Instantaneous phase of rhythmic behaviour under volitional control

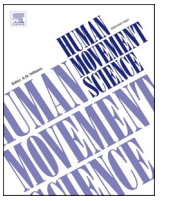
Leonardo Lancia

Human Movement Science 96 (2024) 103249

Contents lists available at ScienceDirect

Human Movement Science

journal homepage: www.elsevier.com/locate/humov



Instantaneous phase of rhythmic behaviour under volitional control

Leonardo Lancia *

Laboratoire Parole et Langage, Aix-Marseille Université / CNRS, 5 av. Pasteur, 13100 Aix-en-Provence, France

ARTICLE INFO

Keywords:
Instantaneous phase
Empirical mode decomposition
Masked sifting
Demodulated oscillations
Intentional actions
Biological rhythms

ABSTRACT

The phase of signals representing cyclic behavioural patterns provides valuable information for understanding the mechanisms driving the observed behaviours. Methods usually adopted to estimate the phase, which are based on projecting the signal onto the complex plane, have strict requirements on its frequency content, which limits their application. To overcome these limitations, input signals can be processed using band-pass filters or decomposition techniques. In this paper, we briefly review these approaches and propose a new one. Our approach is based on the principles of Empirical Mode Decomposition (EMD), but unlike EMD, it does not aim to decompose the input signal. This avoids the many problems that can occur when extracting a signal's components one by one. The proposed approach estimates the phase of experimental signals that have one main oscillatory component modulated by slower activity and perturbed by weak, sparse, or random activity at faster time scales. We illustrate how our approach works by estimating the phase dynamics of synthetic signals and real-world signals representing knee angles during flexion/extension activity, heel height during gait, and the activity of different organs involved in speech production.

1. Introduction

1.1. The phase of goal-oriented behaviour

Many goal-oriented behaviours, such as walking, chewing, and speaking, display cyclic patterns. Often the observed cycles are nested within each other, forming hierarchies of oscillatory activity that characterize the behaviour of potentially many observable variables. In studies focusing on the organisational principles underlying these rhythmic patterns of activity, the notion of phase, representing the position in the cycle at each instant, plays a central role (Kelso, 1999; Schöner & Kelso, 1988). Measuring the relations between events in terms of phase lags permits measuring durations in terms that are meaningful to the functioning of the systems under study, because a cycle of activity is a meaningful unit of description of an oscillatory behaviour (Longo, 2021; Pikovsky, Rosenblum, & Kurths, 2003). At the same time, the phase of a signal is an abstract quantity whose meaning does not depend on definition of the systems under study. This is therefore an ideal quantity to characterize the relations between heterogeneous signals and systems. Unfortunately, the implementation of the notion of phase is not straightforward for general time series with arbitrary cycle shapes (Huang et al., 2009). The usual approach consists in projecting the behaviour of an observed real-valued signal x on that of a signal drawing circles on a plane. If the input signal represents the behaviour under study in the tasks' coordinate system, the circle is that

* Corresponding author.
E-mail address: leonardo.lancia@cnrs.fr.

https://doi.org/10.1016/j.humov.2024.103249
Received 26 December 2023; Received in revised form 18 June 2024; Accepted 21 June 2024
Available online 23 July 2024

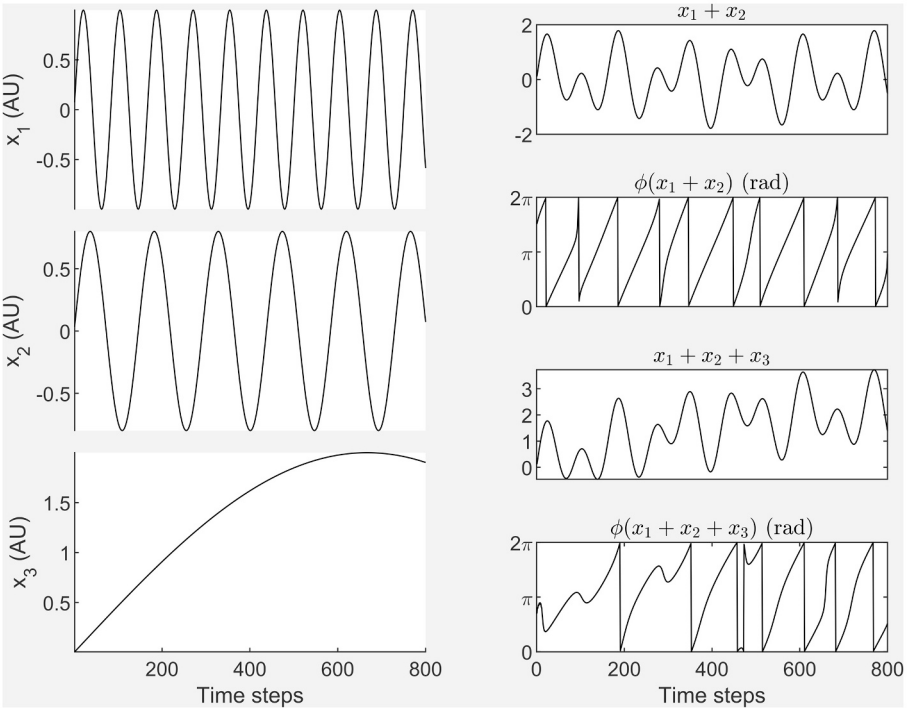


Fig. 1. Issues in the application of Hilbert transform. Left column panels: sinusoidal signals composing the time series displayed in the right column (AU: Arbitrary Units). Right column panels: The sum of the first and second panel from the top in the left column is displayed in the topmost panel. Each extremum is surrounded by two zero-crossing points, but amplitude changes relatively rapidly with respect to phase. Consequently, the dynamics of the signal's Hilbert phase (displayed in the second panel from top) feature a degree of variability not observed in the (first component of) the input signal. The sum of all time series from the left column is displayed in the third panel from top and its Hilbert phase in the bottommost panel. In this case, the phase does not increase from 0 to 2π in each cycle and discontinuities are observed. In both cases signals where normalized with respect to their mean and standard deviation before application of the Hilbert transform.

drawn by the point (x, y) where $y = \dot{x}$ is the derivative of x. In many cases however, the observed univariate signal represents an indirect measure of a behaviour defined in a potentially multidimensional space. In these cases, the Hilbert transform (Rosenblum, Pikovsky, Kurths, Schäfer, & Tass, 2001), can be used to approximate its quadrature signal ($y = \mathscr{H}(x)$). The paired values of x and y constitute the complex-valued Analytic Signal and represent the coordinates of a point that traces a complete circle around the origin of the axes each time that the cyclic pattern of activity in the observed signal repeats itself. The phase of such a signal can be obtained as $\phi = tan^{-1}\left(\frac{y}{x}\right)$.

Principled applications of this definition of phase (henceforth referred to as the Hilbert phase) are limited to time series that can be approximated by a sinusoid with amplitude varying much slower than its phase (i.e. for $x(t) = A(t) \times cos(\Phi(t))$, $\dot{A}(t) \ll \dot{\Phi}(t)$, Rosenblum et al., 2001; Chavez, Besserve, Adam, & Martinerie, 2006; Huang et al., 2009). Failure to meet this requirement results in incorrect phase estimates, as exemplified in Fig. 1. Additionally, the signal should be locally represented by one frequency and by a spread around that frequency. In other words, the signal's phase is expected to change in a locally linear fashion: $\Phi(t) = \omega \times t + u(t)$ with $u(t)$ being a small phase modulation (Gengel & Pikovsky, 2019; Matsuki, Kori, & Kobayashi, 2023). Strategies to estimate a signal's phase that do not rely on the Hilbert transform consist in adopting definitions of phase that are specific to the signal under study (e.g.: Lancia, Chaminade, Nguyen, & Prevot, 2017; Mörtl, Lorenz, & Hirche, 2014), and that often focus on a few notable values characterising each cycle (e.g.: velocity peaks, zero-crossing points, points at which thresholds are crossed, etc., e.g.: Tuller & Kelso, 1984; Schäfer, Rosenblum, Kurths, & Abel, 1998; Kralemann et al., 2013).

The issues related to phase estimation are greatly reduced when this is conducted on signals submitted to band-pass filters or to decomposition techniques aimed at representing a signal via a combination of oscillatory patterns, because both methods return oscillations with slowly varying amplitude and centred on zero (cf. Chavez et al., 2006; Huang et al., 2009). For example, Huang et al. (2009) propose to submit the observed signal to Empirical Mode Decomposition (EMD). This operation produces a small number of oscillatory signals whose amplitude can be further processed so that in each cycle the curves oscillate between -1 and 1 and the quadrature signal can easily be computed. However, approaches based on band-pass filters or on decomposition schemes suffer from two kinds of problem. On the one hand, the selection of the filtering parameters, and the lack of stability of the decomposition algorithms, complicate the identification of the cycles of interest in the signal under study. On the other hand, the application of band-pass filters or of decomposition methods may wash out properties of phase dynamics that make the cycle shape deviate from that of a sinusoid.

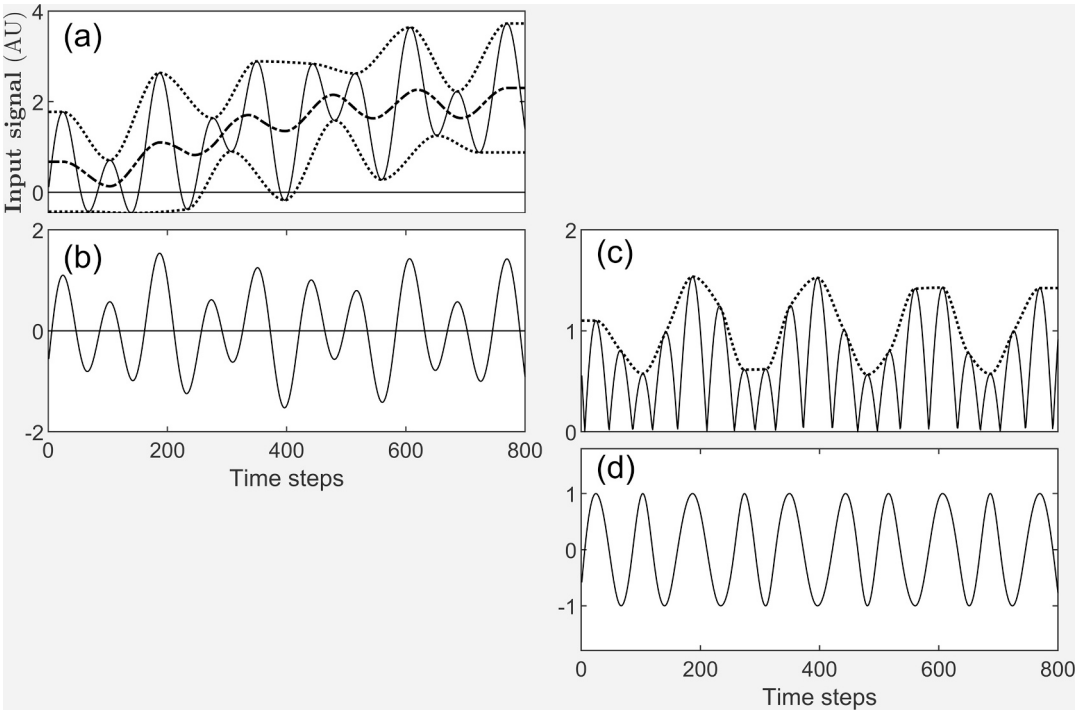


Fig. 2. Centring and demodulation operations. Panel (a): Input signal (continuous line), minimum and maximum envelopes (dotted lines) and mean envelope (dashed line). Panel (b): Output of one centring iteration. Panel (c): absolute value of the centred signal displayed in panel (b), and its maximum envelope. Panel (d): signal obtained by dividing the centred signal in panel (b) by the envelope in panel (c).

1.2. Analysing the phase of signals from goal-oriented behaviours

The work presented in the current paper is based on the observation that we can often identify a goal-oriented behaviour through the pattern of variation of one observable quantity or of a quantity that can be obtained by combining different observables. In experiments eliciting rhythmic behaviour; this usually results in signals that are dominated by one oscillatory pattern permitting to characterize the execution of the task under study.

Consider, for example, running or walking. We can characterize the cycles of activity by measuring the height of the heel from the ground. In speech production, we can compute from the observed acoustic signal a sonority coefficient, which oscillates at the rate of syllables production (Wang & Narayanan, 2007) or a spectral modulation coefficient, which oscillates at the rate of speech sounds' production (Goldstein, 2019). Although such signals often tend to display one main peak per cycle, their shape may deviate from that of a sinusoid (see examples in Sections 3.3 and 3.6 below) and they may contain events that are not the expression of the main oscillatory pattern (see examples in Sections 3.4 below). For example, in cases of complex rhythmic patterns involving several interconnected variables and representing the production of sequences of different actions (e.g. speech production, typing, dancing, etc.), we may observe disturbances in the behaviour of each variable, due to mutual dependencies with the other variables. Hierarchically organised rhythmic patterns (Goswami, 2019) and general non-stationarity of behaviour may produce slow changes that modulate the activity of the dominant oscillation (James, 2014; Melanson, Mejias, Jun, Maler, & Longtin, 2017). Finally, measurement noise may corrupt the observed signal. Still, if we exclude the effect of random noise or of weak and potentially sparse components, a signal collected during the execution of a controlled task can often be modelled as due to the activity of the oscillatory component of interest, potentially modulated by slower oscillatory patterns.

These considerations motivate the approach proposed in the present paper, in which, like Huang et al. (2009), we centre the observed cycles of activity and demodulate them. However, we do not need to submit the observed signal to a full decomposition or to narrow band-pass filters with predetermined coefficients because, once the effect of weak and potentially sparse components is excluded, the oscillatory pattern of interest is expected to be the fastest oscillation present in the signal.

The remainder of the paper is structured as follows: we will first summarise the approach proposed by Huang et al. (2009) to overcome the issues related to the construction of the Analytic Signal. After discussing the limitations of this approach and reviewing variants aimed at mitigating them, we will illustrate how the reviewed methods can be tailored to the analysis of goal-oriented rhythmic patterns. We will complement the proposed analysis pipeline with an original amplitude demodulation schema and we will discuss the integration of the adaptive filtering method proposed by Flandrin, Goncalves, & Rilling, 2004. This step permits to remove high frequency random components often present in the observed signal due to measurement noise. In the second part of the paper, we will use several synthetic and real-life example signals to illustrate the functioning of our algorithm. In the discussion

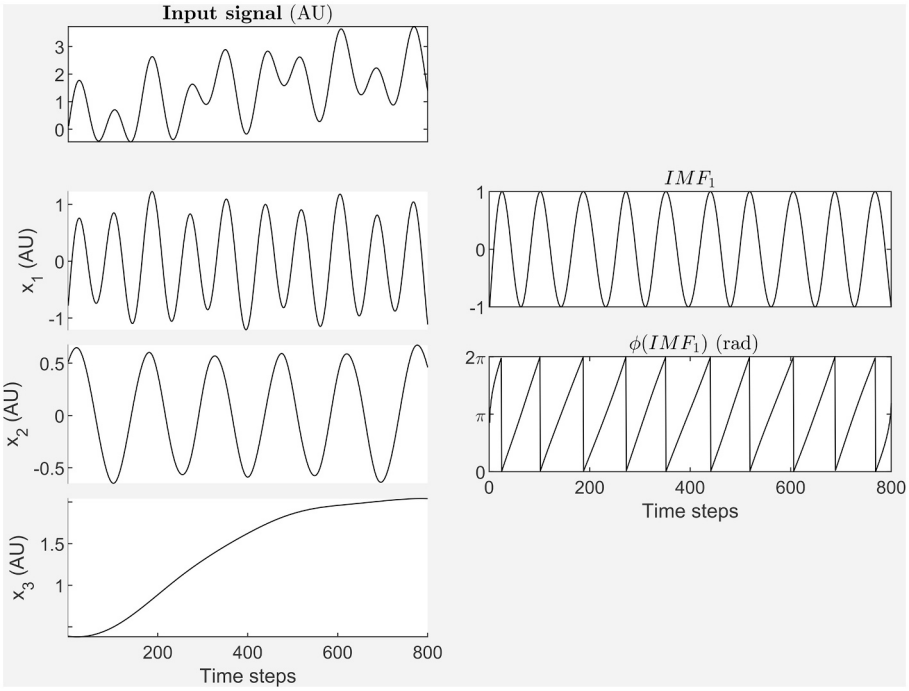


Fig. 3. EHHT analysis of the signal resulted by summing the signals in the left panels of Fig. 1, left column: observed signal (topmost panel) and IMFs from EMD (remaining panels). Right column: 1st IMF after demodulation (topmost panel) and phase from the application of the direct quadrature method (bottommost panel).

section, we will build on the results observed to comment on the features of the proposed methodology. Moreover, we will discuss the implications of our findings for decomposition procedures based on variants of EMD informed by knowledge of the signal's components. The detailed description of the computations performed in this work is included in the Appendix. The code (Matlab and Python) implementing the methods described in the current study and the datasets analysed are available in the following GitHub repository: https://github.com/LeonardoLancia/Get_Phase_Tools.

2. Methods

2.1. Huang et al., proposal

2.1.1. Empirical mode decomposition

The first step of the approach proposed by Huang et al. (2009) is to submit the input signal $x[n]$ (with $n \in \{1, \ldots, N\}$) to Empirical Mode Decomposition (EMD, Huang et al., 1998). This is an iterative procedure in which the input signal submitted to the k^{th} iteration is centred on its time-varying mean. The signal obtained ($d_k[n]$) is referred to as an Intrinsic Mode Function (IMF) and is considered to represent a component of the input signal. The residual signal r_k, obtained by subtracting d_k from the input signal, is used as the input for the next iteration (therefore $r[n]_k = d_{k-1}[n] - d_k[n]$).

The implementation of the centring operation is based on a further iterative process termed Sifting (henceforth referred to as $S_q(\bullet)$, with q indicating the maximum allowed number of iterations). At each iteration of Sifting, two envelopes ($e^+$and e^-) are obtained by interpolating separately via cubic splines the signal's local minima and maxima; then the mean of the two envelopes is subtracted from the input, and the output of each iteration is used as an input to the following one (see Fig. 2). The process is stopped before the maximum allowed number of iterations has been reached in which the signal's local maxima are all positive and its local minima are all negative, or if other more or less stringent criteria are met.

2.1.2. Demodulation and phase estimation

The IMF obtained from EMD that matches the frequency of the oscillation of interest is submitted to the following demodulation operation (henceforth denoted as $D(\bullet)$). Firstly, the local maxima of the signal's absolute values are identified and submitted to cubic Interpolation to produce an energy envelope $A[n]$. Secondly, the signal values are divided by the corresponding energy envelope values ($\widehat{x}[n] = x[n]/A[n]$, see right column panels in Fig. 2). The procedure is iterated, taking as input at each iteration the output of the previous iteration, until no extremum has an absolute value higher than $1 + \varepsilon$ (where ε represents a tolerance parameter) or until a generally small maximum number of iterations has been reached.

The obtained signal can be submitted to the Hilbert transform to obtain a quadrature signal and compute the phase. Hilbert phase,

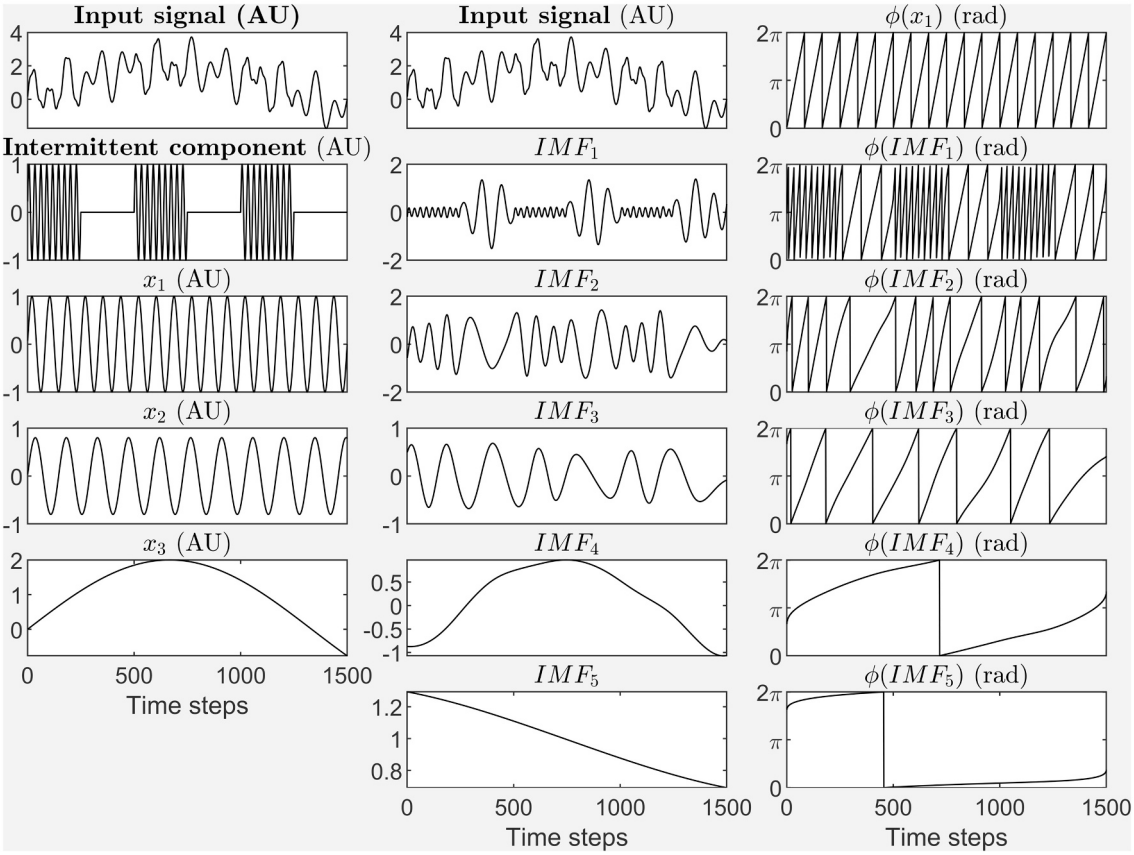


Fig. 4. EMD of a longer stretch of the signal in Fig. 3 (displayed in the topmost panels of the first and second columns) added to an intermittent oscillation (displayed in the second panel from the top of the leftmost column) multiplied by 0.2. The remaining panels of the leftmost column display the other components of the signal. The central column displays the same input signal in the topmost panel and the IMFs obtained by applying EMD in the remaining panels. The topmost panel of the third column displays the true phase of the second component of the input signal, while the following panels of the same column display the phase values extracted from each IMF after submitting it to the demodulation procedure proposed by Huang et al. (2009).

however, may become imprecise in the presence of modulation of the signal's phase (representing local deviations of the signal phase from a linearly changing phase), which tend to be smoothed out (Gengel & Pikovsky, 2019; Matsuki et al., 2023; Varlet & Richardson, 2011). In such cases, (and at the expense of numerical robustness) Huang et al. (2009) propose to obtain the quadrature signal by direct transformation of the demodulated signal as $[n] = -sgn(\partial_{\hat{x}}[n])\sqrt{1-\hat{x}[n]^2}$, where $\partial_{\hat{x}}$ indicates the estimated derivative of the demodulated signal. We will refer to this approach as the Enhanced Hilbert-Huang transform (EHHT, HHT being commonly used to denote the same approach without demodulation). The functioning of this approach is illustrated in Fig. 3, where it is applied to the signal resulting by summing the sinusoids in the left panel of Fig. 1.

2.1.3. Issues with the EHHT

Being based on extrema detection, EMD is sensitive to the presence of noise and intermittent behaviour, observed when an underlying oscillatory component is not constantly active. Furthermore, EMD biases the analysis to account for variability through changes in the amplitude of the obtained analytic signal. This choice is not motivated by theoretical considerations, but by the limitations of the Hilbert transform approach, which requires a signal with a very slowly changing amplitude. In EMD, such a signal is obtained by iterating the sifting operation, which reshapes its input again and again and gradually makes it similar to a sinusoid (Wang, Chen, Qiao, Wu, & Huang, 2010; Wu & Huang, 2010). Consequently, local accelerations and deceleration patterns can be gradually smoothed out.

Fig. 4 illustrates how the phase estimation procedure is affected by the addition of a weak, fast-frequency component with intermittent behaviour to the signal analysed in Fig. 3. The input signal (displayed in the topmost panels of both the leftmost and middle columns) has been obtained by scaling and adding the signal in the second panel from the top of the leftmost column to the signals displayed in the remaining panels of the same column. Although the intermittent component is barely visible in the input signal at the magnification level adopted (the scaling coefficient being 0.2), its impact on the decomposition procedure is dramatic. Due to its intermittent nature, the additional component characterizes part of the first IMF. That leaves the space for portions of the second

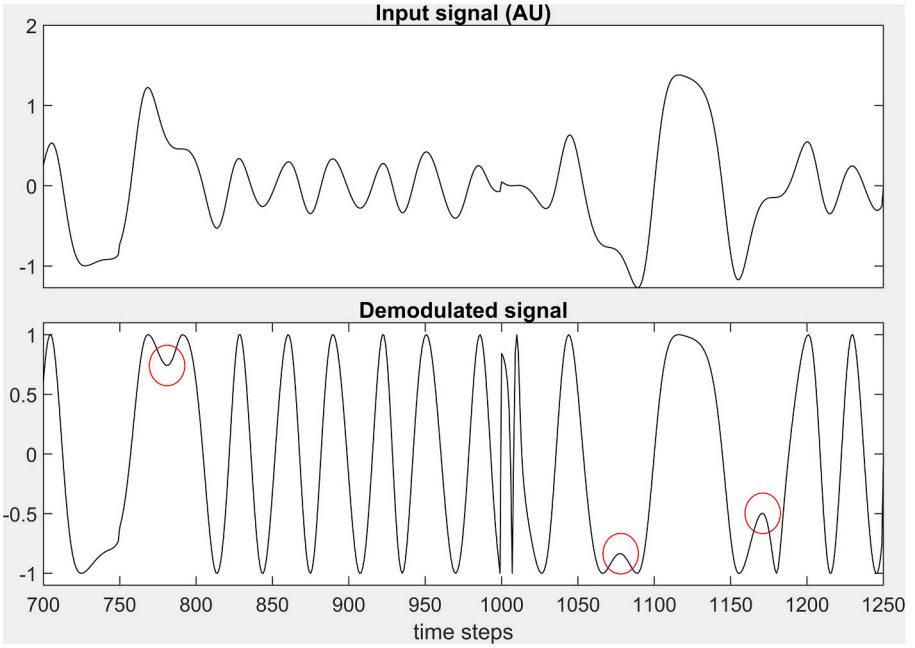


Fig. 5. Local extrema escaping the demodulation introduced by Huang et al. (2009). Top panel: The original signal. Bottom panel: The signal after demodulation. Circles indicate positive local minima and negative local maxima.

component to appear in the remaining parts of that same IMF (second panel from the top of the middle column). Other portions of the second component however characterize the second IMF. Consequently, the true phase of the second component of the input signal, which is the fastest visible component (and which is displayed in the topmost panel of the rightmost column), is captured by different IMFs during different time intervals (as displayed by the remaining panels of the same column).

A further limitation of the EHHT is related to the implementation of the demodulation schema based on cubic spline interpolation. According to Huang et al. (2009), due to the inherent smoothness of the adopted basis functions, rapid changes of amplitude may not be captured by the obtained amplitude envelope. This produces negative local maxima and positive local minima whose presence may introduce phase jumps and/or intervals characterised by negative frequencies (see Fig. 5).

A Final limitation of the EHHT approach concerns the method adopted to avoid the smoothing of the phase modulation, and the consequent loss of rapid phase changes due to the application of the Hilbert transform (Gengel & Pikovsky, 2019; Matsuki et al., 2023). While the direct quadrature method adopted by Huang et al. (2009) permits addressing this limitation (see Section 2.1.2), it lacks robustness around time points where the signal crosses 0 (Sandoval & De Leon, 2017), making it not ideal for the analysis of real-life signals.

2.2. *Masked sifting, masked EMD and iterated masking EMD*

The EHHT approach relies on the performance of EMD and more specifically of the Sifting operator, permitting to separate the observed signal's components. Several studies propose variants of this operator aimed at improving its performances, especially in the presence of intermittent behaviour. Most methods add to the signal submitted to Sifting a disturbance signal that can take the shape either of withe noise (e.g.: Colominas, Schlotthauer, & Torres, 2014; Wu & Huang, 2009) or of a sinusoid (e.g.: Deering & Kaiser, 2005; Fabus, Quinn, Warnaby, & Woolrich, 2021; Fosso & Molinas, 2018; Wang, Hu, & Lo, 2018). An advantage of methods based Sifting perturbed by sinusoidal signals (and generally referred to as **masked Sifting**) is due to the fact that the frequency of the sinusoid added to the input signal acts as a lower bound on the frequency of the signal extracted by the sifting process, allowing explicit control on the frequency selectivity of the Sifting operation.

In its more general definition, masked Sifting (henceforth $MS_s(r, \omega_{mask}, A_{mask}, J)$) takes as input a signal $x[n]$ containing N values collected with sampling period Δt, a frequency value ω_{mask}, an amplitude value A_{mask} and a positive integer J. The index s represents the maximum allowed number of iterations in the sifting process. First, J phase-shifted copies of a masking sinusoidal signal are created:

$$mask_j = A_{mask} \times cos\left(n \times \Delta t \times \omega_{mask} + \phi_{0,j}\right), with\ j \in \{1, \ldots, J\}, and\ n \in \{1, \ldots, N\}.$$

The obtained J signals will have the same amplitude A_{mask} and frequency ω_{mask} but different initial phases $\Phi_{0,j}$. The input signal is independently summed with each $mask_j$ and submitted to the Sifting operator. Partial IMFs are obtained by subtracting the masking signals from the results of the Sifting operations ($\breve{x}_j = S_q\left(mask_j + x\right) - mask_j, \forall j \in \{1, \ldots, N\}$). Finally, the partial IMFs are averaged to produce an independent mode function ($\breve{x} = \sum_{j=1}^{J} \breve{x}_j$).

The presence of the masking signals prevents the sifting operation from capturing oscillations with frequencies smaller than ω_{mask} and thus effectively removes the effects of low-frequency components. The amplitude of the masking signal A_{mask} is set at a fixed proportion of the signal's standard deviation ($A_{mask} = C_a\sigma(x)$, with $\sigma(x)$ indicating the signal's standard deviation).

A critical step in the application of EMD based on masked Sifting (henceforth **masked EMD**) is the selection of the masking signals' frequencies. Deering and Kaiser (2005) propose to extract a temporary first IMF via the application of classical EMD and to average its amplitude-weighted instantaneous frequencies to obtain the masking frequency to use in the extraction of the first IMF via masked Sifting. The masking frequencies used in the computation of the following IMFs should be obtained by dividing the first masking frequency by increasing powers of two. In the approach proposed by Fabus et al. (2021), named **iterated masking EMD** (henceforth itEMD), the masking frequencies are obtained by including the masked EMD algorithm in a further iterative process. At each iteration, the masking frequencies are estimated from the IMFs extracted at the previous iteration (as amplitude-weighted averages) and used to extract new modes via masked EMD. The procedure is stopped when the differences between the average frequencies obtained from consecutive iterations get smaller than a predefined threshold. itEMD provides an effective mean to decompose stationary signals containing intermittent components and it has the advantage of being robust to the initial choice of the masking frequencies. However behavioural data, except when coming from very strictly controlled experimental conditions (e.g. as when driven by external pacemakers), generally lack stationarity. This constitutes an issue because masking frequencies are not allowed to vary over time and because their estimation may become unstable for short, noisy time series. Importantly, no variant of EMD guarantees that the oscillation of interest is captured by the same IMF when analysing separately different recordings of the analysed behaviour (or different chunks of a long signal in order to deal with stationarity). Therefore, given the decompositions of several signals representing different recordings of the same process at different points in time, a further analysis step is required in order to individuate the main component among the IMFs extracted by each signal.

2.3. Computation of phase

In order to increase the robustness of the direct quadrature method (where the quadrature signal is given by $y[n] = -sgn(\partial_{\hat{x}}[n])\sqrt{1-\hat{x}[n]^2}$), Sandoval and De Leon (2017) propose to substitute the values of $y[n]$ close to zero and those in the time intervals containing these values with interpolated values. This approach breaks down however if the cycles of the demodulated signal still show secondary peaks. These cases can be handled through the method proposed by Kralemann, Cimponeriu, Rosenblum, Pikovsky, and Mrowka (2008). Instead of computing the angle of the analytic signal, these authors proceed in several steps. The first is the computation of the cumulative curve-length of the trajectory drawn by the point on the plane at each time step:

$$L[n] = \sum_{n=1}^{N} \sqrt{(\delta(\hat{x}[n]))^2 + (\delta(y[n]))^2},$$

where $\delta(\hat{x}[n])$ indicates the derivative of an amplitude demodulated signal $\hat{x}$ (or more generally of a signal with constant amplitude) and $\delta(y[n])$ indicates the derivative of its quadrature. The cycles' onsets (n_i^*, with $i \in \{1, \ldots, I\}$ and I being the number of cycles) are then located by looking at the intersection between the trajectory and an appropriate Poincaré section (e.g. a line passing from the origin). Finally, given two consecutive cycles' onsets located at time steps n_i^* and n_{i+1}^*, the phase is computed at each time step n occurring between them according to:

$$\Phi[n] = 2 \times \pi \times \frac{L[n] - L[n_i^*]}{L[n_{i+1}^*] - L[n_i^*]} + 2 \times \pi \times i$$

In the original implementation of this approach, the analytic signal was built via the Hilbert transform, a feature that introduces a degree of phase modulation smoothing. The variant proposed by Gengel and Pikovsky (2019) gradually removes the smoothing through the iterative application of the phase computation step. At the first iteration, phase is computed either as $\phi = tan^{-1}\left(\frac{y}{x}\right)$ or based on the computation of the curve length. At each following iteration, the phase computation is performed on an analytic signal indexed over time by the phase values obtained at the previous iteration.

2.4. Our proposal

As mentioned in section 1.2, a key assumption of our approach is that the oscillatory pattern of interest can be captured by centring the observed signal without running a full decomposition. To this aim we use masked Sifting. However, unlike the original Sifting algorithm designed as a high-pass filter, our modified version centers the main oscillatory component. Two key modifications achieve this behaviour. First, in contrast to masked EMD and ITEMD, we estimate the masking frequency using only one Sifting iteration. This ensures we capture only the most significant fluctuations. Second, similar to EHHT (and unlike masked EMD), the centred signal undergoes amplitude demodulation before instantaneous frequency estimation. However, we employ an improved demodulation strategy that eliminates the possibility of negative local maxima and positive local minima. Estimating the masking frequency from these demodulated signals allows us to remove the influence of sparse or weak components, guaranteeing that the masked Sifting output retains the relevant oscillation. Finally, the obtained signal is demodulated again using the modified procedure and used to construct an analytic signal for phase calculation (see Table 1 for a summary of the processing steps).

Table 1
Processing steps of the two parts of our proposed approach. Leftmost and rightmost columns display the formulas implementing the operations described in the centre column. The slash separates operations applied at corresponding steps in the two parts. In the second part, the centred signal is obtained via masked sifting based on the average instantaneous frequency obtained as output of the first part of the algorithm.

PART I	Operation executed	PART II
$\widetilde{x} = S_1(x)$	Centring / masked sifting	$\dot{x} = MS_s\left(x, \overline{\omega}_{\dot{x}}, A_{mask}, J\right)$
$\widehat{x} = \varkappa(\widetilde{x})$	Demodulation	$\widehat{\dot{x}} = \varkappa(\dot{x})$
$y^{(0)} = \mathscr{H}(\widehat{x})$	Hilbert transform	$y = \mathscr{H}(\widehat{\dot{x}})$
$\overline{\omega}_{\dot{x}} = \langle SGd(tan^{-1}(y^{(0)}/\widehat{x}), o_{sg}, l_{sg}) \rangle$	Frequency / phase estimate	$\phi(x) = tan^{-1}(y/\widehat{\dot{x}})$

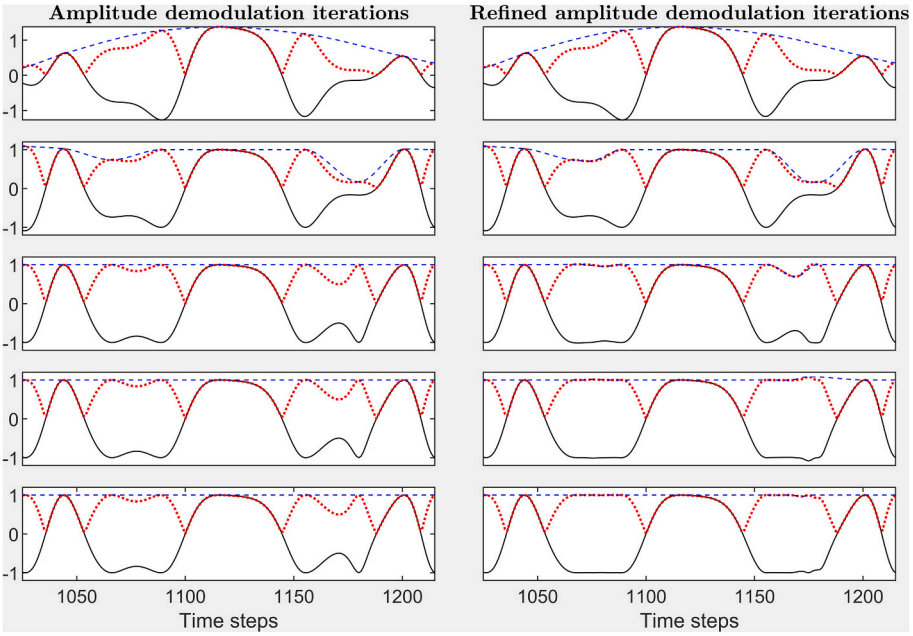


Fig. 6. Five iterations of the demodulation based on the local maxima of the absolute values of the input signal (left column panels) and based on the local extrema of the input signal (right column panels). Each panel displays the input signal (continuous black line), its absolute values (dotted red line), and the envelope obtained by interpolating the extrema identified (dashed blue line). Signals from the first iteration are displayed in the topmost panels. With both approaches, one local maximum smaller than 0 emerges at the second iteration around time step 1070, and another at the third iteration around time step 1180. These extrema do not affect the envelopes displayed in the left column panels because the points do not correspond to local maxima of the absolute values signal. However, the peaks are captured by the envelopes in the right column panels, which are based on the locations of all the extrema of the input signal. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4.1. First part I: Centring

The first part of the algorithm is aimed at estimating the dominant oscillatory frequency of the input signal $x[n]$ (with $n \in \{1, \ldots, N\}$). The first step consists in centring $x[n]$, which is obtained through one iteration of the Sifting process ($\widetilde{x} = S_1(x)$). This early stopping is expected to reduce the effect of weak high-frequency component, which are less likely to be captured at the first iteration.

2.4.2. First part II: Improved demodulation

In the EHHT, some extrema escape demodulation because they are extracted from the signal's absolute values. Indeed, the absolute value of a signal containing positive local minima or negative local maxima will not display any local maximum at these locations (see Fig. 6 caption and panels in the leftmost column). The envelope obtained does not reflect the presence of these local extrema, which therefore are not affected when the input signal is divided by the envelope. This issue can be easily avoided by adopting a refined demodulation procedure (henceforth referred to as $\varkappa(\bullet)$) in which we select local minima and local maxima separately in the input signal, we compute their absolute values, and we obtain the signal's envelope via interpolation of these values (see the right column panels of Fig. 6).

2.4.3. First part III: Frequency estimation

We first compute $\Phi[n]$ through application of the Hilbert transform ($\Phi_{\check{x}}[n] = tan^{-1}\left(\frac{\mathcal{H}(\widehat{x}[n])}{\widehat{x}[n]}\right)$). The use of the Hilbert transform to obtain the quadrature signal required for the computation of $\Phi_{\check{x}}[n]$ is motivated by its numerical stability and by the fact that the smoothing out fast phase modulation does not affect frequency estimation (Matsuki et al., 2023). The instantaneoys phase is unwrapped, so as to increase monotonically over time, and $\omega_{\check{x}}[n]$ is estimated via the application of a Savitzky-Golay differentiator ($\omega_{\check{x}}[n] = SGd(\Phi[n], o_{sg}, l_{sg})$, `with order: $o_{sg} = 5$, and length: $l_{sg} = 16$ time steps).

2.4.4. Second part I: Masked sifting

In the second part of our algorithm, we use the frequency $\overline{\omega}_{\check{x}} = \langle \omega_{\check{x}}[n] \rangle$ (with hooks representing averaging) to process the signal $x[n]$ via masked sifting (Deering & Kaiser, 2005; Fosso & Molinas, 2018).

2.4.5. Second part II: Demodulation, quadrature signal and phase estimation

The oscillatory pattern obtained by application of the masked sifting operator to the input signal ($\check{x} = MS_s\left(x, \overline{\omega}_{\check{x}}, A_{mask}, J\right)$), is submitted to refined demodulation ($\widehat{x} = \varkappa(\check{x})$) and the demodulated signal is used to compute the quadrature signal and the instantaneous phase. The approach to be adopted for the computation of the quadrature signal should probably depend on the application. Hilbert phase smooths out fast modulations. Smoother phase signals may be desirable when analysing complex or irregular vibratory patterns (as those in analysed in Sections 3.4–3.6 below). Using the direct quadrature approach to build the analytic signal and the curve length method (Kralemann et al., 2008) to compute the phase permits better capturing fast phase modulations. With this strategy, the adoption of the Hilbert transform is completely avoided, so that no phase modulation smoothing is expected. Importantly, phase estimates based on the curve length of the trajectory drawn by the obtained analytic signal (Kralemann et al., 2008) are marginally affected by numerical instabilities due to the computation of the quadrature signal.

2.5. Processing noisy signals via masked EMD filtering

The approach described so far deals with smooth signals, such as those obtained by low-pass filtering the data. In order to deal with signals containing high-frequency noise, we propose an adaptive filtering step based on the one introduced by (Flandrin et al., 2004). Originally, the approach is based on the application of EMD to the input signal and on the identification of the modes dominated by random behaviour. The denoised signal is then obtained by summing the remaining IMFs. In our implementation, we substitute EMD with masked EMD, which we modified to benefit from the frequency estimation procedure proposed in sections 2.4.1–2.4.3.

2.5.1. Masked EMD

At iteration k of the masked EMD procedure adopted in our work, the input signal (x), or the residual of the previous iteration (r_{k-1}), is submitted to centring ($S_q(\bullet)$) and the results is amplitude demodulated (through the application of $\varkappa(\bullet)$). A frequency value is computed based on the application of Hilbert transform and used to drive masked Sifting ($MS_s(\bullet)$), as illustrated in Section 2.2. The output of masked sifting at the k^{th} iteration is stored as an IMF ($d_k = MS_s(r_{k-1})$) and the residual ($r_k = r_{k-1} - d_k$) is submitted as input to the next iteration. The process is stopped when the residual obtained does not display local extrema or when a predetermined number of modes has been extracted.

The amplitude of the masking signal A_{mask} is set at a fixed proportion of the first mode's standard deviation ($A_{mask} = C_\alpha \sigma(d_1)$), or, at the first iteration, of the centred input signal. The choice of taking the first estimated mode as a reference for the amplitude of the masking signal is aimed at making this amplitude comparable to that of the signal's random content, which by hypothesis changes faster than its deterministic content. If this is the case, the first extracted IMF is expected to contain mainly noise (Flandrin et al., 2004).

2.5.2. EMD based denoising

The denoising procedure starts by estimating the Hurst exponent H_1 of the first IMF obtained from the application of masked EMD (which should reflect the features of the signal's random content, under the hypothesis that the random content of the signal changes faster than its deterministic content (Flandrin et al., 2004)). Then, for each mode d_k obtained from the application of EMD to the observed signal, we test the hypothesis that its energy is larger than that of the mode $d_k^{H_1}$ obtained from the application of EMD to a random signal with the Hurst exponent H_1. If the null hypothesis that the energies of the two modes are not different cannot be excluded, the mode is flagged as random. To this aim, the energy of each IMF d_k, is compared to the energy values of the corresponding IMFs obtained by submitting r random signals with Hurst exponent equal to H_1 to EMD. The clean signal is finally obtained by summing the modes not flagged as random.

2.6. Parametrization

The following parameters were adopted in the analyses reported in Section 3. Demodulation tolerance: $\varepsilon = e^{-10}$; Savitsky Golay differntiator order: $o_{sg} = 5$ and length: $l_{sg} = 16$; amplitude coefficient of the masking signal: $C_A = 8$ (which, under the hypothesis that the signal is normally distributed should roughly amount to twice the signal's range); number of masking signals: $J = 22$; number of random signals generated to estimate the confidence intervals in the denoising algorithm: $r = 1000$ (but results did not changed with

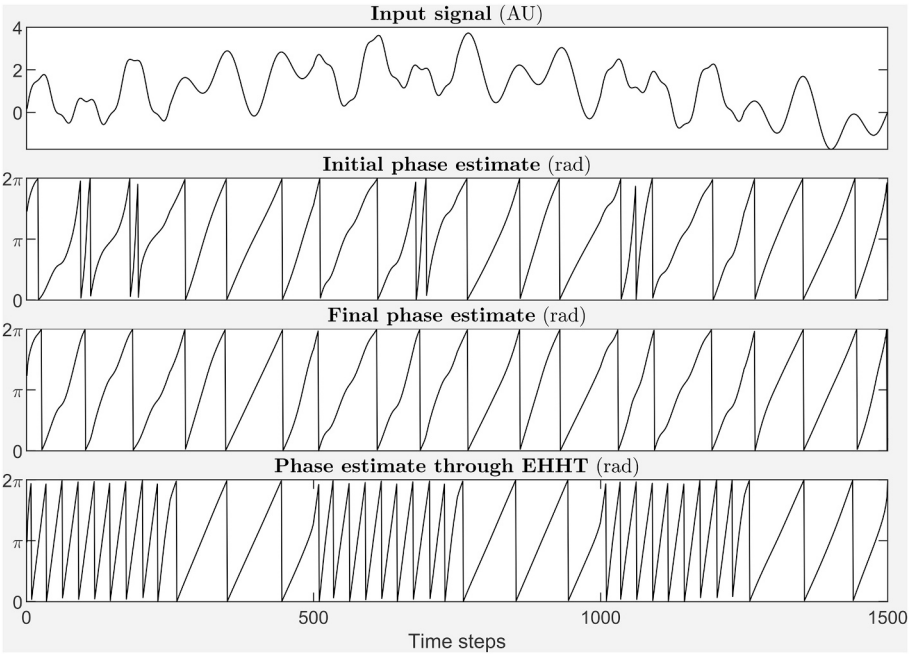


Fig. 7. Application of our approach to the analysis of the signal in Fig. 4 (here displayed in the topmost panel). Second panel from top: initial estimate of the signal phase after application of centring and refined demodulation. Third panel from top: final estimate of the signal's phase. Bottommost panel: phase estimated through EHHT.

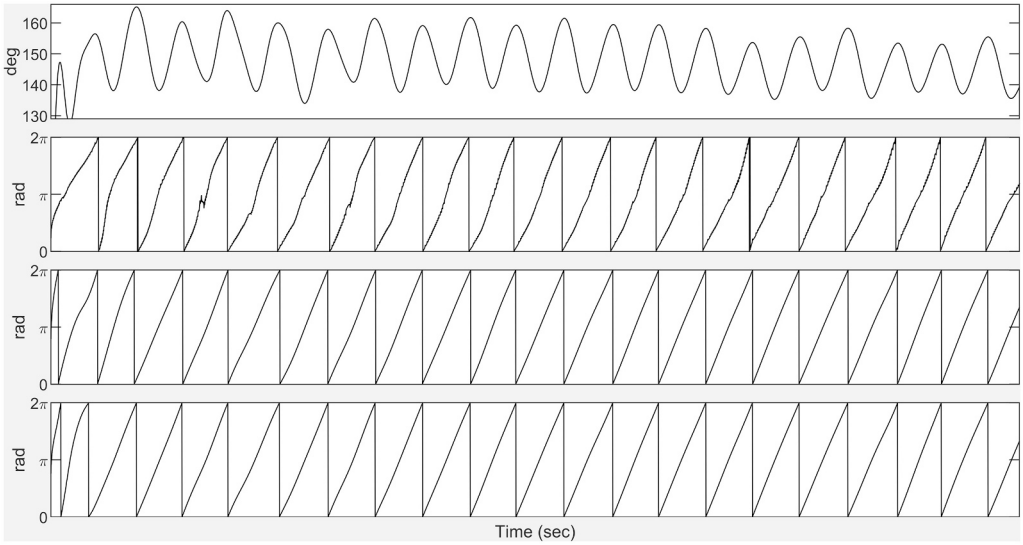


Fig. 8. Knee angle data observed during a flexion / extension task. The observed time series is displayed in the topmost panel while the remaining panels display its phase values computed through application of z-score and Hilbert transforms (second panel from top), through the EHHT (third panel from top), and through the method proposed in this paper (bottommost panel).

$r = 100$); number of masked sifting iterations: $q = 10$ (as suggested in by Wang et al., 2010). This choice was adopted also for the sifting operation in applications of the EHHT. Phase values displayed in the figures are Hilbert phase values. This choice was motivated by the aim of focusing our analyses on the extraction of the main oscillatory components of the signals under study. In Appendix 7.6 we show how fast phase modulation can be captured by obtaining the quadrature signal via the direct quadrature method and by computing the phase via the curve length method.

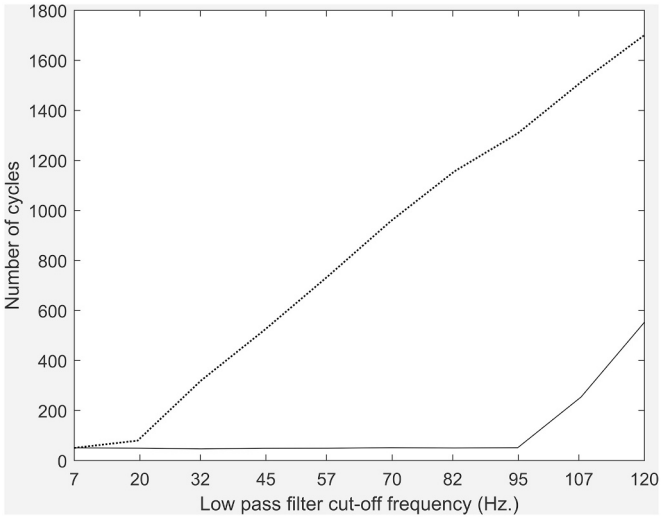


Fig. 9. Estimates of number of cycles based on the instantaneous phase values as a function of the cut off frequency of the low pass filter applied to the signal during pre-processing. The continuous line displays the results obtained through the application of our algorithm. The dotted line displays the results obtained through the adoption of EHHT. Both methods identify the correct cycle number when the cut-of freq. is set at 7 Hz.

3. Results

3.1. Synthetic data

Fig. 7 displays the phase obtained by the application of our algorithm to the signal analysed in Fig. 4. The topmost panel displays the analysed signal, the second panel from the top displays the first estimation of the signal phase via the Hilbert transform approach after centring ($S_1(\bullet)$), and refined demodulation ($\varkappa(\bullet)$). The third panel from the top displays the output of our algorithm while the bottommost panel displays the phase obtained through the EHHT (based on several iterations of the centring operation) by selecting the first IMF and submitting it to demodulation ($D(\bullet)$).

Only the final output of our algorithm permits recovering the oscillatory cycles visible in the signal. However, the comparison between the second and last panels from the top shows that limiting the sifting operator to one iteration of the centring algorithm reduces the sensitivity of the analysis to the small perturbations induced by the weak intermittent component.

3.2. Knee flexion data

Fig. 8 displays the analysis of a signal representing knee joint angles recorded while participants in standing position had to move rhythmically their torso up and down by flexing and extending their knees and keeping their feet still (Miyata & Kudo, 2014). The observed cycles, low-pass filtered at 7 Hz, have a quite sinusoidal shape. Moreover, the changes in cycles' amplitudes are relatively slow when compared to their durations. As shown in the second panel from top even the phase obtained by applying the Hilbert transform to the z-scored signal permits the correct identification of all the cycles. However, this signal presents a potentially challenging feature because its values appear to be quantised; therefore, it contains discontinuities all over its duration (not visible in the figure due to insufficient magnification level). Methods based on extrema detection (as those discussed so far) are expected to break down in presence of such discontinuities. Although this issue can be addressed by filtering the signal (as done in computing the phase values displayed in Fig. 8), the presence of the discontinuities permits appreciating how sensitive the different methods are to the presence of small non-random fluctuations. Indeed, we can gradually increase the cut-off frequency of the low-pass filter applied to the knee angle signal, count the number of cycles detected in the phase signals obtained through the different approaches, and observe how this number depends on the filter cut-off frequency. We can estimate the number of produced cycles via: $\frac{\sum_{n=1}^{N-1}\phi^u[n+1]-\phi^u[n]}{2\pi}$, where N is the length of the time series and $\phi^u[n]$ are the unwrapped phase values.

The cycle counts displayed in Fig. 9 were obtained by using 10 integer valued cut-off frequencies uniformly distributed over the interval [7, 120]Hz. While our approach stably approximates the correct number of cycles over a wide range of cut-off frequencies, the approach based on the EHHT starts deviating from the correct behaviour already at the first change of the cut-off frequency value.

3.3. Low pass filtered foot position during gait

The topmost panel on the left of Fig. 10 represents the vertical movement of a sensor attached on the Achille's tendon during walking. The signals were made available by the authors of the original publication (Hebenstreit et al., 2015) low pass-filtered with cut-off frequency at 6 Hz. The second panel from the top in the same column shows the phase obtained by submitting the z-scored time

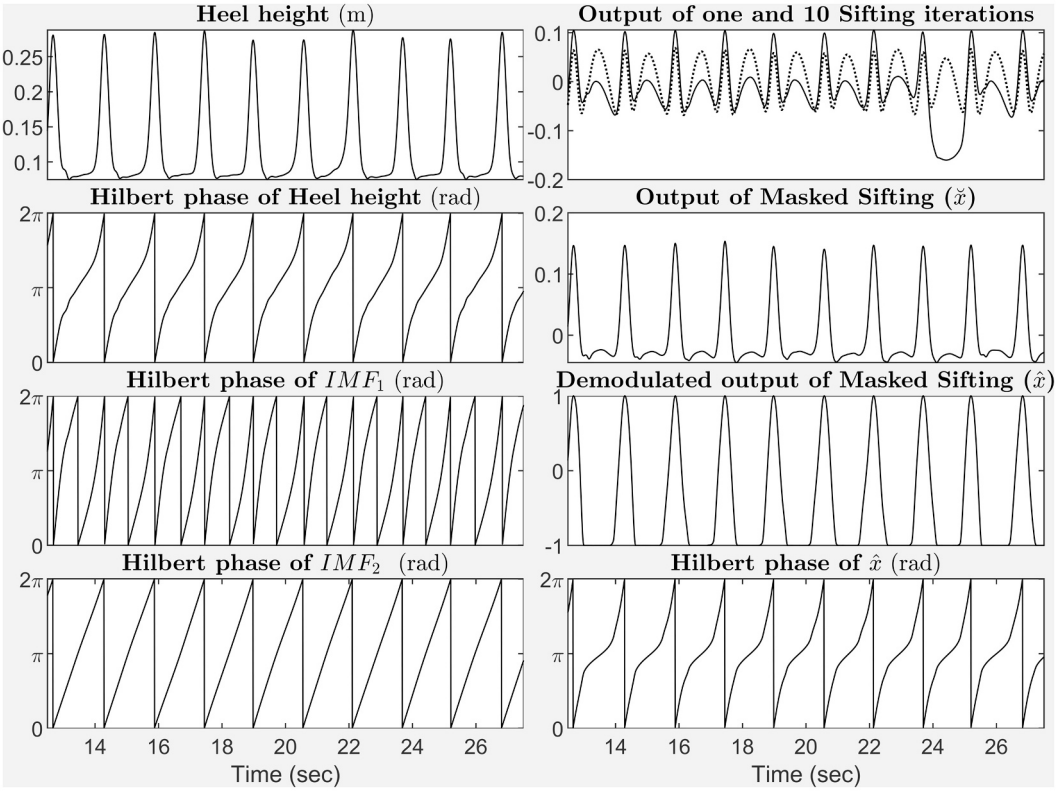


Fig. 10. Processing of foot elevation data recorded during gait. Topmost panel of left column: input signal. Remaining panels from top to bottom: phase of the z-scored input signal, of the first IMF from application of EMD, of the second IMF. Topmost panel of the right column: input signal centered through one iteration of classic Sifting (continuous line) and through ten iterations (dotted line). Remaining panels from top to bottom: output of masked Sifting; output of refined amplitude demodulation; Instantaneous phase of demodulated signal.

series to the Hilbert transform. Note that this approach permits capturing the slowing down of phase dynamics observed at low heights. The third and fourth panels from the top display the phase of the first and second IMFs obtained through EMD (and submitted to demodulation according to the EHHT schema). The second IMF captures the cycles, but information about the phase dynamics is lost. The topmost panel of the right column displays the results of classical sifting with one iteration (continuous line) and 10 iterations (dotted line). Even with a single iteration, spurious cycles emerge during the slower portions of the input signal due to the significant variation in the trajectory's rate of change. Note that the amplitude of these spurious peaks increases as the number of classical sifting iterations is increased. The second panel from the top displays the result of masked sifting, which still exhibits weak spurious cycles. However, this processing step further weakens the relative amplitude of these weak cycles. Therefore, the application of improved amplitude demodulation can eliminate them, allowing for the recovery of the phase dynamics.

3.4. Tip of the tongue, lower lip and jaw motion in repeated speech production

The signals in the panels on the left side of Fig. 11 represent the vertical movement of sensors attached to the tip of the tongue (TTIP, topmost panel), of the jaw (JAW, middle panel) and of the lower lip (LLIP, bottommost panel) collected by Rochet-Capellan and Schwartz (2007) via 2D electromagnetic articulography during the repetition of the utterance 'tapa' (where the vowel is pronounced like in the word 'staff'). Speakers were asked to repeat the two syllables without interruption during 10 s simultaneously to the cues provided by a visual metronome. The metronome frequency increased during the first half of the trial and then gradually came back to its initial value during the second half. Before the analysis, signals were low-pass filtered with a cut-off freq. of 20 Hz to remove measurement noise.

The panels on the right side of the figure display one slow repetition of the utterance (first three panels from top) and one fast repetition (last three panels from top). At slow speech rate, we can clearly see one main cycle of tongue motion (to produce the consonant 't') followed by a much smaller one, and one main cycle of vertical lower lip motion (to produce the 'p') preceded by a slightly smaller one. This pattern is observed because while each articulator organ is activated to produce either 't' or 'p', its movement shows the trace of the other articulator producing the other consonant. This occurs because both articulators sit on the jaw, which rotates to favour the upward movement of each articulator and, in doing that, transmits the movement of one articulator to the other. This also explains why the jaw produces two cycles per utterance. Still, the movement of the tip of the tongue seems to have a larger

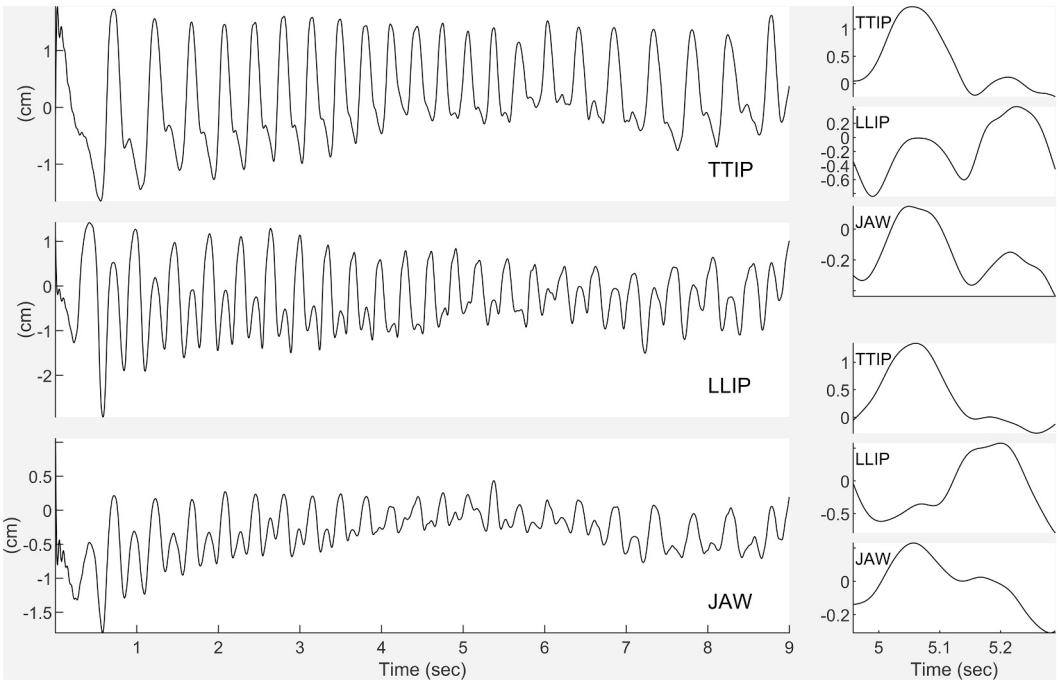


Fig. 11. Articulator vertical motion during the repeated production of 'tapa' with increasing-decreasing speech rate. The panels on the left display the vertical movement of the tip of the tongue sensor (TTIP: topmost panel), of the lower lip sensor (LLIP: middle panel) and of the jaw sensor (JAW: bottommost panel). The panels of the right column display the trajectories of these three articulators during the production of an utterance at slow speech rate (first three panels from top) and at fast speech rate (last three panels from top).

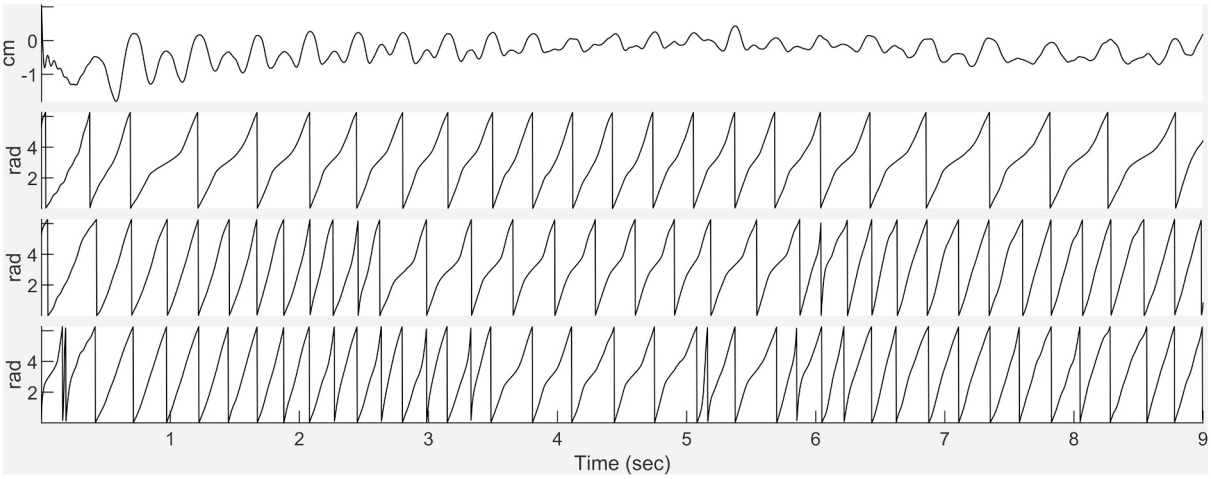


Fig. 12. Phase of articulator motion recorded during the repeated production of 'tapa'. Topmost panel: vertical movement of the jaw. Middle panel: tip of the tongue's phase. Third panel from top: lower lip's phase. Bottommost panel: jaw's phase.

impact on the movement of the other articulators than the cycle of vertical lip motion (Rochet-Capellan & Schwartz, 2007). At fast speech rate, this difference becomes even larger and the effect of lower lip displacement on the movement of the jaw appears as a perturbation of a larger cycle corresponding to the upward movement of the tip of the tongue. This is consistent with the observation that in this task at fast speech rate the coupling between tip of the tongue and jaw becomes stronger than the coupling between the lower lip and the jaw (Lancia & Rosenbaum, 2018).

The signals displayed in Fig. 12 show how the phase values computed through our approach reflect this situation. The topmost panel of the figure displays vertical jaw motion to provide a reference. The phase of the tip of the tongue (second panel from top) captures one cycle per utterance practically all along the sequence. The lower lip phase (third panel from top) displays one cycle per utterance only at fast speech rate. Finally, the jaw (bottommost panel) produces two cycles per utterances during longer portions of the

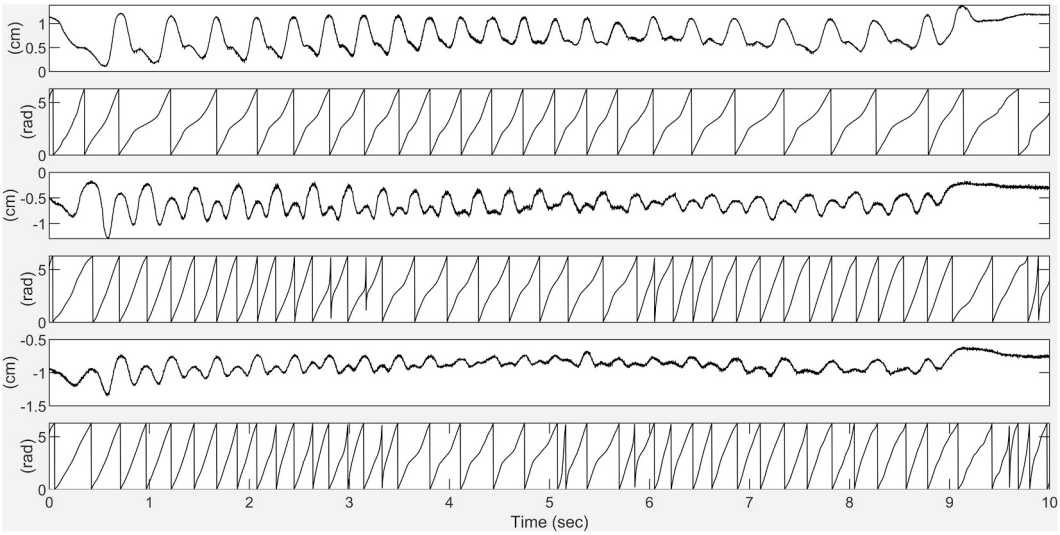


Fig. 13. Processing of speech articulator unfiltered data. Odd panels display from top to bottom the upward movements of the jaw, the lower lip and the tip of the tongue. Even panels display the phase angles extracted (see main text for details).

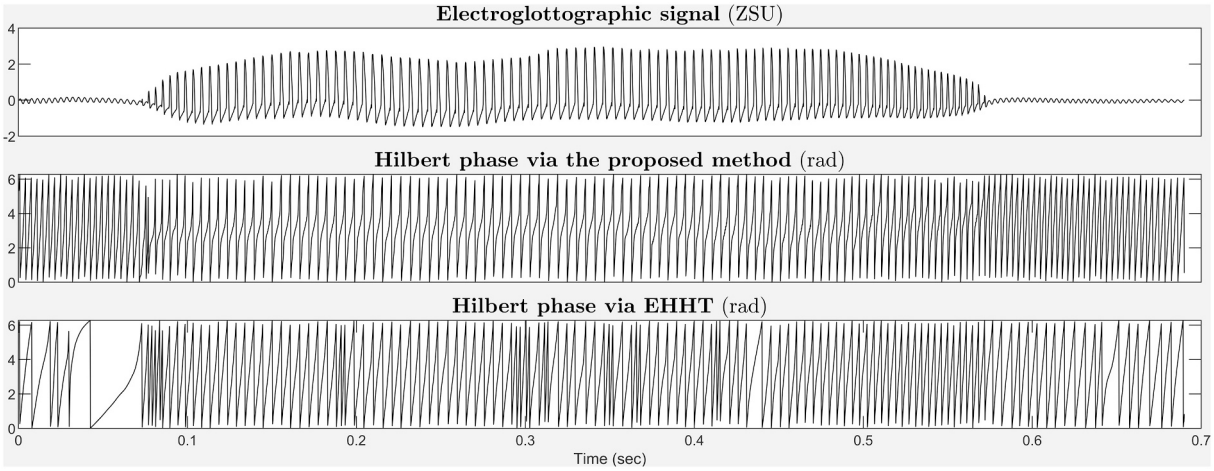


Fig. 14. Vocal fold vibration cycles. Topmost panel: estimate of vocal folds' contact area as captured through an electroglottograph (ZSU: z-scored units). Second panel from top: Phase of the vocal fold vibration cycles obtained through our algorithm. Third panel from top: Phase obtained through the application of the EHHT by selecting the second IMF.

sequence with respect to the lower lip, but as the lower lip, at fast speech rate it displays one cycle per utterance. Note also that at fast speech rate the tip of the tongue and the lower lip are in anti-phase, as it is expected due the alternation of the two consonants. When all articulators display one cycle per utterance, the jaw is in phase with the tip of the tongue. This is expected if the coupling between these two articulator organs is stronger than that between the jaw and the lower lip, as proposed by (Lancia & Rosenbaum, 2018).

3.5. Unfiltered speech articulators data

The odd panels of Fig. 13 starting from top display the articulator movements analysed in Section 3.4 before the application of the low-pass filter. The even panels display the phase obtained after the application of the denoising algorithm illustrated in Section 2.50. The results obtained by substituting the low pass filter with the adaptive filtering algorithm closely match those described in Section 3.4.

3.6. Vocal folds vibration cycles

The signal in Fig. 14 represents the vibration of the vocal folds during the production of the phrase "ma mine" from a female French

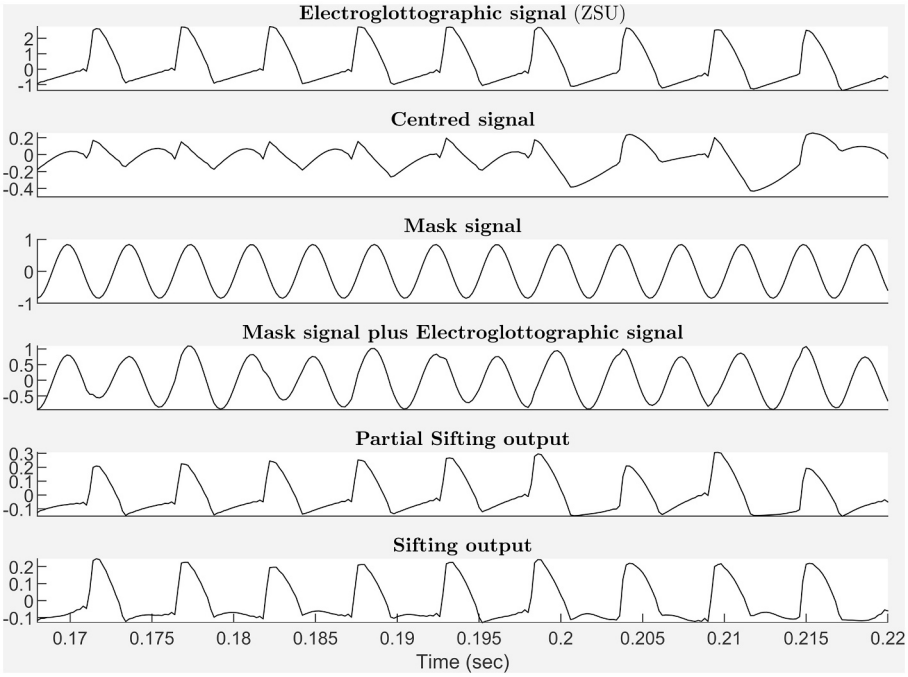


Fig. 15. Signals produced in the analysis of a portion of the vocal folds vibration signal in Fig. 14. Topmost signal: Vocal folds' vibration. Second signal from top: vocal folds' vibration centred on zero. Third signal from top: masking signal with $\phi_0 = 0$. Fourth signal from top: vocal folds' vibration plus masking signal. Fifth signal from top: partial output of the masked sifting process after removal of the masking signal. Bottommost signal: final output of masked sifting.

speaker. More precisely, the signal represents a measure of the contact area between the vocal folds, recorded through an electroglottograph. The shape of the opening and closing cycles, characterising the production of voiced sounds such as the vowels and the nasal consonants composing our signal, displays substantial deviation from the shape of a sinusoid (see also the topmost panel of Fig. 15). This feature constitutes a problem for the construction of an analytic signal without the application of a decomposition method. Moreover, contrarily to the signals analysed so far, speakers cannot control the instant-by-instant behaviour of glottal opening (although they have some degree of control over the global shape of the cycles, which is associated to different voice qualities).

Still, the signal displays one main peak per cycle and our algorithm is able to capture the main component of the vocal folds vibration cycles, as seen in the second panel from the top in Fig. 14. The result obtained through the EHHT are much less robust due to mode splitting. Note, that even using variants of EMD (e.g.: Colominas et al., 2014) that with this kind of signals prevent the occurrence of mixing/splitting phenomena, would not solve the problem of selecting the correct IMF, given that this may change from one chunk of signal to the other.

Fig. 15 shows the signals produced in the process of extracting the phase through our algorithm. The topmost panel displays a portion of the signal in Fig. 14. The second panel from the top shows the centred signal, featuring two kinds of cycles: a smoother one, reflecting the slow varying portions of the vocal fold vibration cycles, and a shaper one reflecting the faster changing portions.

The estimated frequency of this signal will be higher or equal to that of the vocal folds' vibration cycle, but not faster than the local instantaneous frequency observed around the maximum contact instants. One masking signal produced with the observed frequency is shown in the third panel from top of Fig. 15 and the signal submitted to the sifting operation in the panel just below. This signal displays the cycles of the masking signal, deformed by the input signal. Slow changes of the input signal result in changes of amplitude of the summed signal. As such, they are captured by the envelope computed during the sifting process and removed from the signal (see the second panel from the bottom in Fig. 15). Variability in the partial IMF due to the phase difference between the masking signal and the input signal is handled by averaging between the results obtained by using different initial phase values for the masking signals (see the bottommost panel of Fig. 15).

4. Discussion and conclusions

We proposed an analysis pipeline which, based on the masked sifting operator, permits extracting the main oscillatory component from an observed experimental time series so to obtain a continuous time varying estimate of the phase. The pipeline permits processing oscillatory signals whose behaviour deviates from that of a narrowband signal with relatively slow changing amplitude. Moreover, it permits the analysis of patterns of activity deviating to some extent from a sinusoidal shape.

The proposed pipeline is based on the masked Sifting operation, which was originally conceived to function as an adaptive high pass filter, and which in the present approach has been tailored to isolate the main component of the observed oscillatory pattern. This

result is obtained thanks two original features. Firstly, the masking signal frequency is estimated on a signal centred on zero via a single iteration of the classical Sifting operation. Secondly, the centred signals are submitted to an improved amplitude demodulation schema which, by pruning spurious cycles of activity, permits more precise estimates of the masking frequency and of the final instantaneous phase.

The example in Section 3.1 shows that limiting the number of Sifting iterations to one when estimating the masking frequency permits a more robust estimate of the signal phase with respect to a full sifting in presence of weak or sparse oscillations which are faster than the oscillations of interest. This explains the robustness of the obtained phase values to the presence weak components with frequencies higher than that of the target oscillation observed in section 3.2.

Fig. 10 permits appreciating how the two original features introduced in the present work interact in preserving the phase of non-sinusoidal shapes. In classical EMD, non-sinusoidal shapes are often captured by more than one IMF (see the two bottommost panels in the left side column). As discussed by Fabus et al. (2021), masked Sifting can avoid this splitting. This however is only partially true for the heel height patterns analysed. Due to the presence of nonlinearities introducing dramatic changes in the instantaneous frequency, even one iteration of classical sifting makes emerge weak cycles corresponding to slow portions of heel movements (see topmost panel in the right side column). The estimated masking frequency is the average between that of slow and fast cycles and, as a result, the signal obtained from masked Sifting still features spurious cycles, albeit weakened (second panel from top in the right side column). Due to their weaker amplitude, however these cycles are factored out by improved amplitude demodulation.

The examples in section 3.4 show how the proposed approach permits isolating a dominant component in the presence of a weaker or more sparse deterministic component. So the tip of the tongue movement always shows one cycle per utterance because the effect of the lower lip movement on the TTIP sensor is relatively weak. This is not the case for the effect of the tip of the tongue movement on the position of the LLIP or of the JAW sensors, which is relatively strong at slow speech rate, and therefore induces the method to detect two cycles per utterance in the movement of these sensors. In Section 3.5, we showed that the application of adaptive noise removal based on masked EMD permits obtaining the same results in the presence of measurement noise. The example in section 3.1, illustrating the behaviour of the proposed method with synthetic data, deserves some additional comment. Indeed the main component of the observed signal (displayed in the third panel from top in the leftmost column of Fig. 4) does not display the variability in phase dynamics observed in the output of our method (third panel from top of Fig. 7). This occurs because the deformations of the main cycle, that in the synthetic signal are due to the intermittent weak activity of the fastest component, are mainly considered as due to local changes in amplitude or frequency of the dominant oscillatory patterns. Despite in the case illustrated in Section 3.1, this interpretation leads to mix features of two different components in one sequence of phase values, it permits a correct characterization of shapes that deviate from oscillatory behaviour as in the case of the pulse signal analysed in Section 3.2 or of the vibration signal analysed in Section 3.6. Importantly, when the task under study is defined by a pattern of behaviour of an observed quantity, variability in the temporal features of the observed quantity affects the task performance and therefore it should be reflected in the behaviour of the computed phase. This consideration holds true, even when the perturbations may be due to an underlying component oscillating faster than the dominating component (as it is the case for the signal analysed in Section 3.1).

Once the main component underlying an observed signal is extracted, the proposed approach aims to estimate the phase that reflects the underlying cyclical behaviour. Therefore, the phase of a signal with variations in its local rate of change will naturally exhibit strong modulations (see Appendix 7.6). This design feature is important to consider if the analysis goal is not to characterize the dynamics of the observed variables themselves, but rather the behaviour of some abstract system influencing them during a specific task. For example, Kralemann et al. (2008) aimed to estimate the continuous phase of limit cycle dynamics in autonomous oscillators underlying observed three-channel electrocardiogram (ECG) signals. They hypothesized that each channel is driven by the same underlying limit cycle dynamics ($\dot{x}_0(t)$), though observed through different measurement functions $\dot{x}(t) = h\left(\dot{x}_0(t)\right)$ that could introduce their own phase variations. To obtain phase estimates independent of measurement functions, Kralemann et al. (2008) proposed an analysis procedure based on the assumption that limit cycle dynamics exhibit small deviations from linearity due to a constant average frequency. Their method, when applied to a phase signal with modulations due to non-sinusoidal cycle shape, yields a generally linearly increasing phase that is independent of the observed signal's shape and ensures that signal values spaced by 2π are maximally similar.

If the aim of the analysis is that of capturing features of the observed signal's dynamics, which are assumed to reflect the task under study, then the method proposed in this paper can be used directly. However, if the aim of the analysis is to characterize the underlying limit cycle dynamics hypothesized to be driving the observed motor behaviour, then the observed nonlinearly growing phase needs further processing using the approach proposed by Kralemann et al. (2008) or an equivalent method.

The results of our work have implications that extend beyond its scope and should be considered in the context of decomposition strategies. More specifically, our approach lead to a revised version of masked EMD algorithms (Fabus et al., 2021; Wang et al., 2018) in which the dominant frequency of each component needs to be estimated before the application of the masked sifting operator. The results presented in this paper suggest that, for the purpose of estimating the masking signal's frequencies the number of classical Sifting iterations should be limited to one and the obtained centred signal should be submitted to the refined demodulation operation to remove the effects of spurious cycles due to nonlinearities. Future work will be devoted to investigate the benefits this approach for the decomposition of general multicomponent signals.

Author agreement statement

I the undersigned declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

I confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed.

CRediT authorship contribution statement

Leonardo Lancia: Writing – original draft, Visualization, Software, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Data availability

Data and methods described in this manuscript are avialable in the following GitHub repository: https://github.com/Leonardo-Lancia/Get_Phase_Tools

Acknowledgments

This work, carried out within the Convergence Institute ILCB, was supported by grants from France 2030 (ANR-16-CONV-0002) and the Excellence Initiative of Aix-Marseille University (A*MIDEX) and by grants ANR-21-CE28-0015-01 and ANR-21-CE28-0010-02 of the French National Research Agency.

Appendix A. Appendix: Implementational details

A.1. Sifting

The sifting operator (indicated by $S_q(x)$) takes as argument x: a real valued signal of length N and sampling frequency sr and, in the notation adopted here, it is parameterized by the positive integer q.

Let $maxs(\bullet)$ be the function that takes as argument a real valued unidimensional signal and returns its local maxima.

Let $interp(t_0, v, t_1)$ represent the operator that interpolates at time points t_1 the values in v observed at time points t_0. by using monotone piecewise cubic interpolation (Fritsch & Butland, 1984).

Let $isIMF(\bullet)$ represent the operator that accepts as argument a real valued signal and outputs one if it fulfils the criteria that define an IMF according to Rilling, Flandrin, and Goncalves (2003) and 0 otherwise.

$r = x$
$t = \{1, \ldots, N\} \times \frac{1}{sr}$
$stopIt = 0$
$nIter = 0$
while $stopIT = = 0$ and $nIter \leq q$
set $le^+ = argmaxs(r)$
set $le^- = argmaxs(-r)$
$minEnv = interp(t[le^-], x[le^-], t)$
$maxEnv = interp(t[le^+], x[le^+], t)$
$\widetilde{r} = \mathrm{r} - (minEnv + maxEnv)/2$
if $isIMF(\widetilde{r})$
$stopIt = 1$
else
$r = \widetilde{r}$
end if
end while

A.2. Stopping criteria and processing of boundary conditions

Concerning the specific criteria adopted to determine if $\widetilde{r}$ is an IMF, we followed Rilling et al. (2003) who propose a method based on two thresholds θ_1 and θ_2 and one tolerance parameter α. Sifting is stopped when:

$\left|\frac{meanEnv[n]}{maxEnv[n]-minEnv[n]}\right| < \theta_2, \forall n \in \{1, \ldots, N\}$, and $\left\langle \left|\frac{meanEnv[n]}{maxEnv[n]-minEnv[n]}\right| < \theta_1 \right\rangle > \alpha$. Where $meanEnv = (maxEnv + minEnv)/2$.

Another implementational aspect of Sifting, as well of all the other operations involving the interpolation of local extrema, concerns the processing of the boundary conditions. In order to obtain the signal values that precede the first local extremum and follow the last

one, we again followed Rilling et al. (2003) who propose mirroring the extrema at the boundaries (we chose to mirror the first/last two extrema).

The Matlab code provided in the supplementary material is based on the 2007 implementation of these criteria, available at the following address: http://perso.ens-lyon.fr/patrick.flandrin/emd.html. The Python port of this code (used in the Python implementation of our algorithm) is available in the package pyEMD (https://github.com/laszukdawid/PyEMD; Pele & Werman, 2009) through the function 'EMD_matlab.py'.

A.3. Refined demodulation

Let *sort()* be the function that takes as input a vector of real values and sorts them in ascending order.

set $stopIt = 0$
while $nIter < maxIterN$ and $stopIt == 0$
$el^{+} = argmaxs(x)$
$el^{-} = argmaxs(-r)$
$el = sort(el^{-} \cup el^{+})$
$Z = length(el)$
$z \in \{1, 2, \ldots, Z\}$
$e = interp(t[el], |x[el]|, t)$
$r = r/e$
if $|r[el[z]]| \leq 1 - \epsilon, \forall z \in \{1, 2, \ldots, Z\}$
$stopIt = 1$
end if
end while

A.4. Masked sifting

Masked sifting takes as arguments x: a real valued signal of length N, a sampling frequency sr; a masking frequency $\omega_{mask} \in R^{+} \mid \omega_{mask} < sr/2$; a relative amplitude value $A_{mask} \in R^{+}$; and an integer representing the number of masking signals to be applied $n_{masks} \in N^{+}$. The operation is indicated as $MS_q(x, \omega_{mask}, A_{mask}, n_{masks})$, where the index q refers to the maximum number of iterations of the sifting process $S_q(\bullet)$.

First, the partial sifts of the signal modified by addition of the different masking signals are computed and the masking signals subtracted from the results:

$\breve{x}_j = S_q(mask_j + x)\text{-}\, mask_j$, $\forall j \in \{1, .., J\}$ and $\forall n \in \{1, .., N.\}$

with $mask_j[n] = A_{mask} cos\left(n\Delta t\omega_{mask} + \phi_{0,j}\right)$. Then the obtained signals are averaged:

$$\breve{x} = \frac{1}{J}\sum_{j=1}^{J}\breve{x}_j.$$

A.5. Adaptive denoising algorithm

To clean the input signal from high frequency random noise, we first apply Masked EMD in which classical sifting is substituted by masked sifting.

A.5.1. Masked EMD

Let r_d represent the input signal at dth iteration of the algorithm ($r_1 = x$).

Let $S_h(\bullet)$ represent the sifting operator, taking as argument a real valued signal of length N, with h indicating the maximum number of iteration of the sifting process.

Let $\varkappa_k(\bullet)$ represent the demodulation procedure in A.3, with k indicating the maximum number of iterations of the normalization operation.

Let $MS_q(\bullet, \omega_{mask}, A_{mask}, n_{masks})$ represent the masked sifting operation accepting as arguments a real valued signal, a frequency value ω_{mask}, a relative amplitude value A_{mask}, and an integer representing the number of masking signals to be applied n_{masks}, with q indicating the maximum number of iterations of the of the centring operation.

Let $length(\bullet)$ represent the operator that gives the number of elements of its argument.

The following pseudocode illustrates the masked EMD algorithm. Note that the free parameter Ac determines the amplitude of the masking signal as a proportion of the range of the input signal after the application of the centering operator.

$stopIt = 0$
d = 1
$r_d = x$
while $stopIt$ =0
$\widetilde{r}_d = S_1(r_d)$
$\widehat{r}_d = \varkappa_5\left(\widetilde{r}_d\right)$
$y_d = \mathscr{H}(\widehat{r}_d)$
$\omega_{\widetilde{r}_d} = \left\langle SGd\left(tan^{-1}\left(\frac{y_d}{\widehat{\widetilde{r}}_d}\right), o_{sg}, l_{sg}\right)\right\rangle$
If d==1
$A_{\check{r}} = Ac\left(max\left(\widetilde{r}_d\right) - min\left(\widetilde{r}_d\right)\right)$
end if
$\check{r}_d = MS_{10}\left(r_d, \omega_{\widetilde{r}_d}, A_{\check{r}}, nMasks\right)$
$\widehat{r}_d = \varkappa_5(\check{r}_d)$
$y_d = \mathscr{H}(\widehat{r}_d)$
$\omega_{\widetilde{r}_d} = \left\langle SGd\left(tan^{-1}\left(\frac{y_d}{\widehat{\widetilde{r}}_d}\right), o_{sg}, l_{sg}\right)\right\rangle$
$r_{d+1} = r_d - \check{r}_d$
$e^+ = maxs(x)$
$e^- = maxs(-x)$
If length$(e^+ \cup e^-) < 1$
$stopIt = 1$
end if
end while
$nIMF = d$

A.5.2. IMFs significance test and reconstruction of the clean signal

We estimate the Hurst exponent of the first IMF. We check that the energy of the obtained modes is significantly different from that of the modes obtained by decomposing random signals with the Hurst exponent computed. Finally we average the IMFs that passed the test.

Let $Hurst(\bullet)$ be the operator that gives the Hurst exponent of its argument.

Let $getHurstCI_k(H_0, n_H, a_H)$ be the operator that:

generates n_H realizations of a random process with Hurst exponent H_0 and amplitude a_H;

extracts the kth IMF from each realization;

computes the boundaries delimiting the 95% confidence interval of the energy values observed across realizations of that IMF.

The following lines illustrate the remaining steps of the denoising procedure.

$H_0 = husrt(\check{r}_1)$
for k in $\{1, \ldots, nIMF\}$
$lowCI, hiCI = getHurstCI_k(H0, nh, ah)$
if $\sqrt{\frac{\sum_{n=1}^{N}\check{r}_d[n]}{N}} \leq hiCI$
$flag_d = 0$
else
$flag_d = 1$
end if
end for
y=$\sum_{d=1}^{N} flag_d \check{r}_d$

A.6. Phase estimation via corrected direct quadrature and curve length

The aim of the figures in this section is that of illustrating how the curve length approach can capture fast phase modulation. In the figures, Hilbert phase (represented by dotted lines) is compared to phase computed by applying the curve length measure to analytic signals obtained via direct curvature (continuous lines). Here we only report results that differed depending on the phase estimation method adopted. In Fig. 16, we display the same signal analysed in Section 3.1 (top panel); In Fig. 17, we reproduce the analysis of the speech articulators movement from Section 3.4. The comparison of the Hilbert phase and of the phase obtained via curve length in

these figures shows that the former represents a smoothed version of the latter. Fig. 18 displays a portion of the Electroglottographic signal analysed in Section 3.6. In the figure we also display the demodulated main oscillatory component, which is the signal from which the phase value is computed. We can see that the curve length phase does not grow during portions of time in which the demodulated main component does not change.

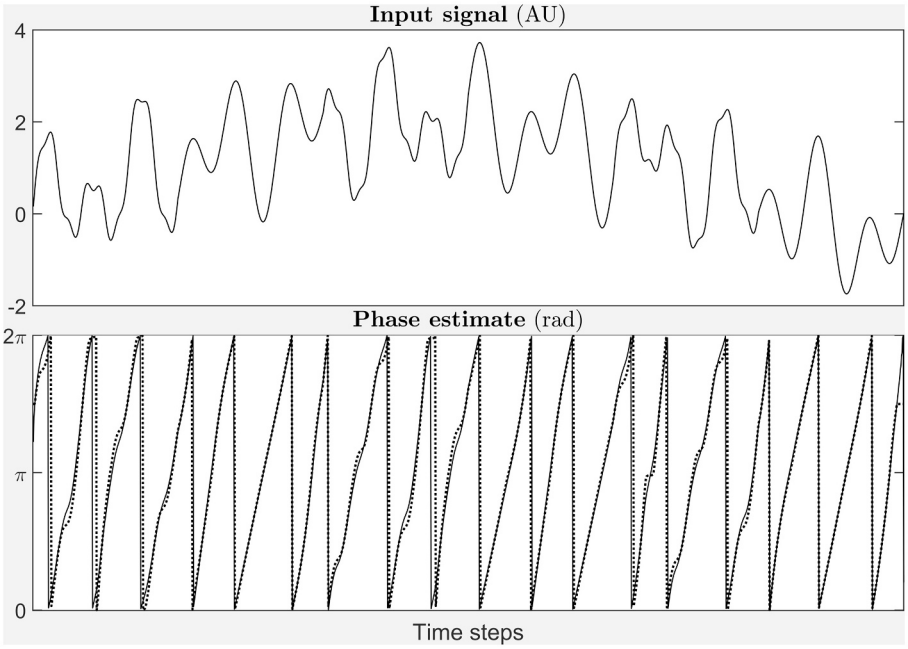


Fig. 16. Reanalysis of the syntetic signal in Fig. 7. Top panel: Input signal. Bottom panel: phase estimates via curve length (continuous line) or Hilbert transform (dotted line).

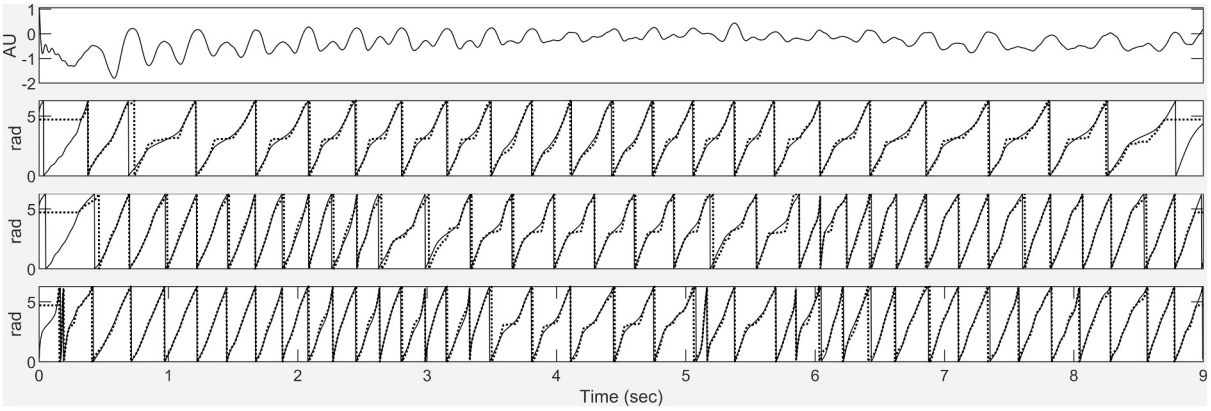


Fig. 17. Reanalysis of the speech articulator trajectories from Section 3.4. Top panel: JAW vertical motion. Remaining panels (from top to bottom): phase of the JAW cycles, of the TTIP cycles and of the LLIP cycles. Continuous lines: curve length based phase. Dotted lines: Hilbert phase.

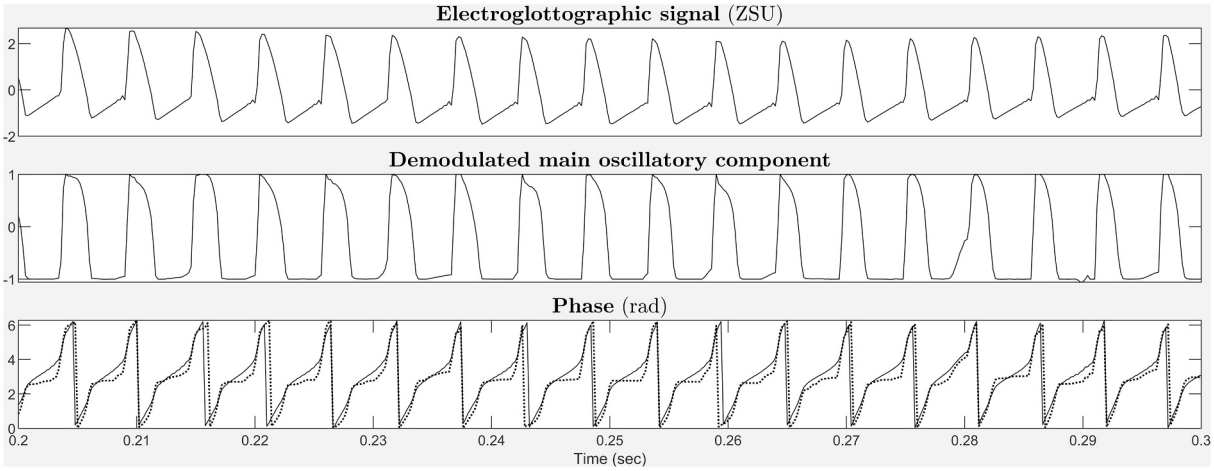


Fig. 18. Phase of a portion of electroglottographic signal from Section 3.6. Topmost panel: *Z*-scored input signal. Middle panel: Demodulated oscillatory component. Bottommost panel: phase. Continuous lines: curve length based phase. Dotted lines: Hilbert phase.

References

Chavez, M., Besserve, M., Adam, C., & Martinerie, J. (2006). Towards a proper estimation of phase synchronization from time series. *Journal of Neuroscience Methods, 154*(1–2), 149–160.

Colominas, M. A., Schlotthauer, G., & Torres, M. A. E. (2014). Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control, 14*, 19–29.

Deering, R., & Kaiser, J. F. (2005). The use of a masking signal to improve empirical mode decomposition. In *, Vol. 4. Proceedings of IEEE international conference on acoustics, speech, and signal processing 2005 (ICASSP'05)* (pp. iv–485). IEEE.

Fabus, M. S., Quinn, A. J., Warnaby, C. E., & Woolrich, M. W. (2021). Automatic decomposition of electrophysiological data into distinct nonsinusoidal oscillatory modes. *Journal of Neurophysiology, 126*(5), 1670–1684.

Flandrin, P., Goncalves, P., & Rilling, G. (2004). Detrending and denoising with empirical mode decompositions. In *12th European IEEE signal processing conference* (pp. 1581–1584).

Fosso, O. B., & Molinas, M. (2018). EMD mode mixing separation of signals with close spectral proximity in smart grids. In *2018 IEEE PES innovative smart grid technologies conference Europe (ISGT-Europe)* (pp. 1–6). IEEE.

Fritsch, F. N., & Butland, J. (1984). A method for constructing local monotone piecewise cubic interpolants. *SIAM Journal on Scientific and Statistical Computing, 5*(2), 300–304.

Gengel, E., & Pikovsky, A. (2019). Phase reconstruction with iterated Hilbert transforms. In *Physics of biological oscillators: New insights into non-equilibrium and non-autonomous systems* (pp. 191–208). Cham: Springer International Publishing.

Goldstein, L. (2019). The role of temporal modulation in sensorimotor interaction. *Frontiers in Psychology, 10*, 2608.

Goswami, U. (2019). Speech rhythm and language acquisition: An amplitude modulation phase hierarchy perspective. *Annals of the New York Academy of Sciences, 1453*(1), 67–78.

Hebenstreit, F., Leibold, A., Krinner, S., Welsch, G., Lochmann, M., & Eskofier, B. M. (2015). Effect of walking speed on gait sub phase durations. *Human Movement Science, 43*, 118–124.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., … Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *, 454(1971). Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* (pp. 903–995).

Huang, N. E., Wu, Z., Long, S. R., Arnold, K. C., Chen, X., & Blank, K. (2009). On instantaneous frequency. *Advances in Adaptive Data Analysis, 1*(2), 177–229.

James, E. G. (2014). Nonstationarity of stable states in rhythmic bimanual coordination. *Motor Control, 18*(2), 184–198.

Kelso, S. (1999). *Dynamic patterns : The self-organization of brain and behavior*. Mit Press.

Kralemann, B., Cimponeriu, L., Rosenblum, M., Pikovsky, A., & Mrowka, R. (2008). Phase dynamics of coupled oscillators reconstructed from data. *Physical Review E, 77*(6), Article 066205.

Kralemann, B., Frühwirth, M., Pikovsky, A., Rosenblum, M., Kenner, T., Schaefer, J., & Moser, M. (2013). In vivo cardiac phase response curve elucidates human respiratory heart rate variability. *Nature Communications, 4*(1), 2418.

Lancia, L., Chaminade, T., Nguyen, N., & Prevot, L. (2017). *Studying the link between inter-speaker coordination and speech imitation through human-machine interactions. (In Interspeech 2017)*. August.

Lancia, L., & Rosenbaum, B. (2018). Coupling relations underlying the production of speech articulator movements and their invariance to speech rate. *Biological Cybernetics, 112*, 253–276.

Longo, G., & Gozzano. (2021). Confusing biological rhythms and physical clocks. Today's ecological relevance of Bergson-Einstein debate on time. In A. Campo (Ed.), *Eintein vs Bergson. An enduring quarrel of time*. De Gruyter.

Matsuki, A., Kori, H., & Kobayashi, R. (2023). An extended Hilbert transform method for reconstructing the phase from an oscillatory signal. *Scientific Reports, 13*(1), 3535.

Melanson, A., Mejias, J. F., Jun, J. J., Maler, L., & Longtin, A. (2017). Nonstationary stochastic dynamics underlie spontaneous transitions between active and inactive behavioral states. *Eneuro, 4*(2).

Miyata, K., & Kudo, K. (2014). Mutual stabilization of rhythmic vocalization and whole-body movement. *PLoS One, 9*(12), Article e115495.

Mörtl, A., Lorenz, T., & Hirche, S. (2014). Rhythm patterns interaction-synchronization behavior for human-robot joint action. *PLoS One, 9*(4), Article e95195.

Pele, O., & Werman, M. (2009). Fast and robust earth mover's distances. In *Proc. 2009 IEEE 12th Int. Conf. On computer vision* (pp. 460–467). Kyoto, Japan.

Pikovsky, A., Rosenblum, M., & Kurths, J. (2003). *Synchronization: A universal concept in nonlinear science*. Cambridge University Press.

Rilling, G., Flandrin, P., & Goncalves, P. (2003). On empirical mode decomposition and its algorithms. In *IEEE-EURASIP workshop on nonlinear signal and image processing. 3*(3) (pp. 8-11).

Rochet-Capellan, A., & Schwartz, J. L. (2007). An articulatory basis for the labial-to-coronal effect:/pata/seems a more stable articulatory pattern than/tapa. *The Journal of the Acoustical Society of America, 121*(6), 3740–3754.

Rosenblum, M., Pikovsky, A., Kurths, J., Schäfer, C., & Tass, P. A. (2001). Phase synchronization: From theory to data analysis. In *, Vol. 4. Handbook of biological physics* (pp. 279–321). North-Holland.
Sandoval, S., & De Leon, P. L. (2017). Advances in empirical mode decomposition for computing instantaneous amplitudes and instantaneous frequencies. In *IEEE international conference on acoustics, speech and signal processing (ICASSP 2017)* (pp. 4311–4315). IEEE.
Schäfer, C., Rosenblum, M. G., Kurths, J., & Abel, H. H. (1998). Heartbeat synchronized with ventilation. *Nature, 392*(6673), 239–240.
Schöner, G., & Kelso, J. A. S. (1988). A synergetic theory of environmentally-specified and learned patterns of movement coordination. *Biological Cybernetics, 58*(2), 71–80. https://doi.org/10.1007/bf00364153
Tuller, B., & Kelso, J. S. (1984). The timing of articulatory gestures: Evidence for relational invariants. *The Journal of the Acoustical Society of America, 76*(4), 1030–1036.
Varlet, M., & Richardson, M. J. (2011). Computation of continuous relative phase and modulation of frequency of human movement. *Journal of Biomechanics, 44*(6), 1200–1204.
Wang, D., & Narayanan, S. S. (2007). Robust speech rate estimation for spontaneous speech. *IEEE Transactions on Audio, Speech and Language Processing, 15*(8), 2190–2201.
Wang, G., Chen, X. Y., Qiao, F. L., Wu, Z., & Huang, N. E. (2010). On intrinsic mode function. *Advances in Adaptive Data Analysis, 2*(03), 277–293.
Wang, Y. H., Hu, K., & Lo, M. T. (2018). Uniform phase empirical mode decomposition: An optimal hybridization of masking signal and ensemble approaches. *IEEE Access, 6*, 34819–34833.
Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis, 1*(01), 1–41.
Wu, Z., & Huang, N. E. (2010). On the filtering properties of the empirical mode decomposition. *Advances in Adaptive Data Analysis, 2*(04), 397–414.