On the quantification of heart rate variability spectral parameters using time–frequency and time-varying methods

BY LUCA T. MAINARDI*

Dipartimento di Bioingegneria, Politecnico di Milano, 20133 Milano, Italy

In the last decades, one of the main challenges in the study of heart rate variability (HRV) signals has been the quantification of the low-frequency (LF) and high-frequency (HF) components of the HRV spectrum during non-stationary events. At this regard, different time–frequency and time-varying approaches have been proposed with the aim to track the modification of the HRV spectra during ischaemic attacks, provocative stress testing, sleep or daily-life activities. The quantitative evaluation of power (and frequencies) of the LF and HF components has been approached in various ways depending on the selected time–frequency method. This paper is an excursus through the most common time–frequency/time-varying representation of the HRV signal with a special emphasis on the algorithms employed for the reliable quantification of the LF and HF parameters and their tracking.

Keywords: time-varying autoregressive models; time–frequency representations; Wigner–Ville distribution; wavelet transform; low-frequency and high-frequency spectral components; autonomic nervous system

1. Introduction

The heart rate variability (HRV) is the beat-to-beat oscillation of RR intervals around its mean value. It is the result of complex regulation mechanisms through which the autonomic nervous system (ANS) rules heart rate (HR) and keeps cardiovascular parameters within physiological ranges. Power spectral analysis of RR series is a well-established tool for the analysis of this kind of data, and it provides significant non-invasive information on ANS activity (Task Force 1996).

A typical spectrum of RR series is characterized by two main components (figure 1): a high-frequency (HF) component (synchronous with respiration, generally in the band 0.2–0.45 Hz) and a low-frequency (LF) one (centred approx. 0.1 Hz, band 0.03–0.15 Hz). While the HF peak is mainly modulated by the parasympathetic (vagal) branch of the ANS, producing the respiratory sinus arrhythmia (RSA), the LF component is mainly the result of
sympathetic modulation, even if the influence of vagal tone may be observed. Consequently, the ratio of LF to HF power is often considered a measure of sympathovagal balance.

Traditional power spectral analysis of HRV data (employing either parametric or non-parametric approaches) implies the signal to be stationary. To overcome this limitation and to describe the changes in HRV spectra during transients of various nature, in the last two decades, tools for simultaneous time–frequency analysis of HRV data have been proposed. They include, among others, the short-time Fourier transform, time-varying autoregressive (AR) analysis (Bianchi et al. 1993), time–frequency representations (TFR; Novak & Novak 1993) and wavelet decomposition (Toledo et al. 2003). While in early 2000 (and before), the emphasis was focused on the selection of the most appropriate tool for time–frequency analysis of HRV data, more recently attention has been moved towards the development of algorithms for the reliable quantification of the LF and HF parameters. This paper aims to review the most popular tools for the TFR of the HRV signal and the algorithms employed for the tracking of LF and HF parameters.

Figure 1. (a) (i) RR interval series during a rest-to-stand manoeuvre and (ii) the corresponding time–frequency plot obtained by the recursive least-squares method; trend of LF (thin curve) and HF power (thick curve) expressed in normalized units (Task Force 1996). (iii) The trend evidences an augmenting sympathetic tone after tilt; tilt occurred around \( t = 120 \) s. (b) Example of spectra extracted from the TFR during (i) rest and (ii) tilt. The pole diagrams are added to evidence the relationship of poles and spectral components.
2. Time-varying AR analysis

Time-varying analysis of HRV supposes the sequence of RR intervals, $x(\cdot)$, to be generated by a time-varying AR model of $p$-order

$$x(t) = \varphi(t)^T \cdot \theta(t) + d(t), \quad (2.1)$$

where $t \in \{-\infty, \ldots, 1, 2, 3, \ldots\}$ is the discrete-time index; $\theta(t) = [a_1(t), a_2(t), \ldots, a_p(t)]^T$ is the parameter vector (i.e. the model coefficients); $\varphi(t) = [x(t-1), x(t-2), \ldots, x(t-p)]^T$ is the observation vector; and $d(t)$ is a scalar, zero-mean, white noise disturbance term of variance $\sigma_d^2 = E[d^2(t)]$, $\mathcal{N}(0, \sigma_d^2)$. In equation (2.1), the parameter vector is made time dependent to take into account that system dynamics may change with time. Without loss of generality, we may describe these changes by

$$\theta(t + 1) = \theta(t) + \delta \theta(t), \quad (2.2)$$

where $\delta \theta(t)$ is known as the drift term. If no a priori information is available, $\delta \theta(t)$ is assumed to be a zero-mean variable with $\langle E[\|\delta \theta(t)\|^2] = \Delta_\theta^2 \rangle$ and independent of $\{x(t), d(t), \delta \theta(t-1)\}$.

The parameter vector is usually unknown and must be estimated from the data. Various algorithms are available for this task and those pertinent to the analysis of HRV series will be reviewed in the next sections. Here, it is sufficient to say that, when the current estimate of $\hat{\theta}(t)$ is available, time–frequency information is obtained by computing a time-varying spectrum

$$T_{AR}(\omega, t) = \frac{\sigma_d^2(t)/f_s}{1 + \theta(t)^T \cdot Z(\omega)^2} = \frac{\sigma_d^2(t)/f_s}{A(z, t) A(1/z, t)} \bigg|_{z = e^{j\omega}}, \quad (2.3)$$

where $Z(\omega) = [e^{-j\omega}, e^{-j2\omega}, \ldots, e^{-jp\omega}]$, $w = 2\pi f / f_s$ and $f_s$ being the sampling frequency, and where $A(z, t) = \prod_{j=1}^p (z - z_i(t))$, with $z_i(t)$’s being the model poles.

(a) Recursive least-squares estimation

The problem of estimating $\hat{\theta}(t)$ can be solved by using the exponentially weighted least-squares (EWLS) algorithm, which is obtained by minimizing the loss function

$$J(\theta) = \sum_{t=-\infty}^{t} \lambda^{t-\tau} \|x(\tau) - \varphi(\tau)^T \theta(\tau)\|^2, \quad (2.4)$$

where $e(\tau) = x(\tau) - \varphi(\tau)^T \theta(\tau)$ is the model prediction error. The coefficient $\lambda$ ($0 < \lambda \leq 1$) is the forgetting factor that progressively discounts old data in favour of the new information and controls the algorithm responsiveness to model changes: the lower the forgetting factor is, the higher the algorithm responsiveness becomes. The minimum of (2.4) is achieved when

$$\hat{\theta}(t) = Q(t)^{-1} R(t), \quad (2.5)$$

*Phil. Trans. R. Soc. A* (2009)
where

\[
R(t) = \sum_{\tau=-\infty}^{t} \lambda^{t-\tau} \varphi(\tau)x(\tau) = \lambda R(t-1) + \varphi(t)x(t)
\]

\[
Q(t) = \sum_{\tau=-\infty}^{t} \lambda^{t-\tau} \varphi(\tau)\varphi(\tau)^T = \lambda Q(t-1) + \varphi(t)\varphi(t)^T
\]

and where \(Q(t)\) is the so-called covariance matrix of the algorithm.

A crucial point is the selection of the appropriate \(\lambda\), which is not a trivial task, and it may affect the performance of the estimation. Campi (1994) showed that the parameter estimation error is bounded

\[
\|\theta(t) - \hat{\theta}(t + 1)\|^2 \leq \alpha_1 \Delta_\theta^2 L + \alpha_2 \sigma_\theta^2 \frac{1}{L},
\]

where \(\alpha_1\) and \(\alpha_2\) are the two coefficients depending on process characteristics and \(L(\lambda) = 1/(1 - \lambda)\) is the memory length of the algorithm indicating approximately how many samples are taken into account to produce the estimate.

The two terms on the right-hand side of (2.7) describe how the parameters’ drift and the noise affect the estimation error. The first term—known as tracking error—increases linearly with \(L\); thus, small \(L\) values should be selected to keep the error bounded in the presence of large \(\Delta_\theta^2\) (i.e. large variations in the parameter vector). The second term—known as steady-state error—is inversely proportional to \(L\); thus, high \(L\) values are required to smooth the effect of noise. This theoretical framework supports the common knowledge that when \(\lambda\) is close to 1 (higher \(L\) value), the EWLS algorithm is insensitive to noise but is not able to track fast changes. Conversely, when \(\lambda\) is decreased (lower \(L\)), the algorithm is more reactive to parameter changes but becomes more sensitive to noise. Equation (2.7) shows that an optimum \(L\), which minimizes the parameter estimation error, does exist, but its value depends on the parameters \(\alpha_1\) and \(\alpha_2\), usually unknown in practice. In the analysis of HRV data, \(\lambda=0.98\) is commonly employed for a memory length of approximately 50 samples (Mainardi et al. 2002).

(i) Recursive formulation

Even if equations (2.5) and (2.6) provide a way to estimate the model parameters, this solution is not efficient as it requires, at every step, the inversion of matrix \(Q(t)\). A more effective implementation makes use of a recursive updating of \(P(t) = Q(t)^{-1}\), leading to the recursive least-squares algorithm (EW-RLS)

\[
\hat{\theta}(t) = \hat{\theta}(t-1) - K(t)[x(t) - \varphi(t)^T \theta(t)]
\]

\[
K(t) = \frac{P(t-1)\varphi(t)}{\lambda + \varphi(t)^T P(t-1)\varphi(t)}
\]

\[
P(t) = \frac{1}{\lambda} (I - K(t)\varphi(t)^T) P(t-1).
\]

For each new sample \(x(t)\), the parameter vector is updated by a quantity proportional to the model prediction error. The larger the error, the bigger the variation in model parameters and, in a picturesque way, we can say that the
Algorithm is learning from its own mistakes. A key role is played by $K(t)$, the gain of the algorithm (updated step by step), which defines the property of the algorithm and modulates the extent to which prediction errors are reflected in the update of model parameters.

(ii) Whale forgetting

In order to overcome the limitation imposed by exponential forgetting, a more generalized recursion rule can be considered (Bianchi et al. 1997)

$$
\begin{align*}
R(t) &= c_1 R(t-1) + c_2 R(t-2) + (1 - c_1 - c_2) \varphi(t) x(t) \\
Q(t) &= c_1 Q(t-1) + c_2 Q(t-2) + (1 - c_1 - c_2) \varphi(t) \varphi(t) T, 
\end{align*}
$$

(2.9)

where $c_1$ and $c_2$ are the two coefficients such that the two roots of $z^2 - c_1 z - c_2 = 0$ lie in $(0, 1)$. The advantage of this generalized recursion is that the compromise between responsiveness and noise insensitivity can be made less dramatic than that in (2.7). The resulting weighting function (figure 2) is comparable to the EW-RLS ones in discharging past samples (i.e. the algorithm is equally reactive to the changes in the parameters), but the average of most recent data allows to make the algorithm less sensitive to noise. Owing to the particular shape of the resulting weighting function, the algorithm is known as the whale-forgetting least-squares method.

**Figure 2.** (a) RR interval series recorded during 3 min cold pressure test. (b) (i) Trend of LF power expressed in normalized unit as obtained by exponential forgetting method (black curve) and whale forgetting (grey curve). Both approaches are able to capture the sympathetic activation induced by the test. Note that the whale-forgetting algorithm provides a more smoothed estimate without loosing in reactivity. (ii) Comparison between exponential (black curve) and whale-forgetting (grey curve) functions.
(b) Smoothed Kalman filter

Equations (2.1) and (2.2) form the state-space model for the process $x(\cdot)$, where the AR model parameters represent the states. Thus, the minimum mean square estimator of $\hat{\theta}(t)$, given the observations $\varphi(t)$, can be obtained using the Kalman filter

$$
\dot{\hat{\theta}}(t) = \hat{\theta}(t-1) - K(t) [x(t) - \varphi(t)^T \hat{\theta}(t)] \\
K(t) = \frac{P(t-1) \varphi(t)}{C_d(t) + \varphi(t)^T P(t-1) \varphi(t)} \\
P(t) = (I - K(t) \varphi(t)^T) P(t-1) + C_{\theta\theta}(t),
$$

(2.10)

where $C_d(t) = \sigma_d^2(t)$ is the observation noise covariance and $C_{\theta\theta}(t) = I \sigma_\theta^2(t)$ is the state noise covariance. These two terms have to be estimated. In the analysis of HRV data, Tarvainen et al. (2006) suggested to update $\sigma_d^2(t)$ using the following recursion $\sigma_d^2(t) = \gamma \sigma_d^2(t-1) + (1 - \gamma) [x(t) - \varphi(t)^T \hat{\theta}(t)]^2$ with $\gamma = 0.95$, and to estimate noise covariance as $C_{\theta\theta}(t) = I \lambda \sigma_\theta^2(t)/\sigma_d^2(t)$, $\sigma_\theta^2(t)$ being a running estimate of the variance of the HRV signal. The latter normalization makes the updating independent of the signal amplitude. The role of $\mu$ is analogous to the role of the forgetting factor in EW-RLS. In fact, by increasing $\mu$, the adaptation speed is increased, but at the expense of an augmented variance of the state estimates. Thus, $\mu$ should be specified in such a way that a desired balance between the filter adaptation and the estimate variance is obtained. The analogy between $\mu$ and $\lambda$ becomes evident if one observes the similarity between algorithms (2.8) and (2.10). For detailed description of the relationships between RLS and Kalman filter, see Sayed & Kailath (1994) and Lippuner & Moschytz (2004).

If a small lag in the processing is allowed and future observations can be used, a more accurate estimate of the state may be obtained using the so-called Kalman smoother. Using a fixed time interval $T$ and for $1 \leq t \leq T$, the smoothing recursions for the state estimate $\hat{\theta}^S(t)$ is

$$
\hat{\theta}^S(t-1) = \hat{\theta}(t) - \frac{I(t-1)}{P(t-1)} [\hat{\theta}^S(t) - \hat{\theta}(t-1)],
$$

(2.11)

where $I(t) = (I - K(t) \varphi(t)^T) P(t-1)$. The philosophy of the smoother is to increase the robustness of the estimates through a weighting combination of previous/past estimates.

(c) HRV parameter estimation

One of the advantages of using a time-varying AR model is the possibility to extract the LF and HF parameters using a spectral decomposition method. The method was originally proposed by Zetterberg (1969) for EEG analysis and successively applied to the decomposition of the HRV spectrum (Baselli et al. 1987).
Using Cauchy’s residue theorem, the signal power, i.e. the area under the AR spectrum (2.3), can be divided into a sum of \( p \) components:

\[
\sigma_x^2 = \frac{\sigma_d^2(t)}{2\pi f_s} \int_0^{2\pi} \frac{1}{A(e^{j\omega}, t)A(e^{-j\omega}, t)} d\omega
\]

\[
= -j \frac{\sigma_d^2(t)}{2\pi f_s} \oint_{|z| = 1} \frac{1}{A(z, t)A(1/z, t)} \frac{dz}{z} = \frac{\sigma_d^2(t)}{f_s} \sum_{i=1}^{p} \gamma(z_i(t)), \tag{2.12}
\]

where the quantity

\[
\gamma(z_i(t)) = \left. \frac{z^{-1}(z - z_i(t))}{A(z, t)A(1/z, t)} \right|_{z = z_i(t)} \tag{2.13}
\]

is the pole residue. Consequently, the spectrum can be decomposed into bell-shaped curves, the spectral components (figure 1), resulting in

\[
P(f, t) = \frac{\sigma_d^2(t)}{f_s} \sum_{i=1}^{p} \gamma(z_i(t)) \left\{ \frac{z_i(t)}{z - z_i(t)} + 1 + \frac{z_i(t)}{1/z - z_i(t)} \right\} \tag{2.14}
\]

The centre frequency \( f \) and the power \( P \) of each spectral component can be computed as

\[
f_i(t) \approx f_s \frac{\angle(z_i(t))}{2\pi} \quad P_i(t) = \mu \frac{\sigma_d^2(t)}{f_s} \text{Re}\{\gamma(z_i(t))\}, \tag{2.15}
\]

where \( \angle(\cdot) \) is the phase expressed in radians, and \( \mu = 2 \) for complex pole pairs and \( \mu = 1 \) for real ones. When dealing with the HRV spectra, the calculation of the LF and HF powers may be performed by summing up the power values of those spectral components whose centre frequency lies within the chosen LF and HF bands, respectively.

Evaluation of equation (2.15) requires the computation of AR poles. Traditional factorization methods for pole estimation are time-consuming procedures, not appropriate for continuous pole detection: they need to be restarted each time a new coefficient set is available, and they do not take into account the previous estimated pole values. Therefore, pole-tracking algorithms were proposed for the analysis of HRV data (Mainardi et al. 1995; Meste et al. 2005).

The simplest way to evaluate the effect of parameter changes in the pole positions is to linearize the relationship

\[
1 + \sum_{k=1}^{p} a_k(t)z^{-k} = \prod_{j=1}^{p} (1 - z_i(t)z^{-1}), \tag{2.16}
\]

\( ^1 \)This approximation is valid when \(|z|\) is sufficiently close to 1. More correctly, the maximum of the spectral component is located at \( \omega \) satisfying

\[
\sin(\omega - \angle z_j)(1 - 2(|z_i|/|z_j|^2 + 1))\cos(\omega - \angle z_j)^2 + \sin(\omega + \angle z_j)(1 - 2(|z_i|/|z_j|^2 + 1))\cos(\omega + \angle z_j)^2 = 0.
\]
obtaining
\[ \Delta z_i = \sum_{k=1}^{p} \frac{\partial z_i}{\partial a_k} \Delta a_k = \sum_{k=1}^{p} \frac{z_i^{p-k}}{\prod_{l=1, l\neq j}^{p} (z_i - z_l)} \Delta a_k. \] (2.17)

Using (2.17), for each new parameter vector, the pole positions can be updated as (Mainardi et al. 1995)

\[ z_i(t) = z_i(t-1) + \Delta z_i = z_i(t-1) + \left[ \frac{\partial z_i}{\partial a_1}, \frac{\partial z_i}{\partial a_2}, \ldots, \frac{\partial z_i}{\partial a_p} \right] (\hat{\theta}(t) - \hat{\theta}(t-1)). \] (2.18)

The above relation is based on a linearization procedure that applies only to small coefficient changes. As the estimation procedure continues, error accumulation can corrupt the pole estimations. For this reason, once the poles are updated, the parameter vector is recalculated on the basis of the new \( z_i \)'s. The procedure may be iterated, if needed. The \( \frac{\partial z_i}{\partial a_k} \) provides a measure of the sensitivity of the \( j \)th pole position to the changes in the \( k \)th model coefficient. This sensitivity depends on the reciprocal position of the poles and increases when the pole distance is reduced. Thus, in the case of closely grouped poles, even small changes in the parameters may overshoot pole changes (Oppenheim & Schafer 1975). Finally, it can be demonstrated that the algorithm is unable to generate a complex pole pair from two real poles; the desired number of complex poles must be provided during initialization.

To overcome these limitations, a second algorithm was presented, which is based on the iteration procedure of the well-known Bairstow algorithm, commonly employed to compute the roots of a polynomial function. Details and comparisons between the two methods applied to HRV data can be found in Mainardi et al. (1995).

In the work of Meste et al. (2005), in order to avoid problems related to variations in pole configurations (a complex pole pair becoming two real poles or vice versa), the AR model of even order is considered and the following factorization is used:

\[ 1 + \sum_{k=1}^{p} a_k(t) z^{-k} = \prod_{j=1}^{p/2} (1 - b_{i,1}(t) z^{-1} - b_{i,2}(t) z^{-2}). \] (2.19)

Tracking is then performed on coefficients \( b_i \) instead of poles \( z_i \), and reflected on pole movements using classical root calculation formula for the second-order polynomials. A formulation analogue to (2.18) was obtained and the efficient way to compute partial derivative is presented using a matrix formed by multiple convolutions. Finally, the authors propose an evolutive AR modelling (i.e. a model in which the \( a_i \) coefficients are constrained to be a linear combination of some basis functions) to provide a time-continuous description of the model coefficients and obtain a robust track.

### 3. Quadratic TFR

The time–frequency distribution of the signal energy may be obtained by quadratic TFR, among which the Wigner–Ville (WV) distribution plays a central role. It is
defined as the Fourier transform of the instantaneous autocorrelation function (ACF), sometimes referred to as the central covariance function \(2\)

\[
T_{WV}(f, t) = \int_{-\infty}^{\infty} R_x(\tau, t)e^{-j2\pi f\tau} \, d\tau = \int_{-\infty}^{\infty} x\left(t - \frac{\tau}{2}\right)x^*\left(t + \frac{\tau}{2}\right)e^{-j2\pi f\tau} \, d\tau. \quad (3.1)
\]

The WV satisfies several desirable mathematical properties for a quadratic TFR (Hlawatsch & Boudreaux-Bartels 1992): it is real valued and preserves time and frequency shift of the signal; it satisfies the marginal properties, and thus the frequency (or time) integral of WV corresponds to the signal’s instantaneous power (or power spectrum); and also the instantaneous frequency can be estimated from the first moment of the WV. The properties of (3.1) have been extensively discussed in Classen & Mecklenbrauker (1980).

Unfortunately, WV is affected by the interference terms (or cross-terms) that prohibit the interpretation of the WV as a pointwise time–frequency energy density and widely limit the readability of the distribution. The cross-terms oscillate in the \(t–f\) plane and can be attenuated by using kernels \(\psi\), obtaining the Cohen’s class (Cohen 1989)

\[
T_C(f, t) = \int \int \psi(t - \theta, f - \nu) T_{WV}(\theta, \nu) \, d\theta \, d\nu, \quad (3.2)
\]

where it is apparent that every member of the class may be interpreted as a two-dimensional, filtered version of the WV and is unequivocally defined by the kernel \(\psi\). Among the various transformations belonging to Cohen’s class, the smoothed pseudo-Wigner–Ville (SPWV) has been largely employed for the analysis of HRV data.

(a) Smoothed pseudo-Wigner–Ville

In its discrete-time and discrete-frequency formulations, the SPWV is defined as

\[
T_{SPWV}(f_m, t) = 2 \sum_{k=-K+1}^{K-1} |h(k)|^2 \left( \sum_{p=-M+1}^{M-1} g(p) r_x(t + p, k) \right) e^{-j2\pi kf_m}, \quad (3.3)
\]

where \(f_m = m/K\); \(h(k)\) is the frequency smoothing symmetric normed window of length \(2K-1\); \(g(p)\) is the time smoothing symmetric normed window of length \(2M-1\); and \(r_x(n, k) = x(n + k)x^*(n - k)\).

The properties of the SPWV depend on the characteristics of \(g(\cdot)\) and \(h(\cdot)\), which control time–frequency resolutions of the distribution. In respect to other representations of the Cohen’s class, the SPWV is characterized by a separable kernel, in which time and frequency smoothing can be tuned independently, resulting in a great flexibility in the selection of smoothing criteria, application and efficient computation. This aspect is also relevant for the analysis of HRV series where the requirements for time–frequency resolutions are different. While frequency resolution is not problematic, as the spectrum of HRV is characterized by a few spectral components (mainly the LF and HF ones), time resolution should be increased in order to track spectral changes when they occur. In HRV data, therefore, time resolution is stressed, while frequency resolution is not.

\(2\)From this point on, \(t\) represents the continuous time.
The selection of the most appropriate smoothing windows for the analysis of cardiovascular variability series has been originally investigated by Novak & Novak (1993), using rest-to-tilt data and controlled reparation experiments. They suggested using a rectangular time-domain smoothing (window length between 5 and 25 samples) and a Gaussian window for frequency smoothing \((h(k) = \exp(\alpha k)^2/(N/2)\) with \(\alpha = 2.5\) and \(N = 128\)). Comparing the performance of the SPWV with other TFRs, Pola et al. (1996) employed a Gaussian (23 samples length) and a hamming (128 samples) window for time- and frequency-domain filtering, respectively. In both works, the length of the time window is short, i.e. time resolution is stressed and limited to a few tens of seconds.

(b) HRV parameter extraction

The parameters of the LF and HF components are usually extracted from the SPWV distribution using the marginal and the first momentum properties (even if they are not strictly valid), thus obtaining, for the LF components (analogous formulation can be used for HF one),

\[
P_{LF}(t) = \sum_{m=m_j}^{m_k} T_{SPWV}(f_m, t) \Delta f
\]

and

\[
f_{LF}(t) = \frac{\sum_{m=m_j}^{m_k} f_m \cdot T_{SPWV}(f_m, t)}{\sum_{m=m_j}^{m_k} T_{SPWV}(f_m, t)}
\]

where the LF band is delimited by \([f_{mj}; f_{mk}]\). Alternatively, the instantaneous frequency is obtained as the maximum in the band

\[
f_{LF}(t) = \arg \max_{f_m \in [f_{mj}, f_{mk}]} T_{SPWV}(f_m, t).
\]

More recently, the work of Mainardi et al. (2004) has proposed a SPWV decomposition method for the quantification of the LF and HF contributions. The method is based on the decomposition of the instantaneous ACF and it assumes that the instantaneous frequencies of the LF and HF components are, locally, linearly varying with time. In particular, we assume that the analytic HRV signal is described as

\[
x(t) = A_{LF} \exp(j2\pi(\alpha_{LF} t^2 + \beta_{LF} t)) + A_{HF} \exp(j2\pi(\alpha_{HF} t^2 + \beta_{HF} t)),
\]

where \(A_{LF}\) and \(A_{HF}\) are the instantaneous amplitudes of the LF and HF components, whose frequencies are \(f_{LF}(t) = 2\alpha_{LF} t + \beta_{LF}\) and \(f_{HF}(t) = 2\alpha_{HF} t + \beta_{HF}\), respectively. If we compute \(r_x(n, k)\) for the signal in (3.7), we get

\[
r_x(t, k) = |A_{LF}|^2 \exp(j2\pi 2f_{LF}(t)k) + |A_{HF}|^2 \exp(j2\pi 2f_{HF}(t)k)
\]

\[
+ A_{LF} A_{HF} \cos(2\pi g(\alpha_{HF}, \alpha_{LF}, \beta_{LF}, \beta_{HF}, k, t)) \exp(j2\pi \frac{f_{LF}(t) + f_{HF}(t)}{2} k),
\]

(3.8)
where the last term is the interference term, which is located in between the two components and oscillates in the time–frequency plane as expected (Hlawatsch & Boudreaux-Bartels 1992). Equation (3.8) is therefore composed by the sum of complex sinusoids including both the signal contributions and the interference terms. Note that the signal components contain information on the instantaneous frequencies and amplitudes of the LF and HF components. Equation (3.8) has been derived for two components only, but the extension to an $N$-component signal is straightforward using the quadratic superimposition principle (Flandrin 1984).

Using an exponential window for frequency smoothing $h(k) = e^{-\gamma|k|}$ and a rectangular window for time smoothing, it can be shown that the quantity

$$e^{-2\gamma|k|} \left( \sum_{p=-M+1}^{M-1} g(p) r_x(t + p, k) \right) \approx \sum_{i=1}^{I} A_i(t) e^{-b_i(t)k} \exp(j2\pi2f_i(t)k),$$

of equation (3.3) can be still approximated with a sum of complex sinusoids, where the influences of cross-terms are vanished by the application of smoothing windows and where the remaining components should be mainly related to signal contributions. For more details on how the assumption holds for a rectangular time-smoothing window, see Bailón et al. (2006a).

The problem of estimating the HRV spectral parameters is now reduced to the problem of estimating the $A_i(t)$, $f_i(t)$, $b_i$ of (3.9). This can be solved using, for example, the method proposed by Kumaresan & Tufts (1982), which provides accurate detection and estimation of exponentially damped sinusoidal signals. The method combines linear backward prediction and singular value decomposition and it is particularly robust against noise. An example of decomposition of the SPWV is shown in figure 3.

Figure 3. (a) (i) RR interval series obtained for a normal volunteer during stress testing, (ii) the TFR obtained by SPWV and (iii) trend of LF power. (b) (i) RR series, (ii) tracking of LF and HF frequencies and (iii) trend of HF power. Note that the decrease in LF power during stress (related to reduction in the overall HRV during effort) is accompanied by the persistence of the HF component even at a higher workload. The RR stress data were collected during the study by Bailón et al. (2006b) and kindly provided by the authors.
Finally, it is noteworthy that a priori information on the HRV spectral components can be easily inserted as constraints in the estimation of parameters of (3.9). The inclusion of knowledge about respiratory frequency is discussed in Bailón et al. (2006a).

4. Wavelet analysis

Wavelet analysis has become a very popular tool for time–frequency signal characterization. The continuous wavelet transform (CWT) of a signal \( x(t) \in L^2(\mathbb{R}) \) is defined as (Riou & Vetterli 1991)

\[
\gamma_x(\tau, \alpha) = \langle x(t)\psi^*_{\tau,\alpha}(t) \rangle = \int_{\mathbb{R}} x(t) \frac{1}{\sqrt{\alpha}} \psi^* \left( \frac{t-\tau}{\alpha} \right) dt,
\]

and it measures the similarity between the signal and a family of functions generated by scaling (contracting or dilating) and translating a function prototype \( \psi(\cdot) \), called mother wavelet. The parameter \( \tau \) gives the position of the wavelet in time, while \( \alpha \) is the scaling factor that governs its frequency.

Using an appropriate selection of \( \psi(\cdot) \), the inverse CWT exists and a signal can be represented by its decomposition into wavelets

\[
x(t) = \frac{1}{C_\psi} \int_{\mathbb{R}} \frac{d\alpha}{\alpha^2} \int_{\mathbb{R}} \gamma_x(\tau, \alpha) \psi_{\tau,\alpha}(t) d\tau,
\]

where

\[
C_\psi = \int |\Psi(\omega)|^2 \frac{d\omega}{\omega}
\]

is a constant that depends on \( \psi(t) \) only, and \( \Psi(\omega) \) being the Fourier transform of the wavelets. The existence of (4.2) is guaranteed when \( C_\psi < +\infty \) (the admissibility condition). This condition imposes \( \Psi(0) = 0 \), i.e. the wavelets have zero mean and are oscillating around the \( t \)-axis. In general, as \( |\Psi(\omega)| \to 0 \) when \( \omega \to +\infty \), wavelets have a bandpass-like spectrum.

Using the Parseval equality, we rewrite relation (4.1) in terms of the Fourier transform of \( x(\cdot) \) and \( \psi(\cdot) \)

\[
\gamma_x(\tau, \alpha) = \int X(\omega)\Psi^*_{\tau,\alpha}(\omega) d\omega = \int X(\omega)\sqrt{\alpha}\Psi^* (\alpha\omega)e^{j\omega\tau} d\omega,
\]

which provides a clear interpretation of the CWT in terms of a constant-\( Q \)-filterbank: a rescaling by \( \alpha \) in the time domain becomes a rescaling by \( 1/\alpha \) in the frequency domain, which has the effect of moving the centre frequency of the wavelet bandpass filter from \( f_\psi \) to \( f_\psi/\alpha \), with a similar rescaling of the 3 dB bandwidth, which is reduced when the scale increases. Therefore, the very appealing feature of the wavelet is that \( \alpha \) makes the time resolution dependent on frequency (scale; and vice versa): at larger scales, a dilated wavelet is used to explore the signal reducing the time resolution but increasing the frequency localization of the components; at lower scales, the contraction of the wavelet provides timely location of HF details, at the expense of frequency resolution.
For the purpose of time–frequency analysis of HRV, we make use of the fundamental observation that CWT is isometric (Grossmann & Morlet 1984) and it preserves the signal energy. We may write

\[ E_x = \int_0^T |x(t)|^2 \, dt = \frac{1}{C_{\psi}} \int_\alpha^\infty \frac{d\alpha}{\alpha^2} \int_\tau^\infty |\gamma_x(\tau, \alpha)|^2 \, d\tau, \quad (4.4) \]

which evidences that the squared modulus of wavelet coefficients can be interpreted in terms of energy distribution, where \( |\gamma_x(\tau, \alpha)|^2 \) represents the ‘energy’ of \( x \) at lag \( \tau \) and scale \( \alpha \). Such a distribution is known as scalogram. A detailed discussion of the CWT as energy distribution can be found in Rioul & Flandrin (1992).

Since the representation of the CWT is redundant, wavelet analysis is performed on a discrete set of time-scale parameters. The most common choice set is \( \tau = k\alpha \) and \( \alpha = 2^{-j} \), where \( j, k \) are the integers (i.e. the CWT is sampled on a ‘dyadic’ grid) so that \( \psi_{\alpha,\tau}(t) = \psi_{j,k} = 2^{j/2} \psi(2^j \tau - k) \).

If the basis functions \( \psi_{j,k} \) form an orthonormal basis, a signal can be represented exactly as the weighted sum of these basis functions (Riou & Vetterli 1991) \( x(t) = \sum_{j,k} d_j(k) \psi_{j,k}(t) \), where the \( d_j(k) \)'s are the wavelet

Figure 4. Tracking of HRV parameters using CWT during reperfusion after MI. Time of reperfusion is marked by a vertical dashed line. (a) The HR signal, (b) the scalogram, (c) the trend of LF component, (d) the trend of HF component and (e) their ratio. Note the reduction in LF components following reperfusion and the slow, progressive increase in parasympathetic activity. Reprint from Toledo et al. (2003).

For the purpose of time–frequency analysis of HRV, we make use of the fundamental observation that CWT is isometric (Grossmann & Morlet 1984) and it preserves the signal energy \( E_x \). We may write

\[ E_x = \int_0^T |x(t)|^2 \, dt = \frac{1}{C_{\psi}} \int_\alpha^\infty \frac{d\alpha}{\alpha^2} \int_\tau^\infty |\gamma_x(\tau, \alpha)|^2 \, d\tau, \quad (4.4) \]

which evidences that the squared modulus of wavelet coefficients can be interpreted in terms of energy distribution, where \( |\gamma_x(\tau, \alpha)|^2 \) represents the ‘energy’ of \( x \) at lag \( \tau \) and scale \( \alpha \). Such a distribution is known as scalogram. A detailed discussion of the CWT as energy distribution can be found in Rioul & Flandrin (1992).

Since the representation of the CWT is redundant, wavelet analysis is performed on a discrete set of time-scale parameters. The most common choice set is \( \tau = k\alpha \) and \( \alpha = 2^{-j} \), where \( j, k \) are the integers (i.e. the CWT is sampled on a ‘dyadic’ grid) so that \( \psi_{\alpha,\tau}(t) = \psi_{j,k} = 2^{j/2} \psi(2^j \tau - k) \).

If the basis functions \( \psi_{j,k} \) form an orthonormal basis, a signal can be represented exactly as the weighted sum of these basis functions (Riou & Vetterli 1991) \( x(t) = \sum_{j,k} d_j(k) \psi_{j,k}(t) \), where the \( d_j(k) \)'s are the wavelet

Phil. Trans. R. Soc. A (2009)
coefficients that can be computed using the discrete wavelet transform (DWT). In this case, the signal energy is obtained as $E_x = \sum_{j,k} |d_j(k)|^2$.

(a) HRV parameter extraction

On the basis of relation (4.4), the quantification of the LF and HF components is generally obtained by integrating the squared modulus of the wavelet coefficients over the desired frequency band $[f_1 \ f_2]$. To do that, we have to move from scales to frequency and interpret the time-scale map in terms of a time–frequency distribution. We observe that the point $(t, f, \tau, f_\psi/\alpha)$ in the time-scale plane is mapped in the point $(t, f, \alpha)$, and then using equation (4.4), we may write

$$P(t) = \frac{1}{C_\psi} \int_{\alpha_1}^{\alpha_2} |\gamma_x(t, \alpha)|^2 \frac{d\alpha}{\alpha^2} = \frac{1}{C_\psi f_\psi} \int_{f_1}^{f_2} |\gamma(t, f_\psi/f)|^2 df,$$

which can be used for the computation of instantaneous power in a given frequency band. An example of tracking of HRV parameters during reperfusion after myocardial infarction (MI) is shown in figure 4.

When the DWT is considered, wavelet expansion divides the frequency axis into dyadic blocks and the frequency contents of each details $d_j$ is $[2^{-j-1} f_\delta \ 2^{-j} f_\delta]$. Instantaneous power of the LF and HF components is measured by

$$P(t) = \sum_j |d_j(t)|^2,$$

where the summation is extended to those scales that cover the LF and HF bands. These scales depend on the sampling frequency of the signal. For example, given $f_\delta = 2$ Hz, the LF band will be covered by scales $j=4,5$, resulting in a band interval $[0.0312 \ 0.125] \text{ Hz}$, while the HF band will be covered by scales $j=2, 3$, resulting in the band interval $[0.125 \ 0.5] \text{ Hz}$. Using a larger band for HF is generally not problematic, because only one spectral component (and its harmonic) exists at frequencies greater than 0.18 Hz (Toledo et al. 2003). However, this is only appropriate when the respiratory frequency lies within this band, but counter-examples exist such as during stress testing (Bailón et al. 2007).

5. Other approaches

A few other strategies have been pursued for the analysis of the HRV signals. Two of them are worth mentioning: the empirical mode decomposition (EMD) and complex demodulation (CD) techniques. In the EMD, the signal is decomposed as $x(t) = m_N(t) + \sum_{n=1}^{N} r_n(t)$, where $m_N$ represents the residual trend and $r_n(t)$ are the intrinsic mode functions (IMFs), which are extracted iteratively as follows: (i) identify the extrema (local minima and maxima) of $x(t)$, (ii) interpolate between minima (maxima), obtaining some ‘envelope’ $e_{\min}(t)$ ($e_{\max}(t)$), (iii) compute the average $m(t) = (e_{\min}(t) + e_{\max}(t))/2$, (iv) extract the detail $r(t) = x(t) - m(t)$, and (v) iterate on the residual $r(t)$.

In practice, the above procedure has to be refined by a sifting process (Huang et al. 1998). According to the way they are extracted, the IMFs are constrained to be zero-mean, amplitude modulation, frequency modulation waveforms. Even

Phil. Trans. R. Soc. A (2009)
though they are not necessarily narrow-band signals—as they can be non-
stationary—the IMF should be, locally, monocomponent signals. Therefore, if
one computes the Hilbert transform of the IMF, \( H(R_n) = |a(t)|e^{j\phi(t)} \), the
tracking of instantaneous power and frequency of the IMF are obtained as \( |a(t)|^2 \)
and \( 1/2\pi(d\phi(t)/dt) \), respectively.

In the CD approach, the signal \( x(\cdot) \) is supposed to have a component of
interest in the band \( [f_0 - f_w, f_0 + f_w] \) and the signal is demodulated (i.e.
multiplied by \( \exp(j2\pi f_0) \)), thus shifting the spectrum and centring the band
of interest on the \( f=0 \) axis. By applying a low-pass filtering with a cut-off
frequency \( f_w \), the desired component, let us name it \( x_{f_0}(t) \), is retrieved. If \( f_w \)
is selected sufficiently small (i.e. we have band-limited signal), we may write
\( x_{f_0}(t) \approx |a(t)|e^{j\phi(t)} \) and instantaneous powers and amplitude estimated as
above. The procedure must be repeated for all the frequencies of interest. It is
noteworthy that this is equivalent to the bandpass (band: \( [f_0 - f_w, f_0 + f_w] \)
signal and apply the Hilbert transform to it. Analogy with EMD appears: the CD
provides a signal decomposition in which the IMFs are simply bandpass versions
of \( x(\cdot) \). The main limitation of the CD is related to the selection of \( f_0 \)'s and the
definition of \( f_w \). This turns out to be a non-trivial task, especially for the HF
components. Therefore, \( a \text{ priori} \) information about the frequency location of the
HRV components must be included. The information may be taken from TFR of
the HRV signal (Monti et al. 2002) or from the processing of other biological
signals (i.e. respiration).

6. Applications

Early applications of the above-mentioned approaches dealt with the assessment
of ANS responses during standard autonomic tests, in which a selective stimulus
elicits sympathetic (rest to tilting, cold pressure test, etc.) or parasympathetic
(deep breathing, valsalva manoeuvre) responses. The pronounced, acute changes
after the stimulus are important for grading the autonomic performances and for
evaluating neuroregulatory disorders.

Early responses to tilt have been first studied by the SPWV (Jasson et al.
1997; Mainardi et al. 2002), evidencing both vagal and sympathetic changes: the
almost instantaneous disappearance of the HF component was accompanied by a
slower, progressive increase in the LF components (Jasson et al. 1997). The
response was dependent on the amplitude of the stimulus: a faster (and larger)
sympathetic response was observed for more pronounced tilt angles (Mainardi
et al. 2002). When the tilt was prolonged, oscillations and non-stationarity
characterized the trend of the LF components, being more evident in patients
with syncope (Furlan et al. 1998). Autonomic unbalance before syncope has also
been assessed using the time-varying method (Mainardi et al. 1997; Furlan et al.
1998) documenting a sympathetic withdrawal preceding syncope, which may be
either sudden or progressive. This withdrawal has also been confirmed by the
analysis of MSNA variability using the CD (Kamiya et al. 2005). More recently,
the DWT has been applied to a variety of autonomic tests (Ducla-Soares et al.
2007) and documented to be effective for the quantification of the first stage of
the autonomic performance. In general, dynamic response to stimulus may help
to better differentiate the class of patients: chronic fatigue syndrome with or
without orthostatic postural tachycardia (Yoshiuchi et al. 2004), patients with predisposition to syncope (Lepicovska et al. 1992) or normal versus mild hypertensive patients (Novak et al. 1994).

Another field of application is the study of relationships between ANS and myocardial ischaemia, either spontaneous or induced. An early result, obtained by time-varying methods (Bianchi et al. 1993), evidenced an LF increase within the 2 min preceding the ischaemic attack. When different ischaemic events are considered, different trends were observed (Bianchi et al. 1997; Vila et al. 1997), all documenting an important involvement of ANS in either generating or sustaining the attack. An acute variation in autonomic tone preceding the attack has been recently confirmed using wavelet analysis and associated with spontaneous coronary spasm in patients with variant angina (Tan et al. 2003). During dipyridamole-induced ischaemia, the trend of spectral parameter was obtained through time-varying models and evidenced, in positive tests, an increase of the LF component in coincidence with the echocardiographic marker of ischaemia, but preceding the ST displacement on the ECG (Petrucci et al. 1996). Using wavelets analysis, this LF increase was related to the degree of pathology (being more evident in patients with triple vessel disease) and directly correlated with the wall motion score index (Petretta et al. 1999). Finally, the CWT was used for explaining the patterns of cardiac rate control during reperfusion induced by thrombolysis in MI patients (Toledo et al. 2003).

The dynamic evolution of RSA and its links to ventilation were investigated during exercise using an evolutive AR model (Meste et al. 2005), evidencing a progressive reduction in RSA modulation during the early part of the exercise (low/medium workloads) followed by a successive increase up to exercise peak (high workloads). The RSA persistence at peak exercise is probably a non-autonomic mechanism related to mechanical stretch of the sinus node, in response to ventilation increase. Dynamic behaviour of the HF component and the respiratory frequency in relation to the anaerobic threshold (AT) was investigated using the Hilbert transform (Anosov et al. 2000) demonstrating significant changes in the AT region. Interestingly, ventilatory thresholds could be assessed from the HRV data during stress by time-varying analysis of RSA (Blain et al. 2005) and during incremental exhaustive running test by the SPWV (Cottin et al. 2007).

An emerging field of application is related to the evaluation of HRV modifications during either normal or pathological sleep for the detection of apnoea events (Roche et al. 2003), for the sleep staging or for the assessment of autonomic changes during sleep (Spicuzza et al. 2003; Mendez et al. 2008). Events of apnoea (hypopnoea) are accompanied by concomitant cyclic variations in HR (a bradycardia/tachycardia pattern), which can be used to detect it (Penzel 2003). WT was proposed as a potential diagnostic screening tool in patients with sleep apnoea–hypopnoea syndrome (Hilton et al. 1999), and was proved to be very well suited to recognize the sleep apnoea-specific cyclic variability of HR (Roche et al. 2003). A comparison of different methods for sleep apnoea detection using ECG, also including time–frequency methods, can be found in Penzel et al. (2002). Recently, indices derived from WT of HRV have been proposed as the indicator of the degree of sleep fragmentation (Sforza et al. 2007). Markedly different autonomic alterations in the modulation of HR during (and after) obstructive and central apnoeas in patients with sleep-disordered
breathing have been observed using the WV transform, suggesting a surge in sympathetic modulation after the obstructive episodes (Spicuzza et al. 2003). Arousal-induced changes in the HRV have been studied using TV methods evidencing that the induced sympathoexcitatory cardiovascular effects are relatively long lasting, compared with MSNA recordings, and may accumulate if repetitive arousals occur in close succession (Blasi et al. 2003).

Other applications deal with the evaluation of sympathovagal balance before non-sustained ventricular tachycardia (Chen 2002) or the assessment of ANS during daily activity (Chan et al. 2006).

7. Discussion and conclusion

This paper presents an overview of the most common time–frequency methods for HRV analysis during non-stationary conditions. The attention was focused on the algorithms employed for the continuous quantification of the LF and HF indices, which ranged from the simple integration of power in defined frequency bands (scales) up to ad hoc decomposition methods designed for the HRV signal. A few examples were presented to emphasize the usefulness of these methods in extracting HRV indices during non-stationary situations and to stress advantages/shortcomings of the methods.

At this regard, the parametric time-varying analysis fits the pseudo-stochastic nature of the HRV signals, and allows great flexibility because time resolution (governed by the forgetting factor) and frequency resolution (depending on model order) can be tuned separately. The AR models provide smoother and narrow frequency peaks, permitting precise frequency identification and the recursive identification is useful for real-time applications. However, tracking of power is more problematic especially when the model poles are close to the unit circle because a small change in their position may result in a large spectral modification.

Computation efficiency and excellent time–frequency resolutions are among the advantages of the SPWVD methods. When these properties are combined with a decomposition method, an efficient tracking of the LF and HF parameters is obtained. The main limitation of this approach is the simplicity of the HRV model (3.7).

The appealing feature of the CWT is that time resolution depends on the frequency of interest and this allows a fast tracking of HF spectral power. However, power computation may be problematic when scales and HRV bands do not properly match or when more than one spectral component is present at a given scale. Despite these limitations, the number of HRV studies employing wavelet transform is considerably increasing in the last few years.

All the above-mentioned approaches, when correctly tuned, provide useful information about ANS dynamic, but further work is needed. A dynamic change in ANS control is usually accompanied by changes (sometimes very pronounced) in the mean HR resulting in the trends that need to be removed to emphasize the beat-to-beat variations. While this is automatically obtained using signal decomposition methods (WT and EMD), pre-processing is requested before the application of time–frequency and time-varying methods (Mainardi et al. 2002; Tarvainen et al. 2002). Another problem, which is often overlooked, is the
selection of the most appropriate frequency band for the computation of the LF and HF parameters. In traditional spectral analysis, these bands are fixed, but this choice is obviously suboptimal when dynamical tracking has to be performed. In these situations, the definition of dynamic boundaries is important and different solutions have recently been proposed (Goren et al. 2006; Bailón et al. 2007).

In conclusion, the methods described in this paper were designed for HRV analysis, but they may be of potential interest in other applicative fields where the tracking of frequency and power of the spectral components is needed.

References


*Phil. Trans. R. Soc. A* (2009)
Quantification of HRV in time–frequency domain


