency representation that is usually referred as "modulation spectrum." Acoustic frequency commonly corresponds to the frequency axis of the first transform and modulation frequency corresponds to the independent variable of the second transform.

This transform, followed by modification or quantization, and then followed by its inverse, is the approach representing an analysis/synthesis or modulation filtering system. Closely related are filterbank or vocoding approaches (e.g. [7]). In all of these cases, an analogous operation for base transform phase is not defined. For reconstruction from modulation spectra, original or quantized input signal phase is directly combined with inverse transformed modulation spectrum magnitude, sometimes with added approximating iterations (e.g. [8]).

While perfect reconstruction is possible for some of the proposed modulation spectral transform approaches (e.g. [5,6]), as we often observe empirically, inverting *modified or significantly quantized* modulation spectra produces mild to profound undesired artifact. The detailed reasons for these artifacts are multiple and beyond the scope and intent of this paper. However, much of the distortion is related to the above lack of a modulation spectral transformation or filtering operation on signal phase.

The main claim of this paper is that the modulation filtering problem, posed as filtering on the magnitudes resulting from a modulation transform, is conceptually incomplete. The explicit assumption of a modulation (multiplying) process producing the observed signal and an explicit assumption of superposition are missing. As we will show, inclusion of these two assumptions via a homomorphic approach solves multiple existing conceptual problems with modulation spectra and also suggests an approach for modulation spectral phase. Furthermore, the previous assumption of a real and non-negative modulator, even for real input signals, is demonstrated to be incomplete.

Our problem statement and analysis below represents a rediscovery of previous homomorphic demultiplication work by Oppenheim [9] and Stockham [10]. Nevertheless, we substantially extend this previous work by integrating it into the problem of modulation spectral analysis/synthesis. In particular, the use of a complex one-sided (in frequency) filterbank for the base transform, before homomorphic deconvolution, allows for a substantially improved fit of an analytic signal to the complex logarithm needed in homomorphic demultiplication. Furthermore, we show that the structure of a homomorphic analytic signal is constrained in useful ways [11] yet still allows the important and common possibility of complex modulation.

2.2. Modulation Model

A fundamental concept for the study of modulation spectra is the notion of a modulator signal multiplying (modulating) a carrier signal, which results in the sub-band signal which we observe. Consistent with previous discussions of modulation spectral analysis, we make these assumptions for each sub-band:

1. $s(t)$ is the observed signal from one complex frequency sub-band or transform index.
2. $c(t)$ is a narrowband complex high frequency "carrier."
3. $m(t)$ is a positive real "modulator" with no significant frequency content close to or above the carrier frequency.

Assumption 3, specifically of a positive real modulator, is, to the best of our knowledge, implicit in all previous published descriptions of modulation spectra. However, as will be seen, this assumption is too restrictive and inconsistent with a goal of true demultiplication.

2.3. Homomorphic Demultiplication

In the 1960's, well before current interest in modulation spectral analysis developed, the question of decomposing modulation in systems satisfying generalized superposition were studied in general by Oppenheim beginning with [9]. This general formulation became popular, and still is, for the different, but dual, problem of homomorphic deconvolution. The first to apply this general formulation to decomposing multiplication was Stockham beginning with [10]. This is the same Stockham who became known for his work on the enhancement of the Caruso and Watergate tapes [12], which were applications of blind deconvolution and not demultiplication.

The review in this section is based directly upon the work of Oppenheim *et al* [13]. Applying the general homomorphic system to modulation results in a homomorphic system P with superposition rule

$$P\left\{s_1(t)\cdot s_2(t)\right\} = P\left\{s_1(t)\right\} + P\left\{s_2(t)\right\} \tag{3}$$

and with homogeneity rule

$$P\left\{s(t)^c\right\} = cP\left\{s(t)\right\} \tag{4}$$

As shown by Oppenheim [14], any homomorphic system can have a canonic representation which is a cascade of a general transformation ϕ , a linear system T , and the inverse transformation ϕ^{-1} . For the above homomorphic system for multiplication, the transformation is $P = \log$ which in general is required to be a complex log.

As noted by Oppenheim *et al* [13], an unconstrained complex logarithm does not produce a unique mapping $\log s(t) = \tilde{s}(t)$. Uniqueness can be obtained by constraining the complex log to be a principal value; however, this definition will violate the additivity of equation (3). Oppenheim *et al* [13] worked around this problem by defining a clever yet complicated procedure which requires knowledge of a continuous signal for more than an instant of time.

2.4. Homomorphic Demultiplication as a Detector

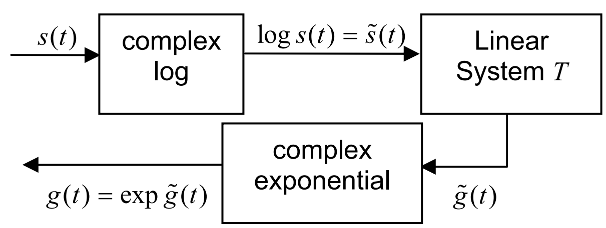

Figure 1. Complex logarithm homomorphic system for multiplication.

For our goal of modulation spectral analysis, we can update the above system by assuming that figure 1 serves as a detector for each channel (sub-band) of a filterbank or transform index of a short-time transform. Identical detectors can be used for each sub-band. Furthermore, as is the goal of modulation spectral analysis or filtering, the linear system T should involve a transform pair. We thus propose the approach below.

We start with a real bandlimited input (e.g. one real sub-band output of a filterbank or short-time Fourier transform) for $s(t)$. An analytic signal $\hat{s}(t) = \left|\hat{s}(t)\right| e^{j\angle\hat{s}(t)}$ is then formed via, for example, Hilbert transforms. Alternatively, the analytic signal can come directly from a positive frequency sub-band output of a one-sided (positive frequency index) complex filterbank. Recalling our modulation model, and assuming the carrier is a unimodular phase signal, namely $\left|\hat{c}(t)\right| = 1$, we end up inferring a decomposition of the input analytic signal

$$\hat{s}(t) = m(t)\cdot\hat{c}(t) = m(t)\cdot\exp\{j\angle\hat{c}(t)\} \quad (5)$$

where by Bedrosian [15], $\hat{c}(t)$ must be the analytic signal of the carrier. Futhermore, from [11], the homomorphic analytic signal $\log\hat{s}(t)$ is minimum phase and not subject to the non-uniqueness problems of the complex logarithm which plagued earlier efforts for complex logs of more general signals.

3. SIMULATION RESULTS

3.1 Real Modulator

For analytic signals, we first assume that

$$\tilde{s}(t) = \log\hat{s}(t) = \log\left|\hat{s}(t)\right| + j\angle\hat{s}(t) = \tilde{m}(t) + \tilde{c}(t) \quad (6)$$

where $\tilde{m}(t) = \log\left|\hat{s}(t)\right|$ wholly represents the modulator and $\tilde{c}(t) = j\angle\hat{s}(t) = j\phi(t)$ wholly represents the carrier. This assumption is consistent with assumption 3 in above section 2.2 and is commonly assumed in modulation spectral analysis.

Figure 4 shows the effect of a high-pass filter for the linear transformation T. In this example, we want to attenuate the frequency components below 250 Hz in a 600 Hz band-limited real modulator which previously multiplied a 3500 Hz carrier. The following steps were taken sequentially:

1. Perform equation 6.
2. High-pass filter $\tilde{m}(t) = \log\left|\hat{s}(t)\right|$. No change to $\tilde{c}(t) = j\angle\hat{s}(t)$.
3. Perform the inverse of equation 6.

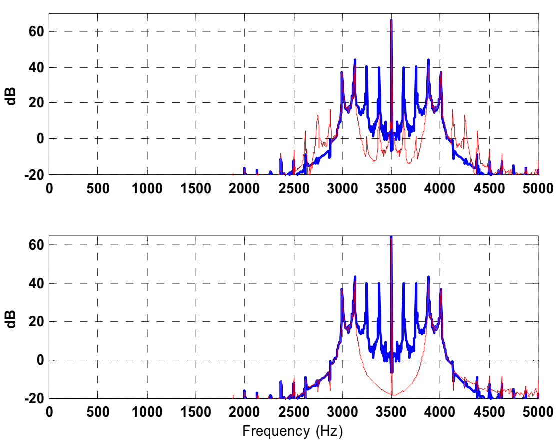

Figure 2. High-pass modulation filtering result in the frequency domain. For the top panel, the thick blue line represents the spectrum of the unfiltered input signal, and the thin red line is the spectrum of the signal processed with equation 6. About 30 dB attenuation is achieved for the low frequencies of the modulator. For the bottom panel, which is used for comparison, the thick blue line represents the same spectrum of the unfiltered input signal, and the thin red line is the spectrum of a signal with pre-filtered modulator.

The spectrum of the filtered signal is plotted in the upper panel of figure 2. For comparison, we synthesized another signal by directly modulating the same carrier with a signal that is originally filtered by the same high-pass filter aforementioned; the resulting spectrum is plotted in the lower panel of figure 2.

We can observe that for the above proposed modulation filtering steps, about 30-35 dB attenuation was achieved at the low frequency bands of the modulator, while about 40-50 dB attenuation was achieved for directly filtering on modulator. This difference in modulation attenuation is insignificant for most speech and audio applications, because the achieved 30 dB attenuation is already substantially beyond the ear's perception of modulation [16].

3.2 A Real Modulator is Too Restrictive

It is commonly assumed that the modulator is positive and real and hence has perfectly symmetric magnitude about the carrier. This assumption gets even more restrictive and impractical for modulation spectra, where $s(t)$ or $\hat{s}(t)$ come from a filterbank or equivalent transform sub-band. In that case, symmetry of the modulator magnitude is considered with respect to the center of the sub-band. Moreover, even when a carrier is centered within a sub-band, side-lobes in neighboring sub-bands will not necessarily be centered. Thus, for an arbitrary signal input, this rare symmetric case requires that the input signal carrier be synchronized with the sub-band center frequency, that the modulator be symmetric about this carrier, and that this symmetry holds in all other sub-bands. This concern about assumed yet incorrect symmetry also hold for previous approaches to modulation spectra, independent of homomorphic processing. Thus, for both theoretical evaluation and for practical cases, the notion of a complex envelope is necessary. While this broader assumption would normally cause an ambiguity between modulator and carrier phase, homomorphic demultiplication, for analytic input signals, removes this ambiguity.

Figure 3 shows a homomorphic demultiplication system which requires no assumption of a real envelope. Poletti [11] used the derivative of the log analytic signal quantity to define instantaneous complex frequency. Our need for a complex modulator is analogous. For a real and positive modulator, the system in figure 3 simplifies to $\tilde{m}'(t) = \log\left|\hat{s}'(t)\right|$. For a complex modulator, it generalizes appropriately. If the linear system T is, for example, a Fourier transform, $\left|X(0)\right|$ represents the carrier frequency. The remaining output magnitude then relates to the Fourier transform of a complex (asymmetric in frequency) modulating envelope. Filtering in modulation is possible via operations on output magnitude and phase, followed by reconstruction by an inverse linear transform, an integral or accumulator (to invert the derivative) and a complex exponential.

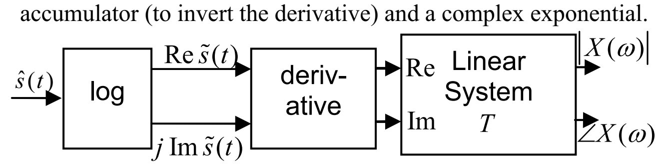

Figure 3. Detector for proposed homomorphic demultiplication system. For a complete modulation spectral analysis system, one of these detectors would be used for each sub-band. For a complete analysis/synthesis system, the inverse of figure 5 would also be required for each sub-band.

Figure 4 demonstrates that the above system, depicted in figure 3, and its inverse can indeed filter complex envelopes. This example shows that all modulation components can be removed only if the complex envelope assumption is made. Similar results would be expected for other input signal possibilities and other types of modulation filtering.

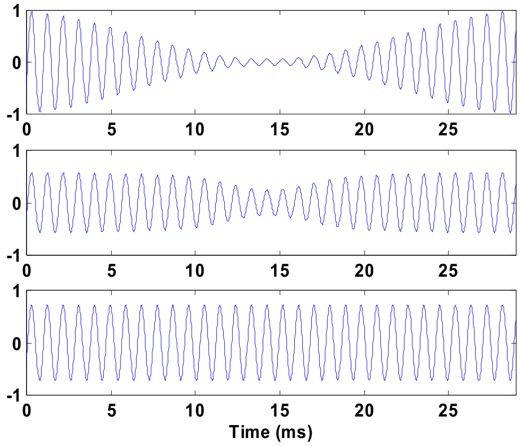

Figure 4. Time domain result of modulation spectral filtering. The top panel shows the input signal, a 1076 Hz carrier multiplied by a 32 Hz modulator and then passed through sub-band analysis, which introduced asymmetries in the modulator. The middle panel shows the resulting reconstructed signal with the assumption of a real envelope and low-pass filtering (removal) of that envelope. The bottom panel shows the resulting reconstructed signal with the assumption of a complex envelope and low-pass filtering (removal) of that envelope via the system in the above figure 3.

While previous work in modulation spectral analysis/synthesis (e.g. [5-7]) had simply substituted the original unfiltered signal's phase (or equivalent) in with the linear filtered signal's magnitude, it can be seen that linear filtering of the complex time derivative is instead required. This result also explains why previous approaches, when used for substantial modulation spectral filtering, produced unwanted artifact.

4. CONCLUSIONS

We have connected the recent concepts of modulation spectral analysis/synthesis to earlier concepts in homomorphic demultiplication. A first result, is that homomorphic demultiplication, when considered from the point of an analytic signal or complex sub-band output, can take advantage of the properties of a homomorphic analytic and narrowband signal. A second result is that, as shown in our simulations, filtering in modulation frequency produces essentially the same result as if the filtered signal was originally produced. As shown in our simulations and expected from the nonlinearity of homomorphic processing, the drawback of our approach is that modulation spectral filtering is in practice limited to modulation frequency stop-band rejection of about 25 dB. For many applications, such as speech and music modification or coding, where humans are known to have a smaller dynamic range of perception of modulation, the effects of this limitation are not expected to be audible. Lastly, and most importantly, we have shown that previous assumptions of a real and positive modulation envelope are incomplete. The homomorphic demodulation system proposed in figure 3 does not require this incomplete assumption.

While the results in this paper are for continuous-time signals, similar results, after accounting for aliasing and circularity, should hold for discrete time. Recursive implementations should be possible for homomorphic demultiplication, thus removing the implementation challenges of the complex log and exponential.

We would like to acknowledge Drs. Sascha Disch and Juergen Herre for their helpful discussions. This research was supported by the Defense Advanced Research Projects Agency, the Fraunhofer Institute IIS-A, and the German-American Fulbright Commission.

References

[1] Zadeh, L. "Frequency analysis of variable networks," *Proc. IRE*, 38(3), 291-299, 1950.
[2] Kailath, T. "Channel characterization: Time-variant dispersive channels." pp. 95-123 in *Lectures on Communication System Theory*, Baghdady, E. (Ed.), McGraw-Hill, New York, NY, 1961.
[3] Gardner, W. *Statistical Spectral Analysis: A Non-probabilistic Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1987.
[4] Gardner, W. "Exploitation of spectral redundancy in cyclostationary signals." *IEEE Signal Processing Magazine*, 14-36, April 1991.
[5] L. E. Atlas and M. S. Vinton, "Modulation frequency and efficient audio coding," in *Proceedings of the SPIE The International Society for Optical Engineering*, 2001.
[6] J. K. Thompson and L. E. Atlas, "A Non-Uniform Modulation Transform for Audio Coding with Increased Time Resolution," in *Proceedings of IEEE ICASSP*, Hong Kong, 2003.
[7] Drullman, R., Festen, J. & Plomp, R., "Effect of temporal envelope smearing on speech reception," *J. Acoust. Soc. Am.* 95, 1053-1064, 1994.
[8] D. Griffin & J. Lim, "Signal Estimation from Modified Short-Time Fourier Transform," *IEEE Trans. ASSP*, ASSP-32, pp. 236-243, April 1984.
[9] A.V. Oppenheim, "Superposition in a class of nonlinear systems," *1964 IEEE Int. Conv. Rec.*, pt. 1, pp. 171-177.
[10] T.G. Stockham, "The Application of Generalized Linearity to Automatic Gain Control," *IEEE Trans. Audio and Electroacous.*, vol. AU-16, No. 2, June, 1968.
[11] Mark A. Poletti, "The Homomorphic Analytic Signal," *IEEE Trans*. SP, Vol. 45, August 1997, pp. 1943-1953.
[12] T. G. Stockham, T. M. Cannon, and R. B. Ingebreston. "Blind deconvolution through digital signal processing," *Proc. IEEE*, pp. 678–692, Apr. 1975.
[13] A.V. Oppenheim, R.W. Schafer, T.G. Stockham, "Nonlinear Filtering of Multiplied and Convolved Signals," *IEEE Trans. Audio and Electroacous.*, vol. AU-16, No. 3, September, 1968.
[14] A.V. Oppenheim, "Superposition in a class of non-linear systems," Research Lab. of Electronics, MIT Tech. Rept. 432, March 31, 1965.
[15] E. Bedrosian, "A Product Theorem for Hilbert Transforms," *Proc. IEEE*, May 1963, pp. 868-869.
[16] B. C. J. Moore, "Masking in the human auditory system," in *Collected Papers on Digital Audio Bit-Rate Reduction*, N. Gilchrist and C. Grewin, Eds., 1996, pp. 9-19.