Industrial Mathematics Institute

2004:12

Analysis of the intrinsic mode functions

R.C. Sharpley and V. Vatchev

IMI

Preprint Series

Department of Mathematics
University of South Carolina

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2004		**00-00-2004 to 00-00-2004**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Analysis of the Intrinsic Mode Functions	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of South Carolina,Department of Mathematics,Columbia,SC,29208**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The Empirical Mode Decomposition is a process for signals which produces Intrinsic Mode Functions from which instantaneous frequencies may be extracted by simple application of the Hilbert transform. The beauty of this method to generate redundant representations is in its simplicity and its effectiveness. Our study has two objectives: first, to provide an alternate characterization of the Intrinsic Mode components into which the signal is decomposed and second, to better understand the resulting polar representations, specifically the ones which are produced by the Hilbert transform of these intrinsic modes.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a REPORT **unclassified**	b ABSTRACT **unclassified**	c THIS PAGE **unclassified**	**Same as Report (SAR)**	**32**	

Analysis of the Intrinsic Mode Functions [1]

by

Robert C. Sharpley and Vesselin Vatchev

Abstract

The Empirical Mode Decomposition is a process for signals which produces Intrinsic Mode Functions from which instantaneous frequencies may be extracted by simple application of the Hilbert transform. The beauty of this method to generate redundant representations is in its simplicity and its effectiveness.

Our study has two objectives: first, to provide an alternate characterization of the Intrinsic Mode components into which the signal is decomposed and second, to better understand the resulting polar representations, specifically the ones which are produced by the Hilbert transform of these intrinsic modes.

1 Introduction

The Empirical Mode Decomposition (EMD) is an iterative process which decomposes real signals $f(t)$ into simpler signals (modes)

$$f(t) = \sum_{j=1}^{M} \psi_j(t). \tag{1.1}$$

Each "monocomponent" signal ψ_j (see [C95]) should be representable in the form

$$\psi(t) = r(t) \sin \theta(t) \tag{1.2}$$

where the amplitude and phase are both physically and mathematically meaningful. Once a suitable polar parameterization is determined, it is possible to analyze the function f by processing these individual components. Important information for analysis, such as the instantaneous frequency and instantaneous bandwidth of the components, are derived from the particular representation used in (1.2). The most common procedure to determine a polar representation is the Hilbert transform and this procedure will be an important part of our discussion.

[1] This work was supported in part by ONR Grants N00014-00-1-0470, N00014-03-1-0051, and NSF Grant DMS-0079549.

Mathematics Subject Classification 2000. Primary 94A12, 92C55, 41A58; Secondary 34B24, 44A15.

Keywords and Phrases. Intrinsic mode function, empirical mode decomposition, signal processing, instantaneous frequency, redundant representations, multiresolution analysis.

In this paper we study the monocomponent signals ψ, called Intrinsic Mode Functions or IMF's, which are produced by the Empirical Mode Decomposition, and their possible representations (1.2) as real parts of analytic signals. Our study of IMF's utilizes mathematical analysis to characterize requirements in terms of solutions to self-adjoint second order ordinary differential equations. In principle, this seems quite natural since signal analysis is often used to study complex vibrational problems and the processes which generate and superimpose the signal components. Once this characterization is established, we then focus on the polar representations of IMF's which are typically built using the Hilbert transform, or more commonly referred to as the analytic method of signal processing.

The difficulty in constructing representations (1.1) is that the expansion must be selected as a linear superposition from a redundant class of signals. Indeed, there are infinitely many nontrivial ways to construct representations of the type (1.1) even in the case that the initial signal f is itself a single "monocomponent". Hence ambiguity of representation, i.e. redundancy, enters on at least two levels: the first in determining a suitable decomposition as a superposition of signals, and the second, after settling on a fixed decomposition, in appropriately determining the amplitude and phase of each component.

At the second stage, it is common practice to represent the component signal in complex form

$$\Psi(t) = r(t) \exp i\theta(t) \tag{1.3}$$

and, after a phase shift by $\pi/2$, to consider $\psi(t)$ as the real part of the complex signal Ψ, as in (1.2). Obviously, the choice of amplitude-phase representation (r, θ) in (1.3) is essentially equivalent to the choice of an imaginary part ϕ:

$$r(t) = \sqrt{\psi^2 + \phi^2} \qquad \theta(t) = \arctan \frac{\phi}{\psi}, \tag{1.4}$$

once some care is taken to handle the branch cut. An analyzing procedure should produce for each signal ψ, a properly chosen companion ϕ for the imaginary part, which is unambiguously defined and properly encodes information about the component signal, in this case the IMF. From the class of all redundant representations of a signal, once a fixed, acceptable representation, with amplitude $r(t)$ and phase $\theta(t)$, is determined, the *instantaneous frequency* of $\psi(t)$ with respect to this representation is the derivative of the phase, i.e. $\theta'(t)$. In this case, a reasonable definition for the *instantaneous bandwidth* is $\frac{r'(t)}{r(t)}$ (see [C95] for additional motivation). The collection of instantaneous phases present at a given instant (i.e. $t = t_0$) for a signal $f(t)$ is heavily dependent upon both the decomposition (1.1) and the selection of representations (1.2) for each monocomponent. The full EMD procedure is obviously a highly nonlinear process, which effectively builds and analyzes inherent components which are adapted to the scale and location of the signal's features.

Historically, there have been two methods used to define the imaginary part of suitable signals, the *analytic* and *quadrature* methods. The analytic signal method results in a complex signal that has its spectrum identical (modulo a constant factor of 2) to that of the real signal for positive frequencies and zero for the negative frequencies. This can be achieved in a unique manner by setting the imaginary part to be the Hilbert transform of the real signal $f(t)$. The *Empirical Mode Decomposition* (EMD) of Huang *et al* [H99] is a highly successful method used to generate a decomposition of the form (1.1) where

the individual components contain significant information. These components were named *Intrinsic Mode Functions* (IMF's) in [H99] since the analytical signal method applied to each such component normally provides desirable information inherent in that mode. Analytic signals and the Hilbert transform are powerful tools and are well understood in Fourier analysis and signal processing, but in certain common circumstances the analytic signal method leads to undesirable and "paradoxical" results in applications which are detailed in [C95]. Results of this paper provide further light on the consequences of using analytic signals, as currently applied, for estimating phase and amplitude of a signal. More results along these lines appear in [V].

Section 2 contains a brief description of the EMD method and motivates the concept of Intrinsic Mode Functions (IMFs) which are the main focus of this paper. Preliminary results on self-adjoint equations are reviewed for background for the results that follow in Section 3. Section 3 contains one of the main results of the paper, namely, the characterization of IMFs as solutions to certain self-adjoint ordinary differential equations. The proof involves a construction of envelopes which do not rely on the Hilbert transform. These envelopes are used directly to compute the coefficients of the differential equations. The differential equations are natural models for linear vibrational problems and should provide further insight into both the EMD procedure and its IMF components. Indeed, signals can be decomposed using the EMD procedure and the resulting IMF's used to identify systems of differential equations naturally associated with the components (see [V] for details).

The purpose of Section 4 of the paper is to further explore the effectiveness of the Hilbert analysis which are applied to IMF's and to better understand some of the anomalies that are observed in practice. Examples are constructed, both analytically and numerically, in order to illustrate that the assumption that an IMF should be the real part of an analytic signal leads to undesirable results. Well-behaved functions are presented, for which the instantaneous frequency computed using the Hilbert transform changes sign, i.e. the phase is non-monotone and physically unrealistic. In order to clarify the notions and procedures, we briefly describe both analytical and computational notions of the Hilbert transform.

Finally we take this opportunity to thank Professor Ronald DeVore (South Carolina) and Professor John Pierce (USNA) for drawing our attention to the Empirical Mode Decomposition and providing the primary references [H98, H99] from which to proceed.

2 The Empirical Mode Decomposition Method

The use of the Hilbert Transform for decomposing a function into meaningful amplitude and phase requires some additional conditions on the function. Unfortunately, no clear description of definition of a signal has been given to judge precisely whether or not a function is a "monocomponent". To compensate for this lack of precision, the concept of "narrow band" has been adopted as restriction on the data in order that the instantaneous frequency be well defined and make physical sense. The instantaneous frequency can be considered as an average of all the frequencies that exist at a given moment, while the instantaneous bandwidth can be considered as the deviation from that average. The most common example is considered to be a signal with constant amplitude, that is $r(t)$ in (1.4) is a constant. Since the phase is modulated, these are usually referred to as frequency modulated, or FM signals.

If no additional conditions are imposed on a given signal, the previously defined notions could still produce "paradoxes". To minimize these physically incompatible artifacts, Huang, *et al* [H99] have developed a method, which they termed the "Hilbert view", in order to study nonstationary and nonlinear data in nonlinear mechanics. The main tools which are used are the Empirical Mode Decomposition Method (EMD) to decompose signals into Intrinsic Mode Functions (IMF's), which are then processed by the Hilbert Transform to produce corresponding analytic signals for each of the inherent modes .

In general, EMD may be applied either to sampled data or to functions of real variable $f(t)$, by first identifying the appropriate time scales that will reveal the physical characteristics of the studied system, decompose the function into modes ψ intrinsic to the function at the determined scales, and then apply the Hilbert transform to each of the intrinsic components.

In the words of Huang and collaborators, the EMD method was motivated "from the simple assumption that any data consists of different simple intrinsic mode oscillations". Three methods of estimating the time scales of f at which these oscillations occur have been proposed:

- the time between successive zero-crossings ;
- the time between successive extrema;
- the time between successive curvature extrema.

The use of a particular method depends on the application. Following the development in [H99], we define a particular class of signals with special properties that make them well suited for analysis.

Definition 2.1 *A function $\psi(t)$ is defined to be an* **Intrinsic Mode Function***, or more briefly an IMF, of a real variable t, if it satisfies two characteristic properties:*

(a.) ψ has exactly one zero between any two consecutive local extrema.
(b.) ψ has zero "local mean".

A function which is required to only satisfy condition (a) will be called a **weak-IMF***.*

In general, the term *local mean* in condition (b) may be purposefully ambiguous, but in the EMD procedure it is typically the pointwise average of the"upper envelope" (determined by the local maxima) and the"lower envelope" (determined by the local minima) of ψ.

The EMD procedure of [H98] decomposes a function (assumed to be known for all values of time under consideration) into a *function-tailored*, fine-to-coarse multiresolution of IMF's. This procedure is extremely attractive, both for its effectiveness in a wide range of applications and for its simplicity of implementation. In the latter respect, one first determines all local extrema (strict changes in monotonicity) and, for an upper envelope, fits a cubic spline through the local maxima. Similarly, a cubic spline is fit through the local minima for a lower envelope and the local mean is the average of these two envelopes. (It is well understood that these are *envelopes* in a loose sense). If the local mean is not zero, then the current local mean is subtracted leaving a current candidate for an IMF. This process is continued (accumulating the local means) until the local mean vanishes or is "sufficiently

small‘”. This process (inner iteration) results in the IMF for the current scale. The accumulated local means from this inner iteration is the version of the function scaled-up to the next coarsest scale. The process is repeated (outer iteration) until the residual is either “sufficiently small‘” or monotone.

In view of the possible deficiency of the upper and lower envelopes to bound the iterates and in order to speed convergence in the inner loop, Huang *et al* suggest [H99] that the stopping criterion on the inner loop be changed from the condition that the “resulting function to be an IMF” to the single condition that “the number of extrema equals to zero-crossings” along with visual assessment of the iterates. This is the motivation for our definition of weak-IMF. Ideally in performing the EMD procedure, all stopping and convergence criteria will be met and $f(t)$ is then represented as

$$f(t) = \sum_{n=1}^{N} \psi_n + r_{N+1}$$

where $r_{N+1} = f_{N+1}$ is the residual, or carrier, signal.

A primary purpose of the decomposition [H99] is to distill from a signal individual modes whose frequency (and possibly bandwidth) can extracted and studied by the methods from the theory of analytic signals. More specifically, quoting from [H99],

> “Having obtained the IMF components, one will have no difficulty in applying the Hilbert transform to each of these components. Then the original data can be expressed as the real part($\Re$) in the following form:
>
> $$f(t) = \Re\left(\sum_{n=1}^{N} A_n(t) \exp\left(i \int \omega_n(t)\, dt\right)\right).$$
>
> The residue r_N is left on purpose, for it is either a monotonic function or a constant.”

The notation above uses $\omega_n = \frac{d\theta_n}{dt}$ to refer to the instantaneous frequency, where the phase of the n-th IMF is computed by $\theta_n := \arctan(H\psi_n/\psi_n)$.

2.1 Initial Observations

The *first step* in a multiresolution decomposition is to choose a time scale which is inherent in the function $f(t)$ and has relevant physical meaning. The scales proposed in [H99] are sets of significant points for the given function $f(t)$. Other possibilities that could be used are the set of inflection points (also mentioned by the authors), the set of zero crossings of the function $f(t) - \cos kt$, k-integer, or some other characteristic points.

The *second step* is to extract some special (with respect to the already chosen time scale) functions, which in the original EMD method are called Intrinsic Mode Functions(IMF's). The definition of an IMF, although somewhat vague, has two parts (a) the number of the extrema equals the number of the zeros and (b) the upper and lower envelopes should have the same absolute value. As it is pointed out in [H99] if we drop (b) we will have a

reasonable (from practical point of view) definition but, in the next stage, this will introduce unrecoverable mathematical ambiguity in determining the modulus and phase.

Therefore any modification of the definition of IMF must include condition (a). The practical implementation of the EMD uses cubic splines as upper ($U(t)$) and lower ($L(t)$) envelopes. The nodes of these two splines interlace and do not have points in common. The absolute value of two cubic splines can be equal if and only if they are the same quadratic polynomial on the whole data span, i.e. if the modulus of the IMF is of the form at^2+bt+c. To overcome this restriction, we can either modify the construction of the envelopes or, instead of requiring $U(t) = -L(t)$ for all t, we can require $|U(t)+L(t)| \leq \epsilon$, for some prescribed $\epsilon > 0$.

Recall that we say a continuous function is a **weak-IMF** if it is only required to satisfy condition (a) in Definition 2.1 of an IMF. One of the main purposes of this paper is to provide a complete characterization of the weak-IMF's in terms of solutions to self-adjoint ordinary differential equations. In a sense this is natural, since one of the uses of the EMD procedure is to study solutions to differential equations and vibration analysis was a major motivation in the development of the Sturm-Liouville theory. In the next section, we list some relevant properties of the solutions of a self-adjoint ODE's which will be useful for our analysis.

2.2 Self-adjoint ODE and Sturm-Liouville systems.

An ODE is called self-adjoint if can be written in the form

$$\frac{d}{dt}\left(P(t)\,\frac{df}{dt}\right) + Q(t)f = 0, \tag{2.1}$$

for $t \in (a,b)$ (a and b finite or infinite), where $P > 0$ and Q is continuous. More generally we can consider a Sturm-Liouville equation (λ real):

$$\frac{d}{dt}\left(p(t)\frac{df}{dt}\right) + (\lambda\rho(t) - q(t))f = 0. \tag{2.2}$$

These equations arose from vibration problems associated with model mechanical systems and the corresponding wave motion was resolved into simple harmonic waves (see [BR]).

Properties of the solutions of self-adjoint and Sturm-Liouville equations

I. Interlacing zeros and extrema If $Q > 0$, then any solution of (2.1) has exactly one maximum or minimum between successive zeros.

II. The Prüfer substitution A powerful method for solving the ODE (2.1) utilizes a transformation of the solution into amplitude and phase. If the substitution $Pf' := r(t)\cos\theta(t)$ and $f(t) := r(t)\sin\theta(t)$ is made, then the equation (2.1) is equivalent to the following nonlinear first order system of ODE's

$$\frac{d\theta}{dt} = Q(t)\sin^2\theta + \frac{1}{P(t)}\cos^2\theta \tag{2.3}$$

$$\frac{dr}{dt} = \frac{1}{2}\left(\frac{1}{P(t)} - Q(t)\right) r\sin 2\theta. \tag{2.4}$$

Notice that if $Q(t)$ is positive, then the first equation shows that the instantaneous frequency of the IMF's is always positive, and therefore the solutions have nondecreasing phase. The second equation relates the instantaneous bandwidth $\frac{r'}{r}$ with $P(t)$, $Q(t)$ and $\theta(t)$. The partial decoupling in this form of the equations is useful in studying the behavior of the phase and amplitude.

III. The Liouville substitution An ODE of the form (2.2) can be transformed to an ODE of the type

$$f'' + (\lambda - q(t))f = 0.$$

Moreover, if $f_n(t)$ is a sequence of normalized eigenfunctions, then

$$f_n(t) = \sqrt{\frac{2}{b-a}} \cos \frac{n\pi(t-a)}{b-a} + \frac{O(1)}{n}.$$

Additional properties of these solutions (e.g. see [BR]) suggest that the description of IMF's as solutions to self-adjoint ODE's will lead to further insight.

3 IMF's and Solutions of Self-adjoint ODE

In this section we characterize weak-IMF's which arise in the Empirical Mode Decomposition algorithm as solutions of self-adjoint ODE's. The main result may be stated as follows.

Theorem 3.1 *Let f be a real-valued function in $C^2[a,b]$, the set of twice continuously differentiable functions on the interval $[a,b]$. If both f and its derivative f' have only simple zeros, then the following three conditions are equivalent:*
(i) *The number of the zeros and the number of the extrema of f on $[a,b]$ differ by at most one;*
(ii) *There exist positive continuously differentiable functions P and Q such that f is a solution of the self-adjoint ODE*

$$(P(t)f'(t))' + Q(t)f(t) = 0; \tag{3.1}$$

(iii) *There exists an associated $C^2[a,b]$ function h such that the coupled system*

$$f(t) = -\frac{1}{Q(t)}h'(t), \quad h(t) = P(t)f'(t), \tag{3.2}$$

holds for some positive continuously differentiable functions P and Q.

Proof: We first prove that condition (i) is equivalent to (ii). That condition (ii) implies (i) follows immediately since Q is a positive function and Property I of the previous section holds for solutions of self-adjoint ODE's (see [BR]).

The proof in the opposite direction ((i) implies (ii)) requires a preliminary result (see Lemma 3.1 that follows) on interpolating piecewise polynomials to be used for envelopes. Let us assume then that there is exactly one zero between any two extrema of f. For

simplicity we assume that the number of zeros and extrema of f on $[a,b]$ are both equal to M. Consider the collection of ordered pairs

$$\{(t_j, |f(t_j)|)\}_{j=1}^{M} \cup \{(z_j, a_j)\}_{j=1}^{M}, \tag{3.3}$$

which will serve as our knot sequence. The points $\{t_j, z_j\}$ satisfy the required interlacing condition ($t_1 < z_1 < t_2 < z_2 < \ldots$), where t_j are the extremal points for f and z_j are its zeros. The data $\mathbf{a} = \{a_j\}$ are any positive numbers which satisfy

$$\max\{|f(t_j)|, |f(t_{j+1})|\} + \eta \leq a_j \tag{3.4}$$

for all $j = 1, \ldots, M$, where $\eta > 0$ is fixed. The following lemma provides a continuous piecewise polynomial envelope for f by Hermite interpolation.

Lemma 3.1 *Let f satisfy the conditions of Theorem 3.1 and the $\{a_j\}$ satisfy the condition (3.4), then there is a continuous, piecewise quintic polynomial R interpolating this data with the following properties, for all j:*

(a) The extrema of R occur precisely at the points t_j, z_j;

(b) $|f| \leq R$ with equality occuring exactly at the points t_j.

(c) R is strictly increasing on (t_j, z_j) and strictly decreasing on (z_j, t_{j+1}).

(d) $R''(t_j) \neq (-1)^{j+1} f''(t_j)$.

Proof: Indeed, let the collection $\{a_j\}$ satisfy (3.4), where $\eta > 0$ is fixed. Interpolate the data specified by (3.3) by a piecewise quintic polynomial R, requiring in addition that $R'(t_j) = R'(z_j) = 0$. On each subinterval determined by the points $\{t_j, z_j\}$, this imposes four conditions on the six coefficients of the local quintic, leaving 2 degrees of freedom for each of the polynomial 'pieces'. Representing such a polynomial in its Taylor expansion about the left hand endpoint of its interval, it is easy to verify that we can force that condition (c) hold at each of the knots, and that we can require $R''(t_j) > 0$. In particular, R has its minima at the maxima of $|f|$ (i.e. the t_j) and its maxima at the zeros of f (the z_j). Therefore, $R''(t_j) > 0 \geq (-1)^{j+1} f''(t_j)$, which verifies condition (d) . □

Remark 3.1 *In general, any piece-wise function R constructed from functions $\varphi_j(t)$ that satisfy the conditions*

$$\begin{aligned} &\varphi(y_1) = v_1,\ \varphi(y_2) = v_2 \\ &\varphi'(y_1) = \varphi'(y_2) = 0,\ \ |\varphi'(t)| > 0 \ \text{ for } \ t \in (y_1, y_2) \end{aligned}$$

will suffice in our construction. In particular, the Meyer's scaling function can be used to produce an envelope R which satisfies properties (a) and (b) of Lemma 3.1 and can be used as a basis for a quadrature calculation of instantaneous phase (see [V]). This idea is implicit in the development that follows.

Having constructed an envelope R for f, we define the phase-related function S by

$$S(t) := \frac{f(t)}{R(t)}. \tag{3.5}$$

By Lemma 3.1, clearly $|S(t)| \leq 1$ for $t \in [a, b]$ and $|S(t)| = 1$ if and only if $t = t_j$ for some $j = 1, 2, \ldots, M$. Since f has exactly one zero between each pair of consecutive interior extrema, then f, and hence S, has alternating signs at the t_j. Without loss of generality, we may assume $\theta(t_1) = \frac{\pi}{2}$, i.e. t_1 is an interior local maximum, otherwise we could consider the function $-f$ instead of f. Endpoint extrema are easily handled separately. As we observed during the proof of Lemma 3.1, the function R was constructed to be strictly increasing on (t_j, z_j) and strictly decreasing on (z_j, t_{j+1}). On intervals (t_j, t_{j+1}), when j is odd, the function f decreases, is positive on (t_j, z_j), and negative on (z_j, t_{j+1}). These properties imply that S decreases on (t_j, t_{j+1}), is positive on (t_j, z_j), and negative on (z_j, t_{j+1}). Similar reasoning shows that for j even, S increases on (t_j, t_{j+1}), is negative on (t_j, z_j), and positive on (z_j, t_{j+1}).

Therefore we can represent S as

$$S(t) =: \sin\theta(t) \tag{3.6}$$

for an implicit function $\theta(t)$ which satisfies $\theta(t_j) = \frac{2j-1}{2}\pi$ and $\theta(z_j) = j\pi$. From these facts, one easily checks that θ is a strictly increasing function. In fact, $\theta(t)$ will be continuously differentiable with strictly positive derivative on $[a, b]$. To see this, first recall that the function R has a continuous first derivative on $[a, b]$, so S is also differentiable and satisfies

$$S'(t) = \frac{f'R - fR'}{R^2}. \tag{3.7}$$

Therefore S' is continuous and by an application of the implicit function theorem applied on each of the intervals (t_j, t_{j+1}), $\theta(t)$ will be continuously differentiable with positive derivative on each of these intervals. We will apply L'Hospital's rule in order to verify the corresponding statement at the extrema t_j. Differentiate formally the relation (3.6) and square the result to obtain on each interval (t_j, t_{j+1}) the identity

$$\theta'(t)^2 = \left(\frac{S'(t)}{\cos(\theta(t)}\right)^2 = \frac{S'(t)^2}{1 - S^2(t)}.$$

So, if $T(t)$ denotes the right-hand side of the above relation, that is

$$T(t) := \frac{S'(t)^2}{1 - S^2(t)}\,, \tag{3.8}$$

then $T(t)$ is clearly continuous except at the t_j where it is undefined. We show, however, that T has removable singularities at these points. Both the numerator and denominator are C^2 functions and vanish at t_j, so an application of L'Hospital's rule shows

$$\lim_{t\to t_j} T(t) = \lim_{t\to t_j} \frac{2S'(t)S''(t)}{-2S'(t)S(t)} = -\frac{S''(t_j)}{S(t_j)}.$$

On the other hand, from (3.7), $S''(t_j) = \frac{(-1)^{j+1} f''(t_j) - R''(t_j)}{f(t_j)}$, and so property (d) of the previous lemma guarantees that this last expression is strictly positive. Hence, θ' is a continuous, strictly positive function on the interval $[a, b]$.

If we use relations (3.5) and (3.6) to write f as $f = R \sin\theta$, then a natural companion is the function h defined by

$$h(t) := -R(t)\cos\theta(t). \tag{3.9}$$

It follows from properties of R and θ that h is strictly decreasing on (z_j, z_{j+1}) when j is odd, is strictly increasing on this interval when j is even, and has its simple zeros at the points t_j. Differentiation of (3.9), provides the identity

$$h'(t) = -R'(t)\cos\theta(t) + R(t)\theta'(t)\sin\theta(t). \tag{3.10}$$

which will be used to complete the proof that condition (ii) of Theorem 3.1 is satisfied. Indeed, define the functions P, Q appearing in equation (3.1) by

$$P(t) := -\frac{h(t)}{f'(t)}, \quad Q(t) := \frac{h'(t)}{f(t)}. \tag{3.11}$$

From the properties of h and f, we see that these are well defined, strictly positive, and with continuous first derivatives on $[a, b]$, except possibly at the set of points $\{t_j\}$ and $\{z_j\}$. That these properties persist at these points as well, we can again apply L'Hospital's rule and use the identity (3.10) together with the fact that θ' is positive.

Obviously, the equations (3.11) are equivalent to

$$P(t)f'(t) = -h(t), \quad Q(t)f(t) = h'(t) \tag{3.12}$$

which in turn are equivalent to equations (3.2). This establishes condition (ii) of Theorem 3.1 and also shows that this condition is equivalent to condition (iii). □

Remark 3.2 *Observe that the function h in condition (iii) of Theorem 3.1 satisfies a related self-adjoint ODE,*

(i) $(\tilde{P}(t)h'(t))' + \tilde{Q}(t)h(t) = 0$, *where* $\tilde{P} := 1/Q$ *and* $\tilde{Q} := 1/P$ *and* P, Q *are the coefficients of Theorem 3.1 .*

Moreover, the coefficients P, Q satisfy the following conditions:

(ii) P, Q *may be represented directly in terms of the amplitude* $R(t)$ *and phase* $\theta(t)$ *by*

$$\frac{1}{P} = \theta' + \frac{R'}{R}\tan\theta, \quad Q = \theta' - \frac{R'}{R}\cot\theta. \tag{3.13}$$

(iii) P, Q *satisfy the inequality*

$$\frac{1}{P} \leq Q,$$

with equality iff $R'(t) = 0$ on $[a, b]$, or, equivalently, if f is an FM signal.

The only statements in this Remark that requires justification are equations (3.13). These follow directly by using the Prüfer substitution in equations (3.13): using (2.3) for θ' and (2.4)

for R'/R.

Theorem 3.1 provides the desired characterization of weak-IMF's, which we summarize in the following corollary.

Corollary 3.1 *A twice differentiable function ψ on $[a, b]$ is a* **weak-IMF** *if and only if it is a solution of the self-adjoint ODE of the type*

$$(P\psi')' + Q\psi = 0,$$

for positive coefficients $P(t)$, $Q(t)$.

If we adopt the definition of an IMF given in Definitiion 2.1, then we have a characterization embodied in the following statements summarizing the results and observations of this section.

Theorem 3.2 *A function ψ is an IMF if and only if it is a weak-IMF whose spline envelopes satisfy the condition that the absolute value of the lower spline envelope is equal to the upper envelope and this common envelope is a quadratic polynomial. Furthermore, the common spline evelope is constant (i.e., ψ is an FM signal) if and only if $Q(t) = 1/P(t)$ for the associated self-adjoint differential equation (3.1).*

The results of this section indicate that we can find a meaningful mathematical and physical description of any weak-IMF in terms of solutions of self-adjoint problems. On the other hand, considering these as the real parts of analytic signals, we show in the next section that there exist functions ψ that are IMF's satisfying both conditions (a) and (b) of Definition 2.1, but the phase produced by using the Hilbert transform is not monotonic, i.e. the instantaneous phase changes sign.

4 Example IMFs and the Hilbert Transform

In this section we analyze several examples that indicate the limitations of the analytic method (i.e. Hilbert transform) to produce physically realistic instantaneous frequencies in the context of IMF analysis. The examples presented show that even for some of the most reasonable definitions for IMFs the Hilbert Transform method will result in instantaneous frequencies which change signs on intervals of positive measure. By a *reasonable* IMF we mean that they satisfy all existing definitions, including the IMF of Huang *et al*, narrowband mono-components, and visual tests. Although our examples are presented in order to identify possible pitfalls in automatic use of the Hilbert transform, in the final analysis, practitioners in signal processing will make the decision on when the use of analyticity is appropriate, and to what extent non-monotone phase is necessary. We mention that the examples, in some sense, also provide a better understanding of many of the paradoxes concerning instantaneous phase and bandwidth which are detailed in Cohen [C95].

4.1 Hilbert Transforms

In order to clarify the discussion, we begin with a brief description of Hilbert transforms and analyticity. In using the terminology "Hilbert Transform method", we mean one of the following

- the *conjugate operator* (or *periodic Hilbert transform*):
  i.e., the transform which is defined for functions ψ on the circle as the imaginary parts of analytic functions whose real part coincides with ψ, see [K, Z] for details. This may be identified with modifying the phase of each Fourier frequency component by a quarter cycle delay, i.e. the *sgn* Fourier coefficient multiplier.
- the *continuous Hilbert Transform*:
  i.e., the transform for functions ψ defined on the real line which is defined as the restriction to $\Re$ of the imaginary part of analytic functions in the upper half plane whose real part on $\Re$ is ψ. This is well defined and understood, for example, on Lebesgue, Sobolev, Hardy, and Besov spaces ($1 \leq p < \infty$ and in certain cases when $p = \infty$). This transform may be realized both as a principal value singular integral operator and as a (continuous) Fourier multiplier. For details see [BS, Z].
- the *discrete Hilbert transform*:
  i.e., a transform on discrete groups which is applied to signals through a multiplier operator of its discrete Fourier transform. The operator is computed by multiplying the FFT coefficients of a signal by *sgn* and then inverting. The multiplier may possibly invoke side conditions such as those as implemented in the built-in version of 'hilbert' in Matlab [M]. We also note that the m-file 'hilbert.m' computes the discrete analytic signal itself and not just the imaginary part.

In each of these cases it is clear that the imaginary part (in the case of continuous functions) is uniquely defined up to an arbitrary numerical constant C. In Fourier and harmonic analysis the choice is usually made based on consideration of the multiplier operator as a bounded isometry on L^2. In some of our examples, we will consider functions on $\Re$ and sample them in order to apply the Discrete Hilbert Transform. For periodic functions and appropriate classes of functions defined on $\Re$, careful selection of the sampling resolution (e.g. Shannon Sampling Theorem [P] in the case of analyzing functions of exponential type) will guarantee that sampling the continuous Hilbert transform of the functions will be equivalent (at least to machine precision) to application of the discrete Hilbert transform to the sampled function. In other words, these numerical operations, when carefully applied, will "numerically comute". It will be clear if there is a distinction between these transforms and, from the context, which one is intended.

One possible remedy in order to try to avoid nonphysical artifacts of the "analytic" method of computing the instantaneous frequency is to require additional constraints in the definition of an IMF. One such condition which immediately comes to mind would be to also require an IMF to have at most one inflection point between its extrema. We show in Example 4.2, however, that even stronger conditions are still not sufficient to prevent sign changes of the instantaneous frequency when Hilbert transforms are used to construct the phase and amplitude for a signal, that is, if one considers an IMF as the real part of an analytic signal. In Propositions 4.1-4.3 we consider the analytical properties of these exam-

ples and show that they are members of large classes of signals that behave similarly when processed by the Hilbert transform, or by the computational Hilbert transform, not matter how finely resolved. Finally, we conclude this section by describing a general procedure that adds a "smooth perturbation" to well behaved signals and leads to undesirable behavior in estimating the instantaneous phase. This indicates the need for the possible consideration of careful denoising of acquired signals before processing IMF's by the Hilbert method.

Before proceeding it is useful to briefly discuss computational aspects of the Hilbert transform and therefore of the corresponding analytic signal. There are several versions of the discrete Hilbert transform, all using the Discrete Fourier Transform (DFT). In the study of monocomponent signals which are Fourier-based and use least squares norms, the choice of the free parameter C is normally chosen so that $\|\psi\|_{\ell^2} = \|H\psi\|_{\ell^2}$, which mimics the corresponding property for transforms on the line and circle. As implemented by Matlab, however, it seems that for many signal processing operations it is preferable to choose the free imaginary constant so that the constant (DC) term of the signal is split between the constant and middle (Nyquist) terms of the DFT of the Hilbert transform. This appears natural since the Nyquist coefficient is aliased to the constant term, see Marple [M] for details. An additional side benefit of this choice of C is that it ensures that the discrete Hilbert transform will be orthogonal to the original signal, which emulates the corresponding property for the Hilbert transform for the line and circle. We note that the discretization process does not permit one to maintain all properties of continuous versions of the transform and some choice on which properties are most important must be made based on the application area.

One serious numerical artifact of the computational Hilbert transform, which typically arises when it is applied to non-continuous periodic functions, is a Gibbs effect. Some care must be taken to insure continuity of the (implicitly assumed) periodic signal, otherwise severe oscillations will occur which often mask the true behavior of the instantaneous frequency. In the examples considered in this section the functions are continuous, although in some cases (see Example 4.2) the higher derivatives are not. In this case, however, the oscillations due to this lack of smoothness are minor, of lower order and do not measurably affect the computations. Typically, we apply the computational Hilbert transform after the supports of our functions are rescaled and translated to the interval $[-\pi, \pi)$.

Since it may rightly be argued that other choices of the free parameter C in the discrete Hilbert transform may possibly alleviate the problem of nonmonotone phase, we focus for the most part on examples for which any choice of the imaginary constant in the analytic signal (and consequently in the definition of the discrete Hilbert transform) will result in undesirable behavior of the instantaneous frequencies obtained by the Hilbert method. Another concern in computational phase estimation is how one numerically 'unwraps' Cartesian expressions to extract phases for polar representations. We offer a technique to avoid ambiguous unwrapping of inverse trigonometric functions by instead computing the 'analytic' instantaneous frequency through the formula

$$\theta'_C(t) := \frac{\psi(t)H\psi'(t) - (H\psi(t) + C)\psi'(t)}{(H\psi(t) + C)^2 + \psi(t)^2} \tag{4.1}$$

where θ_C is the phase corresponding to a given choice of the constant C. We use this identity throughout to compute instantaneous frequencies for explicitly defined functions ψ which are either periodic or defined on the line. Discrete versions using first order differences are also

suitable for computing instantaneous phase for discretely sampled signals. The identity (4.1) follows by implicitly differentiating the expression $\tan\theta_C = (H\psi + C)/\psi$ and using the fact that the Hilbert transform is translation-invariant.

We end this subsection with an general observation concerning the application of the Hilbert transform to IMF's, which follows from Theorem 3.1.

Corollary 4.1 *Suppose that ψ is a periodic, weak-IMF and Ψ is the corresponding analytic function with imaginary part the conjugate operator $H\psi$. If (r, θ) are the corresponding analytic amplitude and phase for the pair $(\psi, H\psi)$, then the coefficients (P, Q) of an associated differential equation (2.1) determined by a Prüfer relationship (3.13) must satisfy*

$$Q = \frac{H\psi'}{\psi}, \quad P = -\frac{H\psi}{\psi'}, \tag{4.2}$$

whenever these two expressions make sense. In particular, a necessary and sufficient condition that the coefficients (P, Q) of the ODE be positive (i.e. a physcially reasonable vibrational system), is that $H\psi$ should be positive exactly where ψ increases, and ψ should be positive exactly where $H\psi$ is increasing.

Proof: This follows immediately from Theorem 3.1 and the Prüfer representation of the coefficients which is given in equation (3.13). □

4.2 Example IMFs

The first examples of IMF's we wish to consider are a family of 2π-periodic functions which have the property that the conjugate operator and the discrete Hilbert transform (applied to a sufficiently refined sampling) differ only by the addition of an imaginary constant.

Example 4.1 *Let ϵ be a real parameter. We consider the family of continuous 2π periodic functions*

$$\psi_\epsilon(t) := e^{\epsilon\cos(t)}\sin(\epsilon\sin(t)). \tag{4.3}$$

Observe that the Hilbert transform of ψ_ϵ is $H\psi_\epsilon(t) = -e^{\epsilon\cos(t)}\cos(\epsilon\sin(t)) + C$, where the constant C is a free parameter that one may choose. In fact, the analytic signal Ψ with real part ψ_ϵ is unique up to a constant and may be written as

$$\Psi_\epsilon(t) = -ie^{\epsilon\, e^{it}} + iC.$$

For particular values of ϵ the function can be used as a model of signals with interesting behavior. For example, ψ_ϵ for $\epsilon \leq 2.9716$ is an FM function and on any finite interval the number of the zeros differs from the number of extrema by a count of at most one.

As one particular example of the Hilbert method for computing instantaneous phase for IMF's, we fix in (4.3) the special choice of $\epsilon_0 = 2.97$ and set

$$\psi(t) := \psi_{\epsilon_0}(t). \tag{4.4}$$

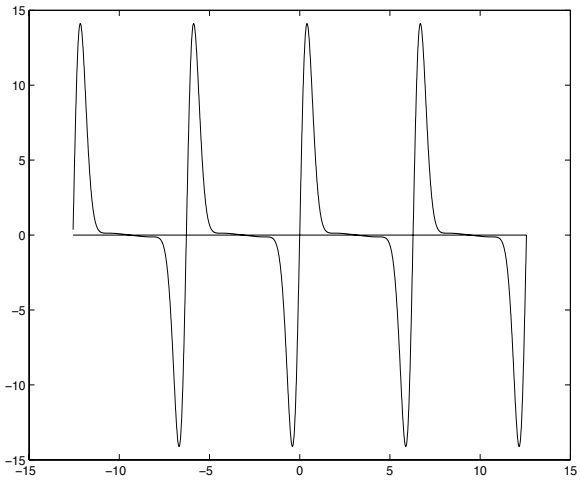


Figure 1: ψ_ϵ, an IMF with poor Hilbert transform.

The graph of $\psi(t)$ is shown in Fig 3.1. In Proposition 4.1 below, we show that $\psi(t)$ is an IMF according to the definition in [H98], but for any choice of the constant C in the Hilbert transform, the instantaneous frequency obtained from the corresponding analytic signal Ψ_{ϵ_0} changes sign.

We first verify the corresponding fact in the case of discrete signals. We sample $\psi(t)$ uniformly with increment $\Delta = \pi/128$ (vector length = 1024) on the interval $[-4\pi, 4\pi - \Delta]$. The graph of the Hilbert transform and corresponding instantaneous frequency of ψ obtained by using Matlab's built-in "hilbert.m" function are shown in Figure 2, parts (a) and (b), respectively. We mention that for this data the choice of constant chosen by Matlab to meet its criteria is $C = 1$. Although other choices for C may decrease the intervals of non-monotonicity of the phase, the artifact will persist for all choices.

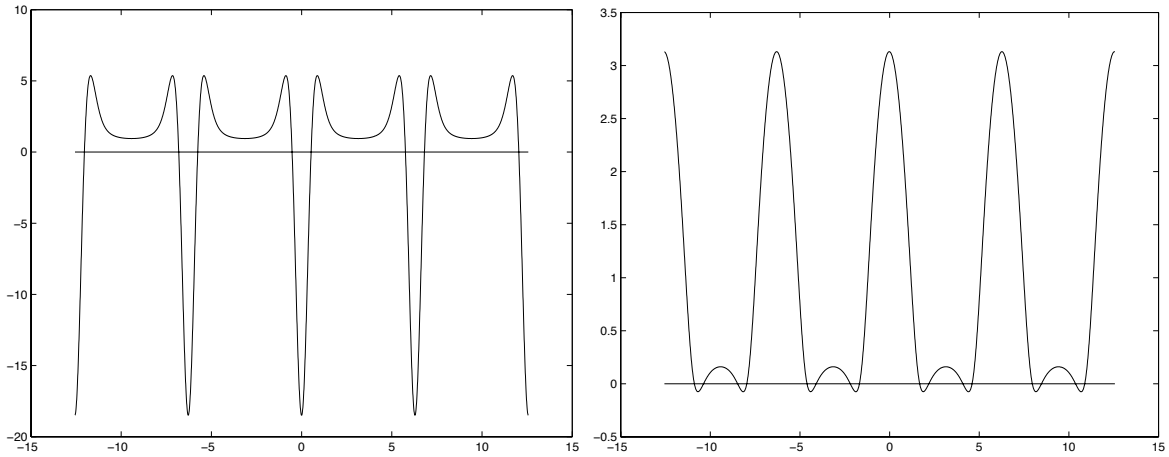
The next proposition shows that the computational observation using the discrete Hilbert transform is a consequence of the continuous transform and cannot be corrected by other choices of the imaginary constant or by a finer sampling rate.

Proposition 4.1 *The function ψ defined by (4.3) is an IMF in the sense of [H98], but its instantaneous frequency computed by the Hilbert transform (with any choice of imaginary constant C) changes its sign on any interval of length at least π.*

Proof. We first show that ψ is a weak-IMF. Clearly ψ is 2π-periodic and an odd function and so we only need to consider it on the interval $[0, \pi)$. The first derivative of ψ is

$$\psi'(t) = \epsilon_0 e^{\epsilon_0 \cos(t)} \cos(t + \epsilon_0 \sin(t)) \tag{4.5}$$

and is zero iff $\nu(t) := t + \epsilon_0 \sin(t) = \frac{2k+1}{2}\pi$ for some integer k. Since $\nu'(t) = 1 + \epsilon_0 \cos(t)$ has exactly one zero z_0 in $[0, \pi)$ (*cos* is monotone in $[0, \pi)$), the function $\nu(t)$ is increasing on $[0, z_0)$, decreasing on (z_0, π), with end values $\nu(0) = 0$ and $\nu(\pi) = \pi$. To show that ψ has



a. **b.**

Figure 2: the Hilbert method for ψ_ϵ: **a.** Imaginary part of discrete analytic signal; **b.** Instantaneous frequency.

only one extremum on $[0, \pi)$, it suffices to show that

$$\nu(z_0) < \frac{3}{2}\pi \tag{4.6}$$

since then the only extremum of ψ on $[0, \pi)$ will be the point e_M where $\nu(e_M) = \frac{\pi}{2}$. At the point z_0 we have $\cos(z_0) = -1/\epsilon_0$, which implies both $\pi/2 < z_0 < \pi$ and

$$\sin(z_0) = \sqrt{1 - (1/\epsilon_0)^2}.$$

Hence from the definition of ν it follows that

$$\nu(z_0) = z_0 + \sqrt{\epsilon_0^2 - 1}.$$

This implies that the condition (4.6) is equivalent to $z_0 < \frac{3}{2}\pi - \sqrt{\epsilon_0^2 - 1}$. But *cos* is negative and decreasing on $[\pi/2, \pi]$, so we see that the desired relationship (4.6) just means that $\cos(z_0) > \cos(\frac{3}{2}\pi - \sqrt{\epsilon_0^2 - 1})$ should hold. The numerical value of the expression on the right is smaller than -0.3382, while $\cos(z_0) = -1/\epsilon_0 > -0.3368$, hence the condition (4.6) holds and ψ has exactly one local extremum in $[0, \pi)$. Finally, since $\epsilon_0 < \pi$, the only zeros of ψ are clearly at the endpoints 0 and π, which verifies that ψ is an weak-IMF.

To see that ψ is in fact an IMF, we need to verify the condition on the upper and lower envelopes. Recall that it is 2π periodic and odd, therefore it has exactly one minimum in the interval $[-\pi, 0]$. The cubic spline fit of the maxima (upper envelope) will be the constant function identically equal to 1. Similarly the cubic spline interpolant of the minima (lower envelope) will have constant value -1. This persists even for sufficiently large intervals if one wishes to take finitely supported functions. Hence the function ψ satisfies the envelope condition for an IMF from [H98]. We note that the general proof to show that for each $0 <$

$\epsilon < \tilde{\epsilon}$, ψ_ϵ is an IMF follows in a similar manner, where $\tilde{\epsilon}$ is the solution to the transcendental equation $1/\tilde{\epsilon} = \sin(\sqrt{\tilde{\epsilon}^2 - 1})$ which arises in the limiting cases above. We observe that $\tilde{\epsilon} \approx 2.9716$.

Next we prove that for any selection of constant C, the corresponding instantaneous frequency θ'_C for ψ which is derived from an analytic method through (4.1) will have nontrivial sign changes. The denominator in formula (4.1) is always positive so it will suffice to prove that the numerator of θ'_C changes sign for any choice of C. Using (4.3), (4.5) and the formula

$$H\psi'(t) = \epsilon_0 \; e^{\epsilon_0 \; \cos(t)} \sin(t + \epsilon_0 \; \sin(t))$$

we can simplify the numerator of θ'_C to the expression

$$\epsilon_0 \; e^{\epsilon_0 \; \cos(t)} \left(e^{\epsilon_0 \; \cos(t)} \cos(t) - C \cos(t + \epsilon_0 \; \sin(t)) \right),$$

and so the sign of the term inside the parentheses

$$N_t(C) := e^{\epsilon_0 \cos(t)} \cos(t) - C \cos(t + \epsilon_0 \sin(t))$$

determines the sign of $\theta'_C(t)$ at each point $t \in [0, \pi)$. First observe that $N_t(C)$ is a linear function of C for fixed t. For each value of C there is a point in $[0, \pi)$ at which θ'_C is negative, in fact $N_{1.9}(C) < -0.04$ for $C < 40$ while $N_{0.1}(C) < -8$ for $C > 30$. Similarly for any value of C there is a point at which θ'_C is positive since $N_{0.1}(C) > 4$ for $C < 13$ and $N_1(C) > 2$ for $C > 0$. By continuity we see that for each value of the constant C the instantaneous frequency θ'_C obtained via the Hilbert transform is positive and negative on sets of positive measure. □

Finally, we mention that the L_2 bandwidth of the analytic signal Ψ corresponding to a signal ψ also depends on the choice of the imaginary constant C. If Ψ is written in polar coordinates as $\Psi(t) = r(t)e^{i\theta(t)}$, the average frequency $\langle\omega\rangle$ and the bandwidth ν^2 have been defined in [C95] as the quantities

$$\langle\omega\rangle \;=\; \int \omega \frac{|S(\omega)|^2}{\|S\|_2^2} \, d\omega = \int \theta'(t) \frac{r^2(t)}{\|r\|_2^2} \, dt, \tag{4.7}$$

$$\begin{aligned} \nu^2 \;&:=\; \frac{1}{\langle\omega\rangle^2} \int (\omega - \langle\omega\rangle)^2 \frac{|S(\omega)|^2}{\|S\|_2^2} \, d\omega \\ &=\; \frac{1}{\langle\omega\rangle^2} \int \left(\left(\frac{r'(t)}{r(t)} \right)^2 + (\theta'(t) - \langle\omega\rangle)^2 \right) \frac{r^2(t)}{\|r\|_2^2} \, dt \\ &=\; \frac{1}{\langle\omega\rangle^2} \int \left(\left(\frac{r'(t)}{r(t)} \right)^2 + (\theta'(t))^2 \right) \frac{r(t)^2}{\|r\|_2^2} \, dt - 1, \end{aligned} \tag{4.8}$$

where $S(\omega)$ is the spectrum (Fourier transform) of $\psi(t)$. The second equation in the displayed sequence (4.8) follows immediately from Plancherel's theorem along with standard properties of the Fourier transform. If one chooses the constant C in the Hilbert transform so that $\|\psi\|_2 = \|H\psi\|_2$, then the computed bandwidth is $\nu^2 = 0.1933$ with mean frequency

$< \omega >= 2.7301$. The discrete Hilbert transform computed by matlab for the sampled ψ has the same L_2 bandwidth and mean frequency.

Summarizing, we observe that the example ψ given in (4.4) is a function which is

(i) an IMF in the sense of [H98];
(ii) a monocomponent in the sense of [C95], i.e. its L_2 bandwidth is small;
(iii) "visually nice",

but the analytic method fails to produce a monotone phase function.

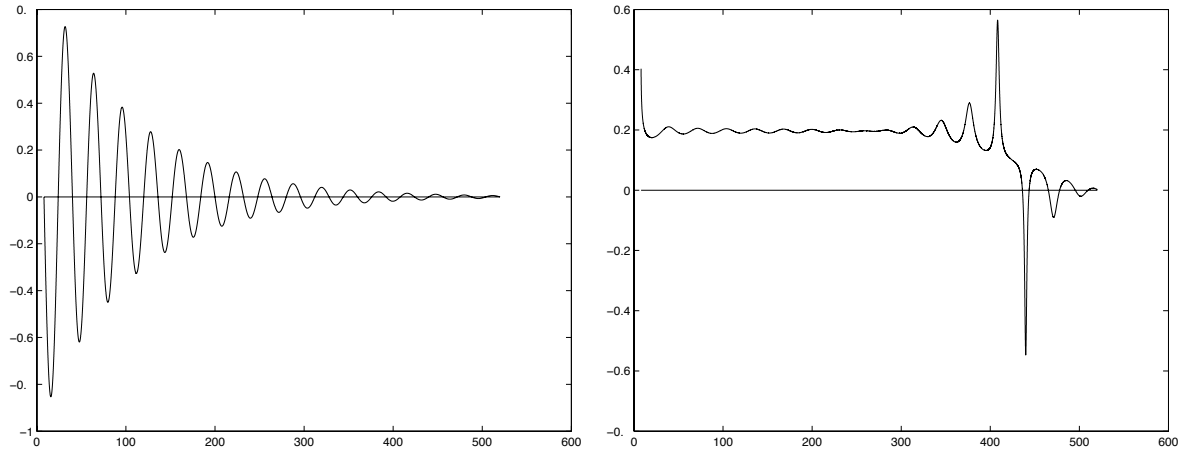
Remark 4.1 *The example considered in Proposition 4.1 also shows that adding the requirement that the Hilbert transform (with a choice of the additive constant $C = 3$) of an IMF must also be a weak-IMF, will not be sufficient to guarantee monotone phase.*

A possible natural refinement of the definition of an IMF that would exclude these functions from the class of IMFs would be to require in addition that the first derivative of an IMF also be an IMF, or at least that the number of the inflection points equals the number of extrema to within a count of one (i.e. a weak-IMF). The next example of a damped sinusoidal signal (*i.e. an amplitude modulated signal*) shows that restrictions along these lines will not be able to avoid the same problem with instantaneous frequencies. We note that this particular signal ψ is considered in [H98], but for the range of t from 1-512 sec. Since the function is not continuous periodic over this range, the expected Gibb's effect at the points $t = 1$ and 512 appears in that example, but is absent here.

Example 4.2 *Let $\psi(t) = \exp(-0.01t)\cos\frac{2}{32}\pi t$, $8 \le t \le 520$, then ψ is a continuous function (of period 32) with a discontiniuty in the first derivative at $t = 8$. The signal ψ and all its derivatives are weak-IMFs. Both the conjugate operator and the computational Hilbert transform (applied to the sampled function with $\Delta t = 1$) result in a sign changing instantaneous frequency for any choice of the constant C. In Figure 3, we have provided a plot of ψ and the optimal instantaneous frequency (over all possible C) which is computed by the Hilbert transform method. The values of both the continuous and computational results are to within machine precision at the plotted vertices.*

In order to verify the properties of this example, we proceed as earlier in Example 4.1 by first verifying the corresponding fact in the case of discrete signals. We sample $\psi(t)$ uniformly with increment $\Delta = 0.1$ (vector length = 5121) on the interval $[8, 520]$. The graph of the Hilbert transform and corresponding instantaneous frequency of ψ obtained by using Matlab's built-in "hilbert" function are shown in Fig 3.3, parts (a) and (b), respectively. The next proposition shows that although other choices of the constant C may decrease the interval where the instantaneous frequency is negative there is no value for C for which it is nonnegative on $[8, 520]$. Analogous to Example 4.1, it can be shown that the instantaneous frequency changes its sign for any choice of the constant C.

Proposition 4.2 *The function ψ in Example 4.2 and all its derivatives are weak-IMFs on the interval $[8, 512]$ whose instantaneous frequencies computed by the Hilbert transform method change sign for any choice of C.*



a. **b.**

Figure 3: Graphs for Example 4.2: **(a)** the IMF; **(b)** its instantaneous frequency.

Proof. To simplify the notation, we denote by $\tilde{\psi}$ the function in Example 4.2 and use ψ to denote $\tilde{\psi}$ under the required linear change of variable from $[8, 520]$ to $[-\pi, \pi]$ in order to apply the continuous Hilbert transforms to the periodic function. In this case, ψ will be of the form

$$\tilde{\psi}(\tau) = c\ \exp(\alpha\tau)\sin(k\tau)$$

where $k = 16$. Next note that the derivatives of ψ are all of a similar form: $\psi^{(n)}(t) = c_1 e^{\alpha t}\cos(kt + c_2)$. In particular, each derivative is just a constant multiple of ψ with a constant shift of phase and hence are weak-IMF's for any $n = 0, 1, \ldots, \infty$. From formula (4.1) applied at the zeros of ψ we have $\theta_C'(z_j) = -\frac{\psi'(z_j)}{H\psi(z_j)+C}$, where the Hilbert transform is defined through the conjugate operator (see [Z, K]) represented as a principal value, singular integral operator

$$H\psi(z_j) = \frac{1}{\pi}\ \text{p.v.}\ \int_{-\pi}^{\pi} \psi(t)\cot(\frac{z_j - t}{2})\ dt. \tag{4.9}$$

The standard identity

$$\sin(kt)\cot(\frac{t}{2}) = (1 + \cos(kt)) + 2\sum_{\ell=1}^{k-1}\cos(\ell t) \tag{4.10}$$

from classical Fourier analysis permits us to evaluate this expression by

$$H\psi(z_j) = \frac{1}{\pi}\int_{-\pi}^{\pi} e^{\alpha t}T(t)\ dt, \tag{4.11}$$

where T is an even trigonometric polynomial of degree 16 with coefficients depending on the z_j. Hence the values of $H\psi$ at the zeros of ψ can be evaluated exactly (with Maple for example). For $z_1 := \frac{7}{8}\pi$ and $z_2 := \frac{15}{16}\pi$, the corresponding values of the Hilbert transform may be estimated by $H\psi(z_1) \le 0.051$ and $H\psi(z_2) \ge 0.073$ which verifies that $H\psi(z_2) > H\psi(z_1)$.

On the other hand, $\psi'(z_1) = -ke^{\alpha z_1} < 0$ and $\psi'(z_2) = ke^{\alpha z_2} > 0$, therefore $\theta'(z_1) = -\frac{\psi'(z_1)}{H\psi(z_1)+C}$ is negative for $C < -H\psi(z_1) \leq -0.051$ and $\theta'(z_2) = -\frac{\psi'(z_2)}{H\psi(z_2)+C}$ is negative for $C > -H\psi(z_2) \geq -0.073$. From the fact that $-H\psi(z_2) < -H\psi(z_1)$ we conclude that for any C the instantaneous frequency is negative for at least one of the points z_1 or z_2. Finally, for the extrema of ψ, say $t = \xi$ we have $\theta'(\xi) = cH\psi'(\xi)\psi(\xi)$, where c is a positive constant for any choice of C and it is easy to verify that there exists a value ξ such that $H\psi'(\xi)\psi(\xi) > 0$. Hence $\theta'(\xi) > 0$ for any choice of C. □

We observe that, under the relaxed condition allowing the difference between the upper and lower envelopes to be within a given tolerance, ψ and its derivatives up to some finite order are (strong) IMF's and the computational Hilbert transform method produces a narrow bandwidth approximately equal to 0.0625.

The next result provides general information about the behavior of the instantaneous frequency θ' from any polar representation of $\psi(t) = r(t)\sin\theta(t)$ in terms of a relation between the amplitude r and ψ.

Lemma 4.1 *Suppose that ψ is a weak-IMF, $r(t) > 0$ is an amplitude such that $\psi(t) = r(t)\cos\theta(t)$. Further, suppose that at some point $t = \tau$, $|\psi(\tau)| \neq r(\tau)$ and $\psi(\tau) \neq 0$. A necessary and sufficient condition for $\theta'(\tau)$ to vanish is that*

$$\frac{\psi'(\tau)}{\psi(\tau)} = \frac{r'(\tau)}{r(\tau)}, \tag{4.12}$$

that is, that the logarithmic derivative of $\frac{\psi}{r}$ should vanish at $t = \tau$.

Proof. Since $r > 0$ we can differentiate the relation $\cos\theta = \frac{\psi}{r}$ and get

$$-\theta'\sin(\theta) = \frac{\psi' r - r'\psi}{r^2} = \frac{\psi}{r}\left(\frac{\psi'}{\psi} - \frac{r'}{r}\right). \tag{4.13}$$

To prove necessity, suppose that $\theta'(\tau) = 0$. Then since $\psi(\tau) \neq 0$, it follows from the identity (4.13), that $\frac{\psi'(\tau)}{\psi(\tau)} = \frac{r'(\tau)}{r(\tau)}$.

To prove sufficiency it is enough to notice that in the event the left-hand side of (4.13) vanishes at $t = \tau$, but $\theta'(\tau) \neq 0$, then $\sin\theta(\tau)$ must vanish. Hence $|\cos\theta(\tau)| = 1$ or $|\psi(\tau)| = r(t)$, which is a contradiction. Hence $\theta'(\tau) = 0$. □

Looking back, one can see that Lemma 4.1 can be used to motivate the proof of the characterization theorem for weak-IMF's (Theorem 3.1). Indeed, for the envelopes constructed there, $\frac{r'}{r}$ and $\frac{\psi'}{\psi}$ were forced to have different signs and therefore they can not be equal at any point that is not a zero of ψ. From Lemma 4.1, it follows that θ' does not change sign between any two zeros of ψ. Since θ' is continuous and was forced to be nonzero at the zeros of ψ, we have that θ' cannot change sign.

Proposition 4.3 *Let $v_a(t)$ be an even function defined on $(-\pi, \pi]$ such that $v_a(0) = 0$ and $e^{v_a}/\|e^{v_a}\|_{L_1} \to \delta$, where δ is the Dirac delta function. Define $\psi_a(t) := e^{v_a(t)}\cos(kt)$. Then*

there exists a value of a_0 sufficiently large such that the analytic instantaneous frequency for ψ_a for any $a > a_0$ changes sign for any choice of the constant used in defining the Hilbert transform $H\psi_a$.

Proof: Recall from (4.9) that a Hilbert transform of ψ_a at a point $t \in (-\pi, \pi]$ is $H\psi_a(t)+C$, where C is an arbitrary real constant and

$$H\psi_a(t) = \text{p.v.}\frac{1}{\pi}\int_{-\pi}^{\pi} \psi_a(\tau)\cot\frac{t-\tau}{2}\,d\tau \tag{4.14}$$

is the conjugate operator for periodic functions. Using two applications of the identity (4.1), we observe that the analytic method produces an instantaneous frequency of the form

$$\theta'_C = \frac{\psi_a H\psi'_a - (H\psi_a + C)\psi'_a}{(H\psi_a + C)^2 + \psi_a^2} = R\theta'_0 - C\,L\psi'_a, \tag{4.15}$$

where R and L are positive functions on $(-\pi, \pi]$. Let $z_j = \frac{2j-1}{2k}\pi$, $-k+1 \le j \le k$ be the zeros of ψ_a, then

$$sgn(\psi'_a(z_j)) = (-1)^j \tag{4.16}$$

and by (4.15), with $C = 0$, it follows that

$$\theta'_0(z_j) = -\frac{\psi'_a(z_j)}{H\psi_a(z_j)} \tag{4.17}$$

and consequently

$$sgn(\theta'_0(z_j)) = (-1)^{j+1} sgn(H\psi_a(z_j)). \tag{4.18}$$

The proof of the lemma will be completed if we can show that for sufficiently large a there is an index J for which two consecutive values of $H\psi_a(z_j)$ have the same sign

$$sgn\,(H\psi_a(z_J)) = sgn\,(H\psi_a(z_{J+1})) =: \sigma. \tag{4.19}$$

When $C = 0$ this follows immediately from equation (4.18). For $C \neq 0$, we use the analogue of (4.18),

$$sgn(\theta'_C) = -sgn(\psi'_a)\; sgn(H\psi_a + C) \tag{4.20}$$

which follows immediately from (4.15). In the case $sgn(C) = \sigma$, this last identity shows that θ'_C has different signs at the endpoints of (z_J, z_{J+1}), since ψ'_a does. For the final case, $sgn(C) = -\sigma$, we observe that $H\psi_a$ and ψ'_a are both odd functions since ψ_a is even. By considering $-z_J$ and $-z_{J+1}$ in place of z_J and z_{J+1}, we see that $sgnH(\psi_a) = -\sigma = sgn(C)$ at these two points and so once again from (4.20), θ'_C has different signs at the endpoints. Hence by the continuity of θ'_C, there are nonempty intervals where the instantaneous frequency takes on opposite sign.

Therefore to complete the proof, we must verify (4.19), i.e., for parameter $a > 0$ sufficiently large, there is a pair of consecutive points z_J, z_{J+1}, such that $H(\psi_a)$ does not change

sign. Evaluating the conjugate operator at the zeros $x = z_j$ in (4.9), we can proceed as in Proposition 4.2 using the periodicity of ψ_a and the trigonometric identity (4.10), to obtain

$$\begin{aligned} H\psi_a(z_j) &= c\int_{-\pi}^{\pi} e^{v_a(t)}\cos(kt)\cot\frac{z_j-t}{2}\,dt \\ &= c\int_{-\pi}^{\pi} e^{v_a(t+z_j)}\sin(kt)\cot\frac{t}{2}\,dt \\ &= c\int_{-\pi}^{\pi} e^{v_a(t+z_j)}P_k(t)\,dt, \end{aligned}$$

where in the last identity $P_k(t)$ is a trigonometric polynomial of degree k. Therefore it follows that

$$\lim_{a\to\infty}\frac{H\psi_a(z_j)}{\|e^{v_a}\|} = P_k(0) = -\cot\frac{z_j}{2}.$$

holds. For any $z_m, z_{m+1} \in (0,\pi)$ it is follows easily that there exists $a > 0$ such that $sgn(H\psi_a(z_m)) = sgn(H\psi_a(z_{m+1}))$. Hence for a is sufficiently large, ψ_a is a weak-IMF. □

We note that the arguments in Proposition 4.3 can also be used to explain the behavior of θ' in Example 4.2.

Example 4.3 *We illustrate in Figure 4 the use of Proposition 4.3 in producing additional weak-IMF's with non-monotone phase. For the sample function ψ, we set $k = 16$ and let v_a be a gaussian with standard deviation $s = .01$ and centered at the origin. The perturbation is applied at both $t_1 = 0$ and $t_2 = \pi/32$. The function ψ is displayed in part (a), its hilbert transform in part (b), and its instantaneous frequency in part (c).*

For this same function, in Figure 4(d) we illustrate the application of Lemma 4.1. The instantaneous frequency changes sign when the logarithmic derivative of $\frac{\psi}{r}$ vanishes at points other than at an acceptable zero: either a zero of (i) ψ or of (ii) its Hilbert transform, i.e. points where $|f| = r$. Notice that the endpoints of the two intervals where the instantaneous frequency becomes negative corresponds precisely to the four (non-acceptable) zeros of the logarithmic derivative of ψ/r

Example 4.4 *An informative example of function which may be considered a true IMF is given by the function*

$$\psi(t) = (t^2+2)\cos(\pi\sin(8t))/16, \quad -4\pi \le t \le 4\pi \tag{4.21}$$

which, along with its instantaneous frequency, is plotted in Figure 5. Notice that t^2+2 may be regarded as an envelope of ψ and that it is close but different than the upper envelope produced by a cubic spline fit through the maxima. Recall from the observation in Section 2.1 that a necessary and sufficient for the IMF envelope condition (b) of Definiton 2.1 to be satisfied is that those envelopes reduce to a quadratic polynomial. This example shows that the sifting convergence criterium in the EMD process for measuring the difference of the absolute values of the upper and lower envelopes should be chosen with care. If fact, it can be

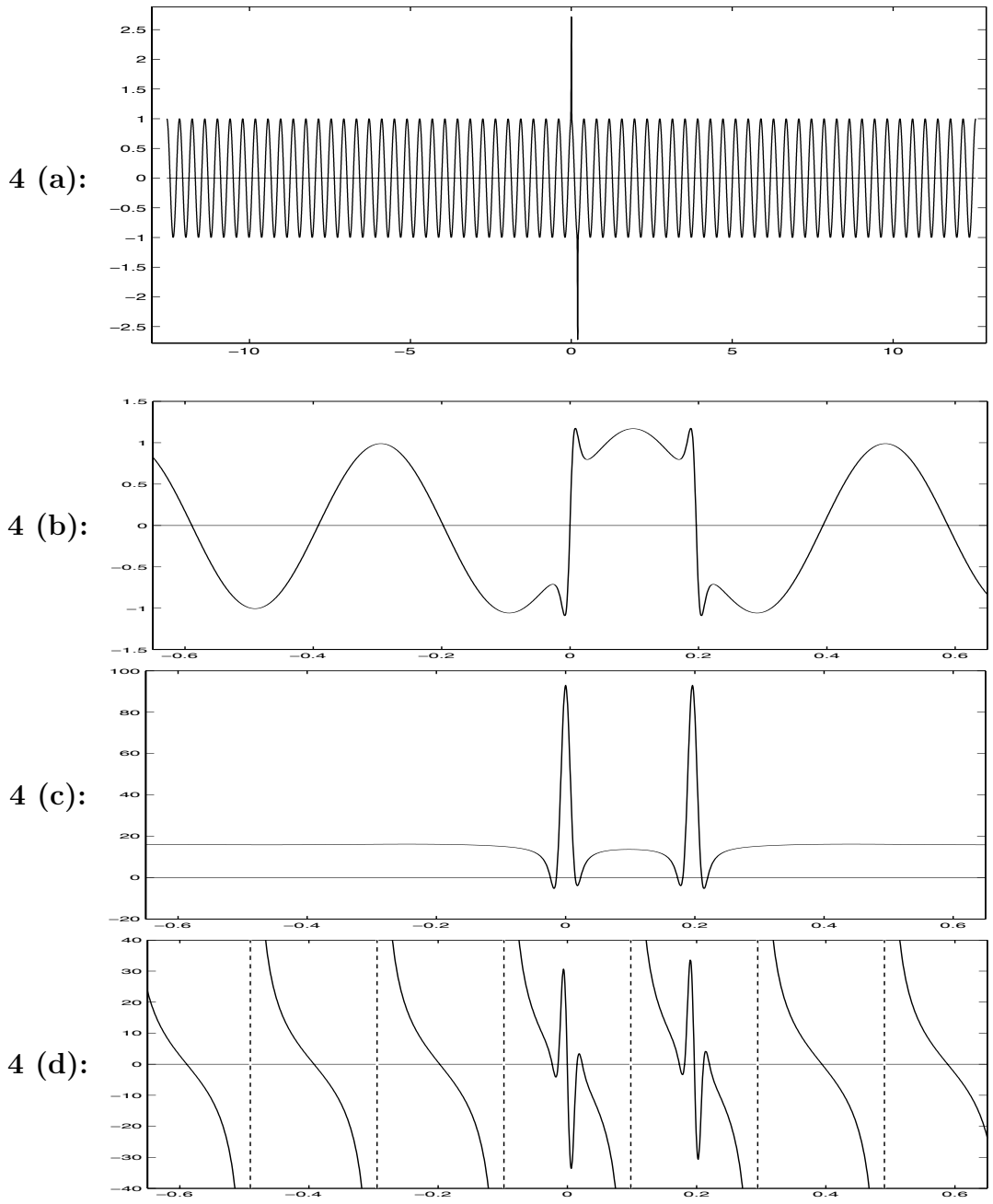
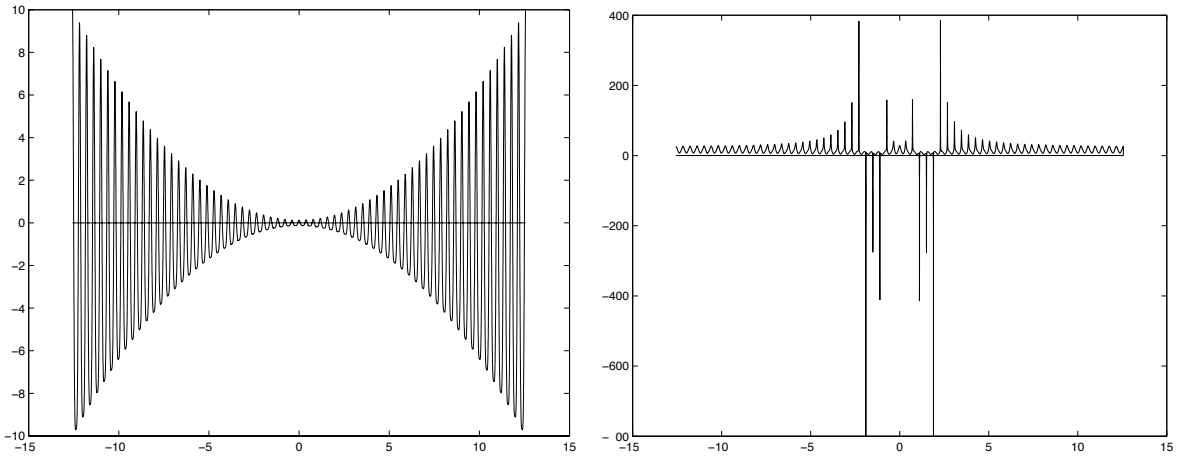


Figure 4: Sample IMF from Example 4.3: **a.** Cosine signal with strong perturbation at 0 and $\pi/16$; **b.** the Hilbert tranform of ψ near the perturbations; **c.** the instantaneous frequency of ψ near the perturbations; **d.** the logarithmic derivative of ψ/r near the perturbations.



a. **b.**

Figure 5: Plot of the example IMF defined by equation (4.21): **a.** The IMF ψ with a parabolic envelope, **b.** Instantaneous frequency.

easily verified that smoothing off the endpoint data of this example will result in a function for which the EMD residual can be made visually negligible after a single sifting. There are many variations of these examples to produce similar behavior.

We conclude this section with a procedure that adds a smooth disturbance at an appropriate scale to any reasonable function in such a way that the function maintains its smoothness and analytical profile, that is, no additional extrema are introduced and the existing extrema are only perturbed, but the resulting function has non-monotone analytic phase. By a *reasonable* function ψ, we will mean an IMF in the strongest sense, which we call a Hilbert-IMF.

Definition 4.1 *A function is called a Hilbert-IMF if it satisfies the following conditions:*

(i.) *ψ is an IMF in the sense of the definition in [H98];*

(ii.) *the analytic signal of ψ (i.e. via the Hilbert transform) $\Psi = re^{i\theta}$, r' and θ' are smooth functions and $\theta' > 0$;*

(iii.) *the weighted L_2 bandwidth is small.*

The idea of the perturbation procedure is based on the fact that by multiplying the analytic function $\Psi = re^{i\theta}$ by another analytic function, say $\Gamma = r_1 e^{i\theta_1}$, the result $\Psi\Gamma = rr_1 e^{i(\theta+\theta_1)}$ is also analytic with analytic amplitude rr_1 and analytic phase $\theta + \theta_1$. Since θ' and θ_1' are smooth functions, in order to force the instantaneous frequency of $\Psi\Gamma$ to change sign it suffices to choose Γ such that $\theta_1'(T) < -\theta'(T)$ at some point T. One way we can insure that the weighted L_2 bandwidth of $\Psi\Gamma$ remains small is to localize the perturbation Γ to a small interval I, i.e. both r_1 and θ_1 should decay rapidly to zero outside I. Further, to guarantee that the real part of $\Psi\Gamma$ is an IMF in the sense of [H98], the added perturbation must only result in a small deviation of the zeros and extrema of the original IMF Ψ, and

should not introduce additional zeros nor extrema. This is achieved by incorporating an additional tuning parameter σ into our perturbation Γ^σ, for small values of σ.

The technique used in the proof of the results below shows that there are many functions that can be used as such perturbation functions. We consider one particular smooth function that is constructed from the modified Poisson kernel

$$y_\rho(t) := \frac{(1-\rho)^2}{1-2\rho\cos t+\rho^2}, \;\; 0 \le \rho < 1. \tag{4.22}$$

It is well known that its conjugate function is $Hy_\rho(t) = \frac{1-\rho}{1+\rho}\frac{2\rho\sin t}{1-2\rho\cos t+\rho^2}$. Although it is an abuse of notation, we will refer to these simply as y and Hy with the understanding that the parameter ρ is implicitly present. We define the perturbation Γ in terms of $y = y_\rho$ by

$$\Gamma = \Gamma_\rho := exp(-Hy + iy). \tag{4.23}$$

and observe that as the parameter ρ approaches 1 from below, the function Γ_ρ becomes very localized.

The perturbed IMF is set to $\Re(\Psi\Gamma^\sigma) = re^{-\sigma Hy}\cos(\theta+\sigma y)$ for some real σ. The idea in brief is to select σ small and ρ sufficiently close to 1 so that the change in the functional values are also small, i.e., the zeros, extrema, and the extremal values are perturbed slightly from the original IMF. On the other hand, the corresponding instantaneous frequency becomes $\theta' + \sigma y'$ (see Lemma 4.2 below). Moreover, y' has one local minimum that is negative with magnitude depending on ρ. In the special case when $r = e^{At}$ and $\theta = mt$ on an interval (the length of the interval can be arbitrarily small), we prove in Corollary 4.2 that there exists a subinterval and values of the parameters σ and ρ such that under mild conditions, the perturbed function satisfies all the properties i-iii), but has non-monotone phase. From the proof and by continuity it is then clear that the new instantaneous frequency can be made negative on an interval while preserving all other properties listed in i-iii). A similar result holding for more general functions is established in Corollary 4.3.

To show that the perturbed function is a weak-IMF we utilize the logarithmic derivative as $\rho \to 1^-$ and the following technical Lemma, where we establish that the maximum of Hy' and its value at the minimum of y' behave asymptotically as a finite multiple of the minimum value of y'.

Lemma 4.2 *Let $0 \le \rho < 1$ and $y = y_\rho$ be defined as in equation (4.22), then the following properties hold:*

(a) for all $\sigma > 0$ the function Γ^σ defined in equation (4.23) is analytic with amplitude $e^{-\sigma Hy}$ and phase σy;

(b) if $-\mu := \min y' = y'(t_0)$, then $\lim\limits_{\rho\to 1^-} \dfrac{Hy'(t_0)}{\mu} = \dfrac{1}{\sqrt{3}}$.

(c) $\lim\limits_{\rho\to 1^-} \dfrac{\mu}{\max|Hy'|} = \dfrac{3\sqrt{3}}{8}$.

(d) The function Hy' is even with exactly one positive zero, t_z which satisfies $0 < t_0 < t_z$ and $\lim\limits_{\rho\to 1^-} \dfrac{y'(t_z)}{\mu} = 0$.

Proof. Part (a) follows from the construction of the analytic function Γ and the fact that $|\Gamma| > 0$. To establish part (b) we determine the minimizer of y', which we denote t_0, from the equation $y''(t_0) = 0$, which is equivalent to the equation

$$2\rho\cos^2 t_0 + (1+\rho^2)\cos t_0 - 4\rho = 0. \tag{4.24}$$

Hence there exists a unique solution t_0 which satisfies the relation $\cos t_0 = \frac{D-(1+\rho^2)}{4\rho}$, where $D := \sqrt{\rho^4 + 34\rho^2 + 1}$. Substituting this explicit formula for $\cos t_0$ into the expression for $y'(t_0)$, the minimum value of y' can be written as

$$y'(t_0) = -2(1-\rho)^2 \frac{\sqrt{2(1+\rho^2)D - (2\rho^4 + 20\rho^2 + 2)}}{(3\rho^2 + 3 - D)^2}. \tag{4.25}$$

Proceeding similarly with the expression for $Hy'(t)$ given by

$$Hy'(t) = -\frac{2\rho(1-\rho)}{1+\rho}\frac{(1+\rho^2)\cos t - 2\rho}{(1-2\rho\cos t + \rho^2)^2} \tag{4.26}$$

we find, after algebraic rationalization and simplification, that

$$\frac{Hy'(t_0)}{\mu} = \frac{2\sqrt{2}\,\rho}{\sqrt{(1+\rho^2)D + \rho^4 + 10\rho^2 + 1}}. \tag{4.27}$$

Part (b) follows immediately by taking the limit as $\rho \to 1^-$.

Part (c) is established in a similar manner by observing from equation (4.26) that $\max|Hy'| = Hy'(0) = \frac{2\rho}{1-\rho^2}$ and so, using the identity (4.25), it follows that $\frac{\mu}{Hy'(0)}$ converges to $\frac{3\sqrt{3}}{8}$ as $\rho \to 1^-$.

Finally, for part (d) we determine from (4.26) that zeros of Hy' are exactly the roots of the equation $\cos t = \frac{2\rho}{1+\rho^2}$. Substituting this expression for $\cos t_z$ into the left-handside of the equation (4.24) for $\cos t_0$, we get a negative value $-\frac{2\rho(1-\rho^2)^2}{(1+\rho^2)^2}$ for the quadratic expression and so $\cos t_z < \cos t_0$, which is equivalent to $t_0 < t_z$. Observing that $\sin t_z = \frac{1-\rho^2}{1+\rho^2}$, we can use this identity to evaluate $y'(t_z)$ to obtain $\frac{y'(t_z)}{\mu} = -\frac{2\rho(1-\rho)}{(1+\rho)\mu} \to 0$ as $\rho \to 1^-$. □

In Corollary 4.2 we prove in the special case $r = e^{At}$, $A \leq 0$ and $\theta = mt$ that we can find values of σ and ρ such that the procedure described above produces a desired function satisfying the properties i-iii) but whose analytic instantaneous frequency changes sign. We first prove a milder version in the following proposition, and then modify the parameter σ to establish the stronger version.

Proposition 4.4 *Let the notation be as in the previous lemma (Lemma 4.2). In particular, let t_0 be the point which provides a global minimum for y', t_z be the positive zero of Hy' and $\mu := -y'(t_0)$. Suppose further that $A \leq 0$. If $\psi(t) := exp(At)\cos mt$, then there exist a constant ρ^* and a point t^* such that for $\rho* < \rho < 1$*

$$\tilde{\psi}(t) = exp(At - \frac{m}{\mu}Hy(t-t^*))\cos(mt + \frac{m}{\mu}y(t-t^*)), \tag{4.28}$$

is a weak-IMF, but its analytic instantaneous frequency vanishes at $t_0 + t^*$. Moreover, the difference between absolute values of the upper and lower cubic spline envelopes is small except at the endpoints in the case that A is negative.

Proof. From the previous discussion and from the representation $\tilde{\psi}(t) = \Re(\Psi(t)\Gamma^{\frac{m}{\mu}}(t-t^*))$ it is clear that the analytic phase of $\tilde{\psi}$ is $\tilde{\theta} = mt + \frac{m}{\mu}y(t-t^*)$. The definition of μ implies that the expression $m + \frac{m}{\mu}y'(t-t^*)$ is nonnegative and vanishes only at the point $t_0 + t^*$, hence $\tilde{\theta}$ is strictly increasing. Furthermore, we show that if ρ is close to 1, the rapid decay of $y(t-t^*)$ away from $t_0 + t^*$ will insure that the zeros of ψ and $\tilde{\psi}$ are the same in number and are separated only slightly from one another.

We may assume that the perturbation $y = y_\rho$ is added between a maximum of ψ and the zero τ_0 immediately following; the other three situations can be handled in the same way with an appropriate changes of the signs of the corresponding expressions. Denote by τ_- the nearest point less than τ_0 which satisfies $\tan m\tau_- = \frac{32}{3\sqrt{3}}$. Any point from the interval (τ_-, τ_0) can be picked for t^*. We select $t^* := \frac{\tau_0+\tau_-}{2}$, $\delta := \frac{\tau_0-\tau_-}{4}$ and set $\Delta := (t^* - \delta, t^* + \delta)$. By construction it is clear that functions y, Hy, y', and Hy' (translated by t^*) tend uniformly to zero outside the interval Δ as ρ approaches 1^-. Hence $\tilde{\psi}$ uniformly tends to ψ outside Δ.

Since ψ and ψ' have only simple zeros, it follows that there exists ρ_1 such that for any $1 > \rho > \rho_1$ the perturbed function $\tilde{\psi}$ is a weak-IMF; even more, for each zero and extrema of ψ there corresponds exactly one zero and extrema of $\tilde{\psi}$. To prove this, we consider the functions on three disjoint sets, a subinterval Δ^* of Δ (to be determined), the set $\Delta \backslash \Delta^*$, and the complement of Δ.

We first consider the set of values t in the complement of Δ. Assume that there is a sequence of $\rho' s$ approaching 1 from below so that there is an extrema of ψ, say τ_e, such that in a neighborhood of that extrema $\tilde{\psi}$ has at least three extrema (the extrema must be odd in number since $\tilde{\psi}$ tends uniformly to ψ). By Rolle's theorem and the uniform convergence as $\rho \to 1^-$, it follows that ψ' has a multiple zero at τ_e, which is a contradiction. Hence outside Δ, $\tilde{\psi}$ is an IMF.

Consider now the interval Δ. To follow the changes of the extrema of $\tilde{\psi}$, we consider the logarithmic derivative of $\tilde{\psi}$ given by

$$\tilde{L} := \frac{\tilde{\psi}'}{\tilde{\psi}} = A + \frac{m}{\mu}(-Hy'(t-t^*)) - (m + \frac{m}{\mu}y'(t-t^*))\tan(mt + \frac{m}{\mu}y(t-t^*)).$$

We show $\tilde{L}$ is negative on Δ. It then follows that $\tilde{\psi}$ has no additional extrema on this interval, and so $\tilde{\psi}$ will be an IMF in the sense of the definition in [H98], but its analytic instantaneous frequency $m + \frac{m}{\mu}y'$ has a zero.

To see that $\tilde{L}$ is negative on Δ, observe that for a fixed ρ $(1 > \rho > \rho_1)$ since $A \leq 0$, $\tan(mt + \frac{m}{\mu}y(t-t^*)) > 0$, and $(m + \frac{m}{\mu}y'(t-t^*)) \geq 0$ it follows that $\tilde{L} < 0$ on the interval $\Delta^* := (t^* - t_z, t^* + t_z)$, where t_z is specified in part (d) of Lemma 4.2. From the proof of that Lemma, we see that t_z approaches 0 as ρ approaches 1^- and hence we can pick ρ_2 $(1 > \rho_2 > \rho_1)$ such that $\Delta^* \subset \Delta$ for each ρ for which $1 > \rho \geq \rho_2$.

On the other hand, for any t outside Δ^*, we have that $m + \frac{m}{\mu}y'(t) > m + \frac{m}{\mu}y'(t_z)$ and from Lemma 4.2(d) it follows that there exists $1 > \rho_3 > \rho_2$ such that the inequality $m + \frac{m}{\mu}y'(t) > \frac{m}{2}$ holds for any $1 > \rho > \rho_3$ and any $t \in \Delta\backslash\Delta^*$. Finally from Lemma 4.2, we can pick $1 > \rho^* > \rho_3$ such that the inequality $\frac{\max|Hy'|}{\mu} < \frac{16}{3\sqrt{3}}$ holds for any $\rho > \rho^*$. The choice of the point t^* and the fact that y is a positive function provide the inequality

$$\tan(mt + \frac{m}{\mu}y(t - t^*)) > \tan(m\tau_-) = \frac{32}{3\sqrt{3}}$$

for any $t \in \Delta\backslash\Delta^*$. Using the above estimates and the assumption $A \leq 0$ we have that

$$\tilde{L} < m\frac{\max|Hy'|}{\mu} - \frac{m}{2}\tan(mt + \frac{m}{\mu}y(t - t^*)) < 0,$$

for any $t \in \Delta\backslash\Delta^*$ and any $1 > \rho > \rho^*$, which completes the proof. □

Corollary 4.2 *Let ψ be the Hilbert-IMF considered in Proposition 4.4. Then there exist $\sigma > \frac{m}{-\mu}$ such that the small perturbation of ψ, $\tilde{\psi} = \Re(\Psi\Gamma^\sigma)$, satisfies parts (i) and (iii) of the definition of a Hilbert-IMF, but has does not have monotone analytic phase.*

Proof. Since $\tilde{\psi}$ and all other related functions considered in Proposition 4.4 depend continuously on σ, and for $\sigma = -\frac{m}{\mu}$ we have that $\tilde{\theta}'$ vanishes only at the point $t_0 + t^*$, it follows that any increase of σ forces the instantaneous frequency to be negative in a neighborhood of that point. On the other hand, by the choice of ρ^*($1 > \rho^* > 0$) from Proposition 4.4 and the uniform convergence, it follows that there exists $\sigma^* > -\frac{m}{\mu}$ such that the perturbed function $\tilde{\psi}$ is a weak-IMF outside Δ and $\tilde{L}_{\sigma^*}$ is negative on Δ; i.e. there are no additional zeros and extrema on Δ. Hence $\tilde{\psi}$ is a nicely behaved function with analytic instantaneous frequency that changes sign on an interval of positive measure. □

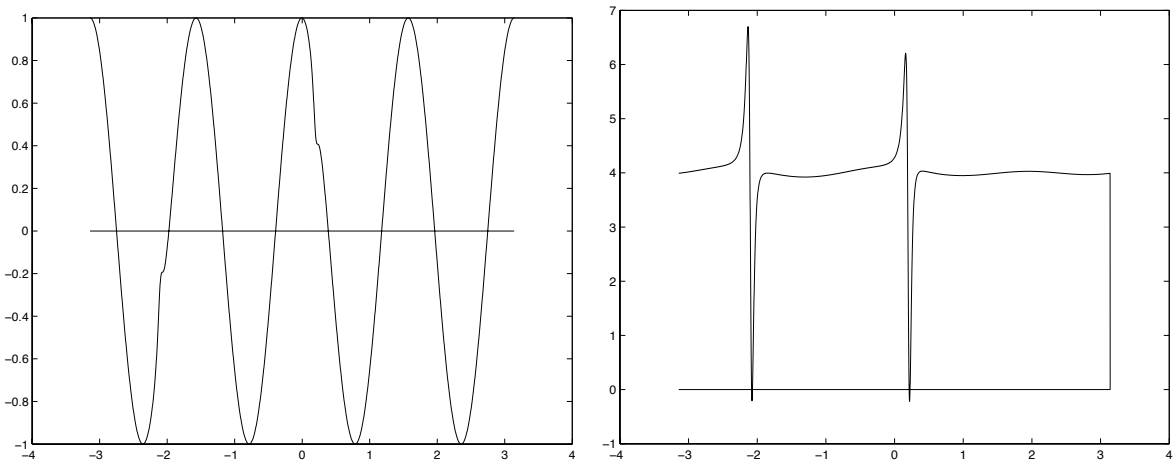
Remark 4.2 *The condition $A \leq 0$ can be relaxed to $A < \frac{m}{\sqrt{3}}$ which agrees with the estimate in Lemma 4.2(b). The proof of the Proposition 4.4 with that restrictiion requires further technical estimates as in Lemma 4.2(d) for a point t_η such that $0 < t_0 < t_\eta$ and*

$$\lim_{\rho\to 1^-}\frac{y'(t_\eta)}{\mu} = \eta - 1$$

for a fixed $0 < \eta < 1$. Details of the estimates are similar so we do not include them here.

Example 4.5 *To illustrate the above construction, we consider the function $\Psi = e^{4it}$, $-\pi < t \leq \pi$, and apply the procedure twice for $\rho = 0.95$ and $\sigma = 0.31$, once at the point $t_1 = -2.1$ and again at the point 0.2. The resulting signal and its analytic instantaneous frequency are shown on Figure 6.*

For any nice function (e.g. a Hilbert-IMF) $\psi = r\cos\theta$ with only simple zeros, it is clear from the identity $\frac{\psi'}{\psi} = \frac{r'}{r} - \tan\theta\,\theta'$ that if there exist a zero for which $r' \neq 0$ then the logarithmic derivative $\frac{\psi'}{\psi}$ and the instantaneous bandwidth $\frac{r'}{r}$ have the same sign in a one-sided neighborhood of that zero. Then if the perturbation is added in that neighborhood,



a. **b.**

Figure 6: Plot of the IMF in Example (4.5): **a.** an IMF perturbed by a smooth perturbation; **b.** Instantaneous frequency.

the proof of Proposition 4.4, without significant modifications, can be used to prove similar results so long as $\frac{r'}{r} + \frac{9}{8\sqrt{3}} \max_I \theta'$ is negative in a neighborhood of that zero, as established in the next Corollary. We note that by choosing ρ closer to 1, the interval I can be made arbitrarily small.

Corollary 4.3 *Let $\psi = r\cos\theta$ be the restriction to the circle of any function analytic in a disk of radius larger than one. Assume that ψ is an IMF in the sense of [H98] with amplitude r and monotone phase θ which are defined using the Hilbert transform. Suppose further that ψ has only simple zeros, that there exists a zero z_0 of ψ for which $\frac{r'}{r}$ and $\frac{\psi'}{\psi}$ have the same sign in a one-sided neighborhood I of z_0, and that $\frac{r'}{r} + \frac{9}{8\sqrt{3}} \max_I \theta' < 0$ on I, then there exist parameters ρ^* and σ^*, and a point t^* such that the function $\tilde{\psi}(t) = \Re(\Psi(t)\Gamma^{\sigma^*}(t - t^*))$ is an IMF with zeros and extrema which are close perturbations of those of ψ, but with an analytic instantaneous frequency $\tilde{\theta}'(t) = \theta' + \sigma y'(t - t^*)$ which changes sign.*

Proof. As in Proposition 4.4, it is enough to prove that $\tilde{\theta}'(T) = 0$ for some point T. We may assume from the hypothesis that $\theta' > 0$, then the modified instantaneous frequency $\tilde{\theta}(t) = \theta'(t) + \sigma y'(t - t^*) > 0$ for small σ. Hence if σ is continuously increased, by continuity we will reach a value σ_1 for which $\tilde{\theta}'(T) = 0$ for some point T and is positive elsewhere. If ρ is then increased close to 1, and σ_1 is adjusted accordingly, we can localize the perturbation on an arbitrarily small interval with $\tilde{\theta}'$ vanishing at a point.

Notice that the instantaneous bandwidth does not change sign on an interval that contains both I and z_0 as an interior point. We may assume that $L < 0$ on I, then the logarithmic derivative of $\tilde{\psi}$ is

$$\tilde{L}(t) = \frac{r'(t)}{r(t)} - \sigma H y'(t - t^*) - \tilde{\theta}'(t) \tan\tilde{\theta}(t).$$

From the choice of σ it follows that $\max_I \theta' - \sigma\mu \geq 0$ and hence $\sigma \leq \frac{\max_I \theta'}{\mu}$. For the new instantaneous bandwidth as $\rho \to 1^-$, we have $\frac{r'}{r} - \sigma H y'(t-t^*) \leq \frac{r'}{r} + \max_I \theta' \frac{max|Hy'|}{\mu} \to \frac{r'}{r} + \frac{9}{8\sqrt{3}} \max_I \theta' < 0$ on I, where the last inequality is the assumption on ψ relating instantaneous bandwidth and frequency. Since I is a finite interval, there exists ρ^* such that $\tilde{L} < 0$ for any ρ satisfying $1 > \rho > \rho^*$. On the other hand, since $\tilde{\theta}' \geq 0$ and $\tan\tilde{\theta} > 0$ on I it follows that $\tilde{L} < 0$ on I. All other steps are the same as in the proof of Proposition 4.4. □

Remark 4.3 *The condition on the logarithmic derivatives (of the function ψ and its amplitude r) to have the same sign in a neighborhood of a zero is equivalent to $r'(z_0) \neq 0$, or $r'(z_0) = 0$ but $r'\psi'\psi > 0$ on I. In other words if all other requirements are met, the procedure works more generally than in the case of envelopes considered in the Theorem 3.1.*

Remark 4.4 *The procedure for adding perturbations to a nice function can be also used for removing certain types of noise. If a function has a negative instantaneous frequency on some small interval, then we can apply the procedure with a perturbation Γ^σ using negative σ in order to remove negative instantaneous frequencies, but still preserve the general features (zeros, local extrema) and smoothness class of the original function.*

References

[BS] C. Bennett and R. Sharpley, Interpolation of Operators, Academic Press, New York, 1988.

[BR] G. Birkhoff and G.C. Rota, Differential Equations, (4-th ed) Wiley, New York, NY, 1989.

[C95] L. Cohen, Time-Frequency Analysis, Prentice-Hall Signal Processing Series, Upper Saddle River, NJ, 1995.

[H98] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. Royal Soc. Lond. A **454** (1998), 903-995.

[H99] N.E. Huang, Z. Shen, and S.R. Long, A new view of nonlinear water waves: the Hilbert spectrum, Annu. Rev. Fluid Mech. **31** (1999), 417–457.

[K] Y. Katznelson, An Introduction to Harmonic Analysis, (2-nd ed.), Dover Pubns, 1976.

[M] S.L. Marple, Computing the discrete-time analytic signal via FFT, IEEE Transactions on Signal Processing **47** (1999), 2600-2603.

[P] M.A. Pinsky, Introduction to Fourier Analysis and Wavelets, Brooks/Cole, Pacific Grove, 2002.

[S] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, 1970.

[V] V. Vatchev, Intrinsic Mode Functions and the Hilbert Transform, Ph.D Dissertation, Department of Mathematics, University of South Carolina, 2004.

[Z] A. Zygmund, Trigonometric Series, Cambridge University Press, Cambridge, 1969.

Industrial Mathematics Institute
Department of Mathematics
University of South Carolina
Columbia, SC 29208

E-mail: sharpley@math.sc.edu, vatchev@math.sc.edu