A GAUSSIAN PROCESS TO DETECT UNDERDAMPED MODES OF OSCILLATION

R.S.MACKAY

Abstract. In many domains of data science, it is desired to detect modes of oscillation of a system, including estimating their frequency, damping rate, mode shape and amplitude. Here a Gaussian process solution is presented.

In memory of Professor Sir David John Cameron MacKay FRS (22 April 1967 – 14 April 2016)

1. Introduction

A dynamical system with an asymptotically stable equilibrium subject to random forcing exhibits oscillatory response if it has underdamped modes of oscillation. It may be desired to identify the mode frequencies, damping rates and shapes.

The example that motivated this paper is to detect inter-area oscillations in electricity transmission networks from phasor measurement unit data [TR+]. Figure 1 shows an example where the oscillations became obvious, but the aim is to detect the least damped modes before they become obvious, so that appropriate control can be put in place.

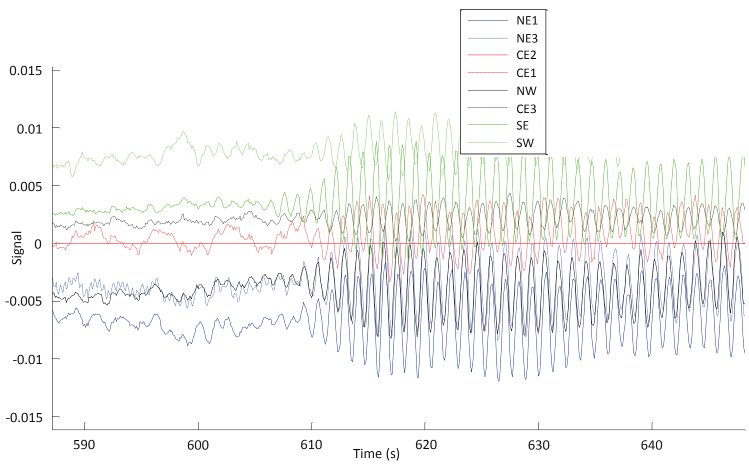

Figure 1. Voltage angles as functions of time for various locations in the GB network, relative to location CE2 (reproduced with permission from [TR+]).

Another example of motivation is to detect soft modes for civil engineering structures such as buildings and bridges, e.g. Ch.13 of [HF]. Yet others are the identification of

Date: January 5, 2018.

modes of oscillation in the sun (helioseismology) [Ko], in gene expression data [PMPR], and the business cycle (e.g. Ch.4 of [Rom]). Finally, I will suggest a potential new approach to using NMR for protein structure determination and imaging, by stochastic stimulation instead of pulses.

A standard approach to detecting oscillations is to identify peaks in the Fourier spectrum [HF]. For example, the response x of the second order system

$$m\ddot{x} + \beta\dot{x} + kx = \eta \tag{1}$$

to noise η with power spectrum P has power spectrum

$$|\hat{x}(\Omega)|^2 = \frac{P(\Omega)}{(k - m\Omega^2)^2 + \beta^2\Omega^2}$$

as a function of frequency Ω. So if the noise is white (P is constant), then the inverse quality factor $Q^{-1} = \frac{\beta}{\sqrt{mk}}$ is precisely the fullwidth at half maximum for the power spectrum $\Omega^2|\hat{x}(\Omega)|^2$ of the velocity $\dot{x}$ (its maximum is at $\Omega_{res} = \sqrt{k/m}$, known as the resonant frequency), and the damping ratio $\zeta = \frac{1}{2}Q^{-1}$ is the halfwidth at half maximum. For P slowly varying on the scale of $\frac{\beta}{\sqrt{mk}}$, the results remain good approximations. This was given a sound grounding in Bayesian analysis (see [Gre] for a survey and [B] for a pedagogical presentation), but still suffers from issues like dealing with trends, choosing windowing functions, missing data, failure to cater for slowly shifting phase, and poor theoretical justification for taking more than the largest peak if one wants to infer more than one mode of oscillation.

Wavelet transforms are popular for resolving signals in both time and frequency (up to the limits of the uncertainty principle), but I am not aware whether they can give an estimate of damping rate.

Another approach is to study the effect of an impulse (the Prony method and variants like MUSIC and ESPRIT, e.g. [PLH]), but many real-world systems may not be subjectable to impulses. For a review of these and some other methods (e.g. Hilbert transform), see [ZD].[1]

Here I present a Gaussian process method for detection of oscillations, following a suggestion of my brother David. For an introduction to Gaussian processes for time-series modelling, see [RO+]. My method allows for slowly varying equilibrium with arbitrary number of underdamped modes, and uses decaying sinusoid kernels rather than the periodic kernels of [M, D+]. It allows some overdamped modes too, because they may form the dominant part of the response, with the oscillations being a decoration on top.

After starting this work, I found the idea had just been proposed independently [PMPR]. I develop it in greater generality, however, and address particularities relevant to the case of the electricity transmission system. Another reference which expresses similar ideas is [RR+] and it has precursors under the terms "linear time invariant systems" and "latent force models", especially [HS].

[1] As yet another method, I learnt back in the mid-1980s that a good way to determine the eigenvalues of an asymptotically stable system from the response to an impulse is to Laplace transform the response numerically and then fit a Padé approximation and read off its poles.

2. Gaussian processes

I begin with a rapid review of Gaussian processes (GP).

A Gaussian process is a probability distribution for functions $F : T \to \mathbb{R}$ from a set T such that for all $n \geq 1$ the marginal density P for the values $f_1, \dots f_n = F(t_1), \dots F(t_n)$ at any finite set $t_1, \dots t_n \in T$ is Gaussian.

Examples for the set T are $\mathbb{R}$ representing time, or the set V of vertices in a graph representing spatial locations in a network, or $\mathbb{R} \times V$ for time and vertices, or $\mathbb{R} \times V \times I$ where I is a set of labels representing components of a vector of values at each vertex and time. We shall end up using $T = \mathbb{R} \times ((V \times \{1, 2\}) \cup E)$ where $\{1, 2\}$ label the frequency and power imbalance at each vertex in a graph and E is the set of edges in a spanning tree.

A basic theorem is that it follows that there is a "mean" function $M : T \to \mathbb{R}$ and a positive definite "covariance" function $C : T \times T \to \mathbb{R}$ such that

$$P(f_1, \dots f_n) = (2\pi)^{-n/2} (\det c)^{-1/2} e^{-\frac{1}{2}(f-m)^T c^{-1} (f-m)},$$

where m is the vector with components $m_i = M(t_i)$ and c is the matrix with components $c_{ij} = C(t_i, t_j)$. C positive definite means that for all $t_1, \dots t_n$, and $v_1, \dots v_n$ not all zero, then $v^T c v > 0$.

There are many introductions to Gaussian processes, e.g. [M, RW, RR, RO+, L+], and software packages, e.g. GPML.

An amusing example of a covariance function for a Gaussian process is given in Appendix B.

3. Linear stochastic systems

Suppose we are faced with an asymptotically stable linear forced system

$$\dot{x} = A(t)x + \eta(t), \tag{2}$$

with $x, \eta \in \mathbb{R}^n$, $\dot{x} = \frac{dx}{dt}$, and A possibly time-dependent. In reality the system may be nonlinear but if it has an asymptotically stable equilibrium (or more generally, an asymptotically stable slowly varying solution) and the forcing is small then it is appropriate to linearise the system about that solution. Then the response $x(t)$ to forcing $\eta(t)$ can be written as

$$x(t) = \int_{-\infty}^{t} H(t, t') \eta(t') \, dt',$$

with H the impulse response (matrix-valued Green function), i.e. the matrix solution of

$$\frac{\partial H}{\partial t} = A(t) H(t, t')$$

for $t > t'$ with $H(t'+, t') = I$. For future use, note that for any $t < t' < t''$,

$$H(t, t'') = H(t, t') H(t', t''). \tag{3}$$

If η is a Gaussian random process on $\mathbb{R} \times \{1, \dots n\}$ with zero mean and covariance function[2]

$$K_{ij}(s, t) = \langle \eta_i(s) \eta_j(t) \rangle$$

[2]Dependence on component label is indicated by subscripts.

then x is Gaussian with zero mean and covariance function

$$L_{ij}(s,t) = \langle x_i(s)x_j(t)\rangle = \int_{-\infty}^{s} ds' \int_{-\infty}^{t} dt' H_{ik}(s,s')K_{kl}(s',t')H_{jl}(t,t'),$$

using the summation convention.

If the system is autonomous (some say time-invariant) then $H(s,s')$ is a matrix-function $h(\sigma)$ of just one variable $\sigma = s - s'$. If the forcing is stationary (which means its statistics are time-invariant) then $K(s,t)$ is a matrix-function $k(\tau)$ of $\tau = t - s$ with $k(-\tau) = k(\tau)^T$. So, assuming both and changing variables to σ and $\tau' = t' - s'$,

$$L_{ij}(s,t) = l_{ij}(\tau) = \int_0^{\infty} d\sigma \int_{-\infty}^{\tau+\sigma} d\tau' h_{ik}(\sigma)k_{kl}(\tau')h_{jl}(\tau+\sigma-\tau'). \tag{4}$$

This is an example of how one Gaussian process (η) produces another (x) by convolution [RW] (also known as "blurring", e.g. [M]).

In particular, if the forcing is white, i.e.

$$k(\tau) = K\delta(\tau)$$

for some symmetric positive semi-definite[3] (psd) matrix K (reusing the notation), then for $\tau > 0$,

$$l_{ij}(\tau) = \int_0^{\infty} d\sigma h_{ik}(\sigma)K_{kl}h_{jl}(\tau+\sigma), \tag{5}$$

and for $\tau < 0$, $l(-\tau) = l(\tau)^T$. Using $h(\tau+\sigma) = h(\sigma)h(\tau)$ for $\sigma, \tau > 0$ from (3), this boils down to

$$l(\tau) = \left(\int_0^{\infty} d\sigma h(\sigma)Kh^T(\sigma)\right) h^T(\tau) \tag{6}$$

for $\tau > 0$, giving the standard result that the covariance of the response of a linear system to white noise is a matrix multiple of the transpose of the impulse response function (e.g. p.105 of [Ga]).

Note that (6) can be written as

$$l(\tau) = \Sigma h^T(\tau), \tag{7}$$

where the symmetric matrix

$$\Sigma = \int_0^{\infty} h(\sigma)Kh^T(\sigma)\, d\sigma \tag{8}$$

satisfies the Sylvester equation [Ga]

$$A\Sigma + \Sigma A^T = -K.$$

The latter has a unique solution for Σ because A has been assumed to have all its spectrum in the open left half plane, so there are no pairs of eigenvalues for A and A^T summing to zero (see the theory of Sylvester equations [BR]).

[3] K positive semi-definite means $u^TKu \geq 0$ for all vectors u.

The idea that filtering white noise produces interesting processes is old, e.g. [Ga, BF, TBT].[4] The resulting Gaussian processes are sometimes called (asymptotically stable) linear time-invariant (LTI) processes.

Fourier transform analysis requires the autonomous and stationary assumptions and often implicitly assumes white noise. We shall use these assumptions too initially, but later on we shall relax to slowly varying system parameters and noise-covariance, and other noise sources.

Note that one can treat noise sources that are themselves (linearly) filtered white noise via hidden variables, representing the state of the filter. So for now, restrict attention to the white noise case.

The one-dimensional case

$$\dot{x} = -ax + \eta,$$

with $\langle \eta(s)\eta(t)\rangle = K\delta(t-s)$ and $a, K > 0$, generates the (stationary) Ornstein-Uhlenbeck (OU) process. The impulse response (i.e. solution when $\eta(t)$ is replaced by $\delta(t)$) is $x(t) = e^{-at}$, so applying (6) the covariance function is exponential:

$$l(\tau) = \frac{K}{2a} e^{-a|\tau|}.$$

A sample is shown in Figure 2. With probability one, samples are continuous but nowhere differentiable [Ad].

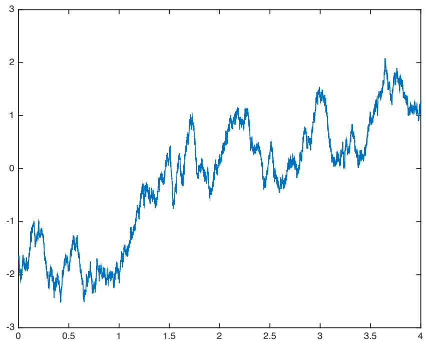

FIGURE 2. A sample from the OU process with $a = 1$, $K = 2$.

The second order case (1) with $m, \beta, k > 0$ and $mk > \beta^2/4$, generates the underdamped linear Langevin (ULL) process. Setting $\alpha = \frac{\beta}{2m}$ and $\omega = \frac{1}{m}\sqrt{mk - \beta^2/4}$, the free solutions of (1) are $x(t) = e^{-\alpha t}(A \cos \omega t + B \sin \omega t)$ for arbitrary constants A and B. Writing (1) as a system

$$\begin{aligned} \dot{x} &= v + \zeta \\ \dot{v} &= -\frac{\beta}{m} v - \frac{k}{m} x + \frac{\eta}{m}, \end{aligned} \tag{9}$$

[4]Also I took a computer music course in Princeton in 1978/9 in which we made human song by filtering white noise.

where an artificial noise source ζ has been introduced to facilitate evaluation of the impulse response matrix (see Appendix A), and applying (6), yields the covariance function for x as

$$l(\tau) = \frac{K}{2\beta k} e^{-\alpha|\tau|}(\cos\omega\tau + \frac{\alpha}{\omega}\sin\omega|\tau|).$$

Samples are shown in Figure 3 for two values of α/ω. Note that the samples are differentiable, as can be proved by applying results from [Ad]. Intuitively, it is because the noise forces the second derivative of the observable, or equivalently $l'(0) = 0$. If one wishes to identify the resonant frequency Ω_{res} and the damping ratio ζ from ω and α then use $\Omega_{res}^2 = \frac{k}{m} = \omega^2 + \alpha^2$ and $\zeta = \frac{\alpha}{\Omega_{res}}$.

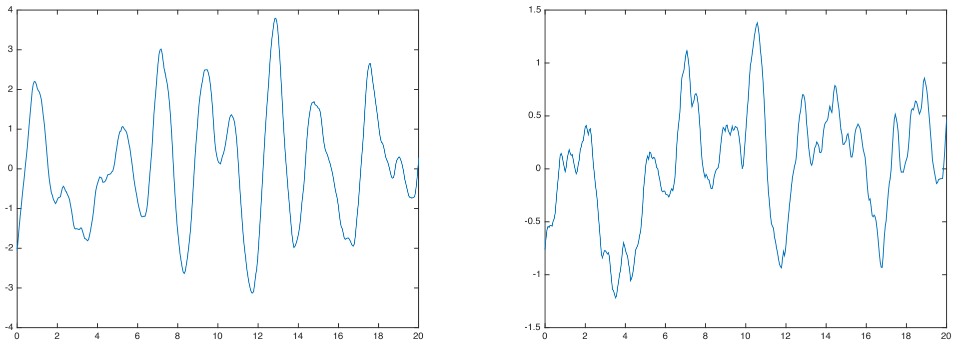

FIGURE 3. Samples from the ULL process for $\frac{K}{2\beta k} = 1$ with (a) $\alpha = e^{-1}, \omega = e$, (b) $\alpha = e, \omega = e^{-1}$.

It is also of interest to look at the process for the velocity $\dot{x}$ of the ULL, which I call the "velocity of ULL" process (VULL). Its covariance function (Appendix A) is

$$l(\tau) = \frac{K}{2m\beta} e^{-\alpha|\tau|}(\cos\omega\tau - \frac{\alpha}{\omega}\sin\omega|\tau|).$$

Samples are shown in Figure 4 for two values of α/ω. The samples are continuous but nowhere differentiable (apply [Ad]).

One can also consider the full 2D system, for which the covariance function is a 2×2 matrix function (32) of time difference, as derived in Appendix A.

The ULL and VULL are extreme cases of one-dimensional observation $y(t) = ax(t) + b\dot{x}(t)$ (for constants a and b) of the underdamped 2D system, which has covariance function proportional to $e^{-\alpha|\tau|}(\cos\omega\tau + A\sin\omega|\tau|)$ for some $A \in [-\alpha/\omega, +\alpha/\omega]$ [XXX perhaps give it explicitly?].

An alternative way to see this, that will be useful for higher dimensions, is that a general underdamped 2D system can be written in the real Jordan normal form

$$\dot{x} = B \begin{bmatrix} -\alpha & -\omega \\ \omega & -\alpha \end{bmatrix} C^T x + \eta \tag{10}$$

for two-dimensional vectors x, η, invertible 2×2 matrix B, $C^T = B^{-1}$ (the transpose is an electrical engineering convention) and $\langle \eta_i(s)\eta_j(t)\rangle = K_{ij}\delta(t-s)$ for some psd symmetric

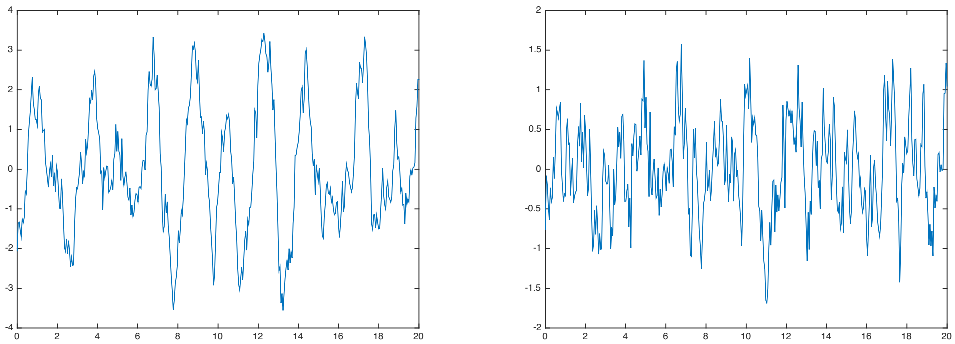

FIGURE 4. Samples from the VULL process for $\frac{K}{2m\beta} = 1$ with (a) $\alpha = e^{-1}, \omega = e$, (b) $\alpha = e, \omega = e^{-1}$ (e is the base of natural logarithms).

matrix K. Writing $y = C^T x$ then

$$\dot{y} = \begin{bmatrix} -\alpha & -\omega \\ \omega & -\alpha \end{bmatrix} y + \tilde{\eta} \tag{11}$$

with $\tilde{\eta} = C^T \eta$, which has covariance matrix $\tilde{K} = C^T K C$. The impulse response is $y(t) = e^{-\alpha t} R_{\omega t}$, where

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{12}$$

So the covariance matrix function for y for $\tau \geq 0$ is

$$l(\tau) = \left(\int_0^\infty e^{-\alpha\sigma} R_{\omega\sigma} \tilde{K} e^{-\alpha\sigma} R_{\omega\sigma}^T \, d\sigma \right) e^{-\alpha\tau} R_{\omega\tau}^T.$$

Write $R_\theta = \cos\theta I + \sin\theta J$ with

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{13}$$

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{14}$$

Then for $\tau \geq 0$,

$$l(\tau) = \frac{1}{4\alpha(\alpha^2 + \omega^2)} e^{-\alpha\tau} (R \cos\omega\tau + S \sin\omega\tau), \tag{15}$$

with matrices [XXX check signs]

$$R = (2\alpha^2 + \omega^2)\tilde{K} + \alpha\omega(J\tilde{K} - \tilde{K}J) - \omega^2 J\tilde{K}J \tag{16}$$

$$S = -\alpha\omega(\tilde{K} + J\tilde{K}J) - \omega^2 J\tilde{K} - (2\alpha^2 + \omega^2)\tilde{K}J. \tag{17}$$

The covariance matrix function for x is $Bl(\tau)B^T$ so has a similar expression.

The covariance function for one component, say y_1, is

$$l_{11}(\tau) = \frac{1}{4\alpha(\alpha^2 + \omega^2)} e^{-\alpha\tau} (R_{11} \cos\omega\tau + S_{11} \sin\omega\tau)$$

with

$$R_{11} = (2\alpha^2+\omega^2)\tilde{K}_{11} - 2\alpha\omega\tilde{K}_{12} - 2\omega^2\tilde{K}_{22}$$
$$S_{11} = \alpha\omega(\tilde{K}_{22}-\tilde{K}_{11}) - 2\alpha^2\tilde{K}_{12}.$$

It follows that for α,ω fixed, the ratio S_{11}/R_{11} goes monotonically from $-\alpha/\omega$ to $+\alpha/\omega$ as

$$\frac{(\omega\tilde{K}_{12}-\alpha\tilde{K}_{11})^2 + (\tilde{K}_{11}\tilde{K}_{22}-\tilde{K}_{12}^2)\omega^2}{\tilde{K}_{11}^2}$$

goes from 0 to infinity. Note that K psd implies that this expression is never negative. The case $-\alpha/\omega$ is the VULL and the case $+\alpha/\omega$ is the ULL. For rotationally invariant $\tilde{K}$ (i.e. a multiple of the identity) then $S_{11}=0$ and the resulting covariance function was called OUosc by [PMPR] and "exponentially damped cosine" by [Ab]. Note that $S_{11}/R_{11} > \alpha/\omega$ does not give a valid covariance function because a covariance function always satisfies $l_{11}(\tau) \le l_{11}(0)$ for all τ. Similarly, $S_{11}/R_{11} < -\alpha/\omega$ is not valid because a covariance function always satisfies $\int l_{11}(\tau)\, d\tau \ge 0$. Thus if one is observing only one component, to keep these constraints it is a good idea to write $S_{11} = \frac{\alpha}{\omega}R_{11}\tanh r$ and use $r \in \mathbb{R}\cup\{\pm\infty\}$ as parameter instead of S_{11}. I call the GP with covariance function

$$l(\tau) = \sigma^2 e^{-\alpha|\tau|}(\cos\omega\tau + \frac{\alpha}{\omega}\tanh r \sin\omega|\tau|)$$

the generalised ULL (GULL). The same form of covariance function is obtained for any component, $a_1y_1+a_2y_2$ (or equivalent for x), but with σ^2 and r depending on (a_1,a_2).

Now proceed to the case of arbitrary dimension N. Suppose that the eigenvalues of A are distinct. Then A can be diagonalised by change of coordinates. If there are complex eigenvalues one needs complex coordinates, but the alternative used here is to block-diagonalise with a 2×2 block for each complex-conjugate pair of eigenvalues and restrict to real coordinate changes (as above). Thus

$$A = B\Lambda C^T$$

with Λ block-diagonal, having a 1×1 block $-\lambda_m$ for each real eigenvalue (the notation has a minus sign because asymptotic stability requires them to be negative) and a 2×2 block $\begin{bmatrix} -\alpha_m & -\omega_m \\ \omega_m & -\alpha_m \end{bmatrix}$ with $\alpha_m,\omega_m > 0$ for each complex conjugate pair $-\alpha_m \pm i\omega_m$, and $C^T = B^{-1}$. The columns of B (taken in pairs for complex eigenvalues) are the mode shapes.

Then the impulse response can be written

$$h(\sigma) = BM(\sigma)C^T,$$

with M block-diagonal, having 1×1 blocks $e^{-\lambda_m\sigma}$ and 2×2 blocks $e^{-\alpha_m\sigma}R_{\omega_m\sigma}$.

It follows that the covariance of the response is (for $\tau > 0$)

$$l_{ij}(\tau) = \sum_{m,n} B_{im}\int_0^\infty M_m(\sigma)D_{mn}M_n^T(\sigma)\, d\sigma\; M_n^T(\tau)B_{jn}, \tag{18}$$

where M_m denotes the diagonal blocks of M, $D_{mn} = C_{km}K_{kl}C_{ln}$ (with summation convention), and the sum over modes m,n has been made explicit. For a complex mode

m, B_{im} is a pair of columns, as already mentioned. The elements D_{mn} are 1×1, 1×2, 2×1 and 2×2 blocks according as m, n are real or complex modes. We write

$$P_{mn} = \int_0^\infty M_m(\sigma) D_{mn} M_n^T(\sigma)\ d\sigma,$$

so

$$l_{ij}(\tau) = \sum_{m,n} B_{im} P_{mn} M_n^T(\tau) B_{jn}. \tag{19}$$

It remains to evaluate the integrals P_{mn}. In the 1×1 case,

$$\int_0^\infty e^{-\lambda_m \sigma} D_{mn} e^{-\lambda_n \sigma}\ d\sigma = \frac{D_{mn}}{\lambda_m + \lambda_n}.$$

In the 1×2 case,

$$\int_0^\infty e^{-\lambda_m \sigma} D_{mn} e^{-\alpha_n \sigma} R^T_{\omega_n \sigma}\ d\sigma = \frac{1}{(\lambda_m + \alpha_n)^2 + \omega_n^2} D_{mn} \begin{bmatrix} \lambda_m + \alpha_n & \omega_n \\ -\omega_n & \lambda_m + \alpha_n \end{bmatrix}.$$

In the 2×1 case,

$$\int_0^\infty e^{-\alpha_m \sigma} R_{\omega_m \sigma} D_{mn} e^{-\lambda_n \sigma} d\sigma = \frac{1}{(\alpha_m + \lambda_n)^2 + \omega_m^2} \begin{bmatrix} \alpha_m + \lambda_n & -\omega_m \\ \omega_m & \alpha_m + \lambda_n \end{bmatrix} D_{mn}.$$

In the 2×2 case, write $D_{mn} = \begin{bmatrix} N+Q & E+F \\ E-F & N-Q \end{bmatrix}$. Then the integral boils down to [XXX CHECK SIGNS!]

$$\begin{aligned} P &= \tfrac{N}{(\alpha_m+\alpha_n)^2+(\omega_m+\omega_n)^2} \begin{bmatrix} \alpha_m + \alpha_n & -\omega_m - \omega_n \\ \omega_m + \omega_n & \alpha_m + \alpha_n \end{bmatrix} \\ &+ \tfrac{Q}{(\alpha_m+\alpha_n)^2+(\omega_m-\omega_n)^2} \begin{bmatrix} \alpha_m + \alpha_n & \omega_n - \omega_m \\ \omega_m - \omega_n & \alpha_m + \alpha_n \end{bmatrix} \\ &+ \tfrac{E}{(\alpha_m+\alpha_n)^2+(\omega_m-\omega_n)^2} \begin{bmatrix} \omega_n - \omega_n & \alpha_m + \alpha_n \\ \alpha_m + \alpha_n & \omega_m - \omega_n \end{bmatrix} \\ &+ \tfrac{F}{(\alpha_m+\alpha_n)^2+(\omega_m+\omega_n)^2} \begin{bmatrix} \omega_m + \omega_n & \alpha_m + \alpha_n \\ -\alpha_m - \alpha_n & \omega_m + \omega_n \end{bmatrix}. \end{aligned}$$

Note that in the case $m = n$ then D_{mn} is symmetric psd so $F = 0$, $N \geq \sqrt{Q^2 + E^2}$ and

$$P = \frac{N}{2(\alpha^2 + \omega^2)} \begin{bmatrix} \alpha & -\omega \\ \omega & \alpha \end{bmatrix} + \frac{1}{2\alpha} \begin{bmatrix} Q & E \\ E & Q \end{bmatrix}.$$

This reproduces the result (15) for the 2D case studied earlier.

In practice, to fit data to a GP with covariance function (19) it is not sensible to seek to fit all modes. Thus a smaller number $K < N$ of modes can be targeted (count each complex mode as 2) by making B an $N \times K$ matrix and P a $K \times K$ matrix.

Also, the parameters of a system and the noise source may change slowly over time. If one wants to model data over a long time then one should allow these parameters to be themselves a random process [XXX but is the result Gaussian?? superstatistics?]. For many purposes, however, it is enough to analyse a time interval long enough to capture the oscillations and short enough that the underlying parameters do not change much. More important is to design the analysis to work on streaming data rather than batch

[XXX [RR] could be useful here, and [RR+] and its references e.g. [HS]? probably need to make a separate section on this!].

4. Application to AC electricity networks

The dynamics of an AC (alternating current) electricity network can be modelled approximately by a connected graph with a node for each rotating machine (synchronous generator or motor) [MBB] (this leaves open the question of how to model DC/AC convertors, such as at wind farms, solar photovoltaic farms and DC interconnected terminals). Let N be the number of nodes. As described in [Rog] (another useful reference is [An]), one can model an AC network at various levels of complexity. If one ignores aspects like the dynamics of the voltages[5], 3-phase imbalances, reactive power control and harmonics, the state can be specified by a phase ϕ_l and frequency $f_l = \dot{\phi}_l$ (as ϕ_l is in radians it might be better to denote f_l by ω_l, but I am already using ω for mode frequencies) at each node l, and dynamics for the vector f of frequencies and phases ϕ given by balancing power (cf. (1) of [SMH] or (17) of [SM]):

$$
\begin{aligned}
I_l f_l \dot{f}_l &= p_l - \Gamma_l f_l^2 - \sum_{l'} V_l V_{l'} \left(B_{ll'} \sin(\phi_l - \phi_{l'}) + G_{ll'} \cos(\phi_l - \phi_{l'}) \right) \\
\dot{\phi}_l &= f_l
\end{aligned}
\tag{20}
$$

where I_l is an inertia, Γ_l a damping constant, V_l is the amplitude of the voltage at l, $B_{ll'}$ is a symmetric matrix of ideal admittances of the line between l and l' ($B_{ll} = 0$), $G_{ll'}$ is a symmetric psd matrix of conductances of the line between l and l' (which produces transmission losses), including self-conductances, and p is a vector of power imbalances (generation minus consumption), which is to be regarded as an external stochastic process (e.g. people switching loads on and off, wind farms producing varying power). For the moment, think of p as fixed.

The system has the special feature of global phase rotation invariance: if one adds the same constant to all the phases then the dynamics produces the same trajectory but with the constant added. One can quotient by this symmetry group, which we denote by S.[6] For example, choose a root node o and a spanning tree in the graph, orient its edges e away from o (other choices are alright but this is to make a definite choice), and let $\Delta_e = \phi_{l'} - \phi_l$ for each edge $e = ll'$ in the spanning tree (there are $N-1$ of these, and we denote the vector of phase differences by Δ. Then the phase difference between any two nodes can be expressed as a signed sum of the Δ_e, and the equations $\dot{\phi}_l = f_l$ can be replaced by $\dot{\Delta}_e = f_{l'} - f_l$.

The quotient system has a manifold of equilibria in the space of all power imbalance vectors p, frequency vectors f and phase difference vectors Δ. For an equilibrium (mod S), each node has the same frequency and the phase differences are constant. The manifold of equilibria is a graph over the common frequency $F \in \mathbb{R}$ and the phase

---

[5] this is relatively easy to incorporate, e.g. [TBP], but a full treatment would require including voltage control, power system stabilisers, and excitor control

[6] In reality, the system operator is required to keep the phases within some interval (of about 100 cycles) around that for a reference rotor at the nominal frequency, so they exert changes to p to achieve this, thereby breaking the phase rotation invariance, but we will ignore that.

differences $\Delta \in (\mathbb{R}/2\pi\mathbb{Z})^{N-1}$:

$$p_l = \Gamma_l F^2 + \sum_{l'} V_l V_{l'} \left(B_{ll'} \sin(\phi_l - \phi_{l'}) + G_{ll'} \cos(\phi_l - \phi_{l'}) \right).$$

Let us restrict attention to the part with F near a nominal reference frequency F_0 (50Hz in Europe, which means $F_0 = 100\pi$ in radians/sec). Then the projection to the space of power imbalances p covers 2^{N-1} times a region around $p = 0$ [XXX this is not always true]. Of these 2^{N-1} equilibria, precisely one is stable, roughly speaking the one with all phase differences between linked nodes being less than $\pi/2$ in absolute value. This can be established by the energy method used in [TBP], modified to include the conductance matrix G and ignore the voltage dynamics [XXX DO IT]. It should noted, however, that inclusion of governors or power system stabilisers in the model can destabilise the equilibrium and produce inter-area oscillations [Rog], presumably by a Hopf bifurcation. The method of the present paper is not well adapted to detecting autonomous oscillations as opposed to damped ones forced by noise.

Suppose the system is near the stable equilibrium for some p. As p moves in time, the response roughly speaking follows it on the manifold of equilibria, but deviations from equilibrium are in general excited and these would relax back to equilibrium if p were to stop moving. For small movements of p about a mean imbalance vector P with corresponding stable equilibrium (F, Δ), it is appropriate to linearise the system. A reference for small-signal stability in power systems is [GPV]. Write δf_l, $\delta\Delta_e$, δp_l for the deviations of f_l, Δ_e and p_l from the equilibrium. Write

$$M_l = I_l F, \quad \gamma_l = 2\Gamma_l F,$$

$$T_{ll'} = V_l V_{l'} (B_{ll'} \cos(\Phi_l - \Phi_{l'}) - G_{ll'} \sin(\Phi_l - \Phi_{l'})).$$

Then

$$\begin{aligned} M_l \dot{\delta f}_l &= \delta p_l - \gamma_l \delta f_l - \sum_{l'} T_{ll'} (\delta\phi_l - \delta\phi_{l'}) \\ \dot{\delta\Delta}_e &= \delta f_{l'} - \delta f_l \end{aligned} \tag{21}$$

for $e = ll'$. Write this as

$$\dot{x} = Ax + C\delta p \tag{22}$$

with $x = \begin{bmatrix} \delta f \\ \delta\Delta \end{bmatrix}$ and $C = \begin{bmatrix} \mathrm{diag} M_l^{-1} \\ 0 \end{bmatrix}$.

Let us model the dynamics of the power imbalances by

$$\dot{\delta p} = -J\delta p + \sigma\xi \tag{23}$$

for some matrix J (with $-J$ asymptotically stable) and (multidimensional) white noise $\sigma\xi$ with covariance matrix $K = \sigma\sigma^T$ (later, J, P, T and K may vary slowly in time). This is a somewhat crude representation, but captures the idea that p has random increments and reversion to a mean. There is evidence that load distribution is close to Gaussian, e.g. fig.14 of [TT+], which is consistent with this model, though that data says nothing about the temporal correlations. One might argue that National Grid's balancing actions are based more on the average frequency and phase differences than the power imbalances, but on the manifold of equilibria these are equivalent.

The resulting system (22, 23) for $(x, \delta p)$ is of the form (2), but it has a skew-product structure that we should exploit, namely $\dot{\delta p}$ does not depend on x (also the x-dynamics has structure in that it is only the frequencies that see δp directly). In reality, perhaps $\dot{\delta p}$ does depend a little on x, e.g. National Grid balancing operations and frequency-sensitive generators and loads, but let us continue with this model. One way to exploit the skew-product structure is to derive the covariance function for δp using (6) and then insert this into the formula (4) for the covariance function of x, but it leads to an integration whose treatment is not simple. Alternatively, we can apply (6) to the joint system (22, 23), exploit the skew-product form of the impulse response, and take the xx-block of the covariance function. I choose the latter approach, subject to the simplifying but generic assumption of simple eigenvalues for the full system.

The impulse response of (23) can be written in matrix exponential notation as $\delta p(t) = e^{-Jt}$. Similarly, the impulse response of (22) can be written as $x(t) = e^{At}$. To compute the response of x to an impulse on $\dot{p}$, it is convenient to assume that A and $-J$ have no eigenvalues in common, as is generically the case. Then there exists a unique solution E to another Sylvester equation

$$AE + EJ = C, \tag{24}$$

and defining $y = x + Ep$ we see that $\dot{y} = Ay + E\xi$. So the response of y to an impulse on $\dot{p}$ is $e^{At}E$. It follows that the response of $x = y - Ep$ to an impulse on $\dot{p}$ is

$$x(t) = h_{xp}(t) := e^{At}E - Ee^{-Jt}.$$

Note that using (24), the time-derivative of h_{xp} at $t = 0$ is just C (this is one place where the matrix exponential notation helps). Thus the impulse response of the full system has the block form

$$h(t) = \begin{bmatrix} e^{-Jt} & 0 \\ e^{At}E - Ee^{-Jt} & e^{At} \end{bmatrix} \tag{25}$$

Then the stationary covariance matrix Σ (8) of the joint process has the block form

$$\Sigma = \int_0^\infty h(\sigma) \begin{bmatrix} K & 0 \\ 0 & 0 \end{bmatrix} h^T(\sigma)\, d\sigma = \int_0^\infty \begin{bmatrix} e^{-J\sigma}Ke^{-J^T\sigma} & e^{-J\sigma}Kh_{xp}^T(\sigma) \\ h_{xp}(\sigma)Ke^{-J^T\sigma} & h_{xp}(\sigma)Kh_{xp}^T(\sigma) \end{bmatrix} d\sigma. \tag{26}$$

It follows from (6) that (for $\tau > 0$)

$$\begin{aligned} l_{xx}(\tau) &= \Sigma_{xp}h_{xp}^T(\tau) + \Sigma_{xx}e^{A^T\tau} \\ &= \left(\int_0^\infty h_{xp}(\sigma)KE^Te^{A^T\sigma}d\sigma\right) e^{A^T\tau} - \left(\int_0^\infty h_{xp}(\sigma)Ke^{-J^T\sigma}d\sigma\right) e^{-J^T\tau}E^T. \end{aligned} \tag{27}$$

Thus the covariance of $x = (\delta f, \delta\Delta)$ is a linear combination of functions from the impulse response of x to $\dot{x}$ and of p to $\dot{p}$.[7]

So now we could try to fit observations of (f, Δ) at as many locations as available (say, M) and as a function of time t to an autonomous GP with mean function of the form $(F\mathbf{1}, \bar{\Delta})$ for some $F \in \mathbb{R}$ and $\bar{\Delta} \in \mathbb{R}^{M-1}$ and covariance function of the form (18).

[7]In the case of common eigenvalues λ to $-J$ and A there would in general also be terms of the form $P(\tau)e^{\lambda\tau}$ with P a polynomial of degree higher than those which might already result from multiplicity in $-J$ or A.

We make the obvious step of shrinking the spanning tree to one for just the observed nodes.

So the proposal is to fit an autonomous GP with mean function $(F, \bar{\Delta})$ and covariance function of the form (19) to observations (f_p, Δ_e) as functions of time t, but with B truncated to having only a small number of columns. There is the question of how many modes to allow. This can be decided by an automatic Bayesian method, but one should expect the basic behaviour to be an OU process for f_p.

Indeed, using GPML, I found that a 2-hour trace of frequency at 1-second intervals, Figure 5, made publicly available by National Grid [NG], fit reasonably well to an OU process with a decay time of about 30 minutes and amplitude 0.045Hz. The time constant is so long compared to the period (about 2 seconds) or decay time (about 20 seconds) of inter-area oscillations that it is hardly relevant, and one could just say that the basic behaviour of f_p is a Wiener process.

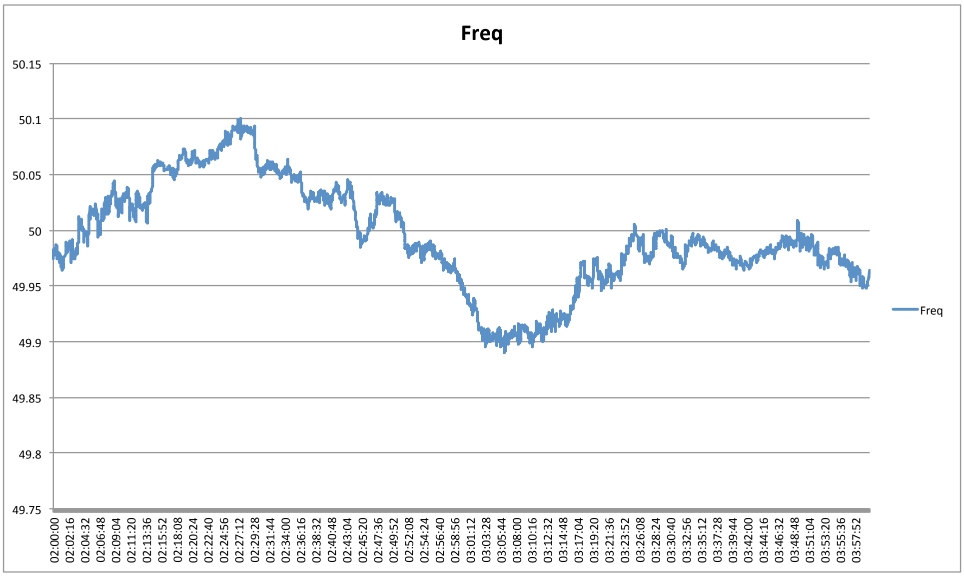

FIGURE 5. A frequency trace over 2 hours from National Grid.

On shorter timescales, however, the data look smooth (Figure 6). A simple model is a first-order filtered OU process (FOU). To justify this, imagine the system is aggregated to a single node. Then we have two equations of the form

$$
\begin{aligned}
M\dot{\delta f} &= -\gamma \delta f + \delta p \\
\dot{\delta p} &= -J\delta p + \sigma \xi
\end{aligned}
\tag{28}
$$

It follows from the second equation that δp is OU with covariance function $k(\tau) = \frac{\sigma^2}{2J} e^{-J|\tau|}$. Then applying (4) we see that δf is a GP with covariance function

$$
C(\tau) = \int_0^\infty ds \int_{-\infty}^{\tau+s} d\tau' h(s) k(\tau') h(\tau + s - \tau'),
$$

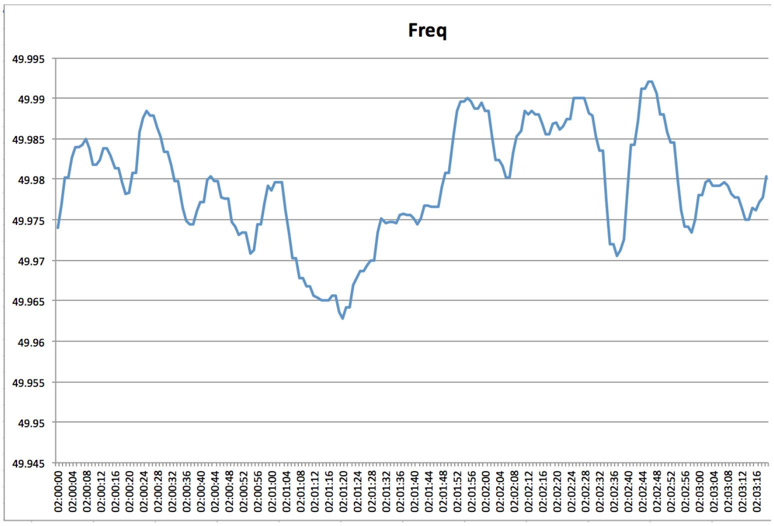

FIGURE 6. The first 3 minutes 20 seconds of the frequency trace.

where h is the impulse response for the first equation, viz. $h(s) = \frac{1}{M}e^{-\Gamma s}$ for $s > 0$, with $\Gamma = \gamma/M$. Computation of the integral (for the generic case $\Gamma \neq J$) yields

$$C(\tau) = \frac{\sigma^2}{2JM\gamma(\Gamma^2 - J^2)}(\Gamma e^{-J|\tau|} - Je^{-\Gamma|\tau|}).$$

A sample from the FOU process is shown in Figure 7.

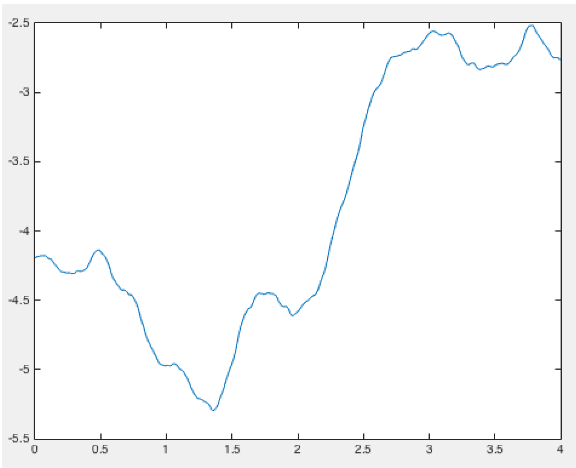

FIGURE 7. A sample from the filtered OU process for $\Gamma = 1/e$, $J = e^2$

Fitting an FOU to the first 3mins 20 secs of the data yields time constants $1/\Gamma$ and $1/J$ around 1.5mins and 2.2 secs, though one can not say which is which.

XXX Distinguishing x modes from p modes!: Timescales??

Note that the same covariance function arises for the overdamped linear Langevin process, with $-\Gamma$ and $-J$ being the two real eigenvalues.

XXX What do we expect for the basic behaviour of Δ_e? (cf. fig in [TR+])

XXX And then what would inter-area oscillations look like?

XXX Then allow the mean to vary slowly in time.

5. Removing redundancy

The specification of our covariance function (19) via a mode shape matrix B and a noise matrix P carries some redundancy: one can multiply each mode shape vector by an arbitrary non-zero constant and rotate the complex mode shape vectors by an arbitrary angle and carry out the inverse operations on P. To infer mode shape vectors from data it is better to remove this redundancy.

One way is to use mode projectors. These are linear operators P of rank 1 or 2 (for real or complex modes, respectively) such that $P^2 = P$ and the image of a vector under P is its component in the given mode.

XXX Work this out more.

6. Inferring parameters on a manifold

The spaces of projectors of given rank are manifolds, called Grassmannians. They can not be covered by a single coordinate system of the form of $\mathbb{R}^n$, so standard methods to infer parameters, such as used by GPML, can not be applied directly. Extension of methods to determine maximum likelihood points on arbitrary manifolds is required.

XXX Work it out.

7. Discussion

We have presented a method to detect oscillations in systems with many components.

One defect of the approach is that the forcing might not be Gaussian. For example, even a Poisson process with independent Gaussian amplitude is not Gaussian. Indeed, a consequence of the Gaussian assumption is that the covariance of the response is time-symmetric, whereas this may not be true for real systems. Evidence for Gaussian distribution of load is given in Fig.14 of [TT+], but they do not report on time-correlation. Load variations are likely to be independent, however, which would make them Gaussian and white. On the other hand, wind power is unlikely to be delta-correlated. There is considerable research on the statistics of wind power, e.g. [DPP].

Another defect of the approach is that it does not allow for nonlinearity. Nevertheless, for small fluctuations around an equilibrium, linearising is a good approach. It will fail to give a good approximation, however, if the eigenvalues of any mode approach or cross the imaginary axis. A big question with power flow oscillations, gene expression and business cycles is whether there is a limit cycle of some underlying deterministic dynamics, or just lightly damped oscillations around an equilibrium forced by noise. Figure 1 suggests to me that there was a Hopf bifurcation, but the assumption in the power system community is that it was just a large kick that set off a lightly damped mode of oscillation. For gene expression this has been addressed by [D+]. For business cycles, most economists decided long ago that they are just a near unit root process (meaning lightly damped oscillations forced by shocks) [Rom], though Grandmont proposed deterministic models with a variety of forms of dynamics [Gra].

To detect periodic components, my brother David [M] proposed the family of stationary covariance functions of the form

$$k(t) = \sigma^2 \exp\left(-\frac{2\sin^2(\omega t/2)}{\lambda^2}\right),$$

for which samples are exactly periodic with period $2\pi/\omega$. A slight modification was used in [L+] to remove the effect of its non-zero mean, namely

$$k(t) = \sigma^2 \frac{\exp(\lambda^{-2}\cos\omega t) - I_0(\lambda^{-2})}{\exp(\lambda^{-2}) - I_0(\lambda^{-2})},$$

where I_0 is a modified Bessel function of the first kind. It has the limiting form

$$k(t) = \sigma^2 \cos(\omega t)$$

as $\lambda \to \infty$, called the Cos kernel, which has the property that it forces anti-periodicity with anti-period π/ω: $f(t + \pi/\omega) = -f(t)$. Although these have found valuable uses, and can be made less rigid by multiplication by a decaying kernel such as $\exp(-\alpha|t|)$ (which with the Cos kernel produces OUosc), it seems to me highly preferable to start from the point of view of a linear system forced by noise.

XXX Real-time updating and connection to Kalman filter, cf [RR].

XXX NMR proposal

Acknowledgements

I am grateful to Ben Marshall of National Grid for proposing the problem of detecting inter-area oscillations, and to Phillip Ashton and him for helpful discussions on the topic and pointers to the literature; to MSc student Tajhame Francis for initial investigations by spectral analysis; to my brother David for telling me to "Use a Gaussian process"; to PhD student Marcos Tello Fraile and postdoc Lisa Flatley for trying to follow my suggestions; to Hannes Nickisch and Colm Connaughton for helping me implement my resulting solutions in GPML; to Rasmussen and Nickisch for creating GPML; to Zoubin Ghahramani for answering some questions about GPs; to Igor Mezic and Yoshihiko Susuki for discussions on modelling AC networks; and to undergraduate summer project student John Prater for coding up the 2×2 ULL for GPML. The beginning of the work was supported by National Grid under a Network Innovation Allowance award. The later parts were supported by the Alan Turing Institute under award TU/B/000101.

Appendix A: Covariance for Underdamped linear Langevin process

Starting from the system (9), an impulse $\eta/m = \delta(t)$ produces response

$$\begin{aligned} x(t) &= \frac{1}{\omega} e^{-\alpha t} \sin\omega t \\ v(t) &= \frac{1}{\omega} e^{-\alpha t}(\omega\cos\omega t - \alpha\sin\omega t) \end{aligned} \tag{29}$$

An impulse $\zeta = \delta(t)$ produces

$$x(t) = e^{-\alpha t}(\cos\omega t + \frac{\alpha}{\omega}\sin\omega t) \tag{30}$$
$$v(t) = -e^{-\alpha t}(\omega + \alpha^2/\omega)\sin\omega t$$

Thus the impulse response matrix (with components in order x, v) is

$$h(t) = e^{-\alpha t}\begin{bmatrix} \cos\omega t + \frac{\alpha}{\omega}\sin\omega t & \frac{1}{\omega}\sin\omega t \\ -(\omega + \frac{\alpha^2}{\omega})\sin\omega t & \cos\omega t - \frac{\alpha}{\omega}\sin\omega t \end{bmatrix}. \tag{31}$$

The covariance matrix for the noise is

$$K = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\sigma^2}{m^2} \end{bmatrix}$$

($\zeta = 0$). So applying (8) we obtain

$$\Sigma = \int_0^\infty h(s)Kh^T(s)\ ds = \frac{\sigma^2}{4m^2\alpha}\begin{bmatrix} \frac{1}{\alpha^2+\omega^2} & 0 \\ 0 & 1 \end{bmatrix}.$$

So the end result from (7) is (for $\tau > 0$)

$$C(\tau) = \frac{\sigma^2}{4m^2\alpha}e^{-\alpha\tau}\begin{bmatrix} \frac{1}{\alpha^2+\omega^2}(\cos\omega\tau + \frac{\alpha}{\omega}\sin\omega\tau) & -\frac{1}{\omega}\sin\omega\tau \\ \frac{1}{\omega}\sin\omega\tau & \cos\omega\tau - \frac{\alpha}{\omega}\sin\omega\tau \end{bmatrix}. \tag{32}$$

One can use $2m\alpha = \beta$ and $\alpha^2 + \omega^2 = k/m$ to simplify this in terms of the original parameters if preferred.

Thus in particular, the xx-component of $C(\tau)$ is $\frac{\sigma^2}{2\beta k}e^{-\alpha\tau}(\cos\omega\tau + \frac{\alpha}{\omega}\sin\omega\tau)$, as claimed for the ULL. Similarly, the vv-component gives the covariance for the VULL.

Appendix B: Riemann's ξ-function

An amusing remark is that Riemann's ξ-function [Ri] is the covariance function for a stationary Gaussian process. This is because it is even and the Fourier transform of a positive function:

$$\xi(t) = 2\int_0^\infty \Phi(\omega)\cos\omega t\ d\omega,$$

where $\Phi(\omega) = \sum_{n\geq 1}(4\pi^2 n^4 e^{9\omega/2} - 6\pi n^2 e^{5\omega/2})e^{-\pi n^2 e^{2\omega}}$. It generates infinitely smooth samples, much like the "squared exponential" covariance,[8] but the pseudo-random oscillatory decay of $\xi(t)$ as $t \to \infty$ might be useful in some contexts.

References

[Ab] Abrahamsen P, A review of Gaussian random fields and correlations functions (2nd ed, 1997)
[Ad] Adler RJ, The geometry of random fields (Wiley, 1981).
[An] Andersson G, Power system analysis (ETH lecture notes, 2012)
[BR] Bhatia R, Rosenthal P, How and why to solve the operator equation $AX - XB = Y$, Bull Lond Math Soc 29 (1997) 1–21.
[BF] Boyle P, Frean M, Dependent Gaussian Processes, Adv Neural Info Processing Sys (2005)
[B] Bretthorst GL, Bayesian spectrum analysis and parameter estimation, Lect Note Stats 48 (Springer, 1988).

[8] which would be better called the exponentiated square covariance

[BZA] Brincker R, Zhang L, Andersen P, Output-only modal analysis by frequency domain decomposition, Proc ISMA25 (2000), vol.2, ...

[DPP] D'Amico G, Petroni F, Prattico F, Wind speed prediction for wind farm applications by extreme value theory and copulas, J Wind Eng Ind Aerodyn 145 (2015) 229–36.

[D+] Durrande N, Hensman J, Rattray M, Lawrence ND, Detecting periodicities with Gaussian processes, arXiv:1303.7090v2 (2016)

[Ga] Gardiner CW, Stochastic Methods (Springer, 1985, 2009)

[GPV] Gibbard MJ, Pourbeik P, Vowles DJ, Small-signal stability, control and dynamics performance of power systems (Univ Adelaide Press, 2015)

[GS] Grainger JJ, Stevenson WD, Power system analysis (McGraw-Hill, 1994).

[Gra] Grandmont J-M, On endogeneous competitive business cycles, Econometrica 53 (1985) 995–1045.

[Gre] Gregory PC, A Bayesian revolution in spectral analysis, in: Bayesian inference and maximum entropy methods in Science and Engineering, ed Mohammad-Djafari A, Am Inst Phys Conf Proc 568 (2001) 557–...

[H+] Han S, Rong N, Sun T, Zhang J, An approach for estimating mode shape for participation of inter-area oscillation mode, IEEE Internat Symp on Circuits & Systems (ISCAS), (2013)...

[HS] Hartikainen J, Särkkä S, Kalman filtering and smoothing solutions to temporal Gaussian process regression models, Proc IEEE Internat Workshop on Machine Learning for Signal Processing 2010

[HF] He J, Fu Z-F, Modal analysis (Butterworth-Heinemann, 2001).

[Ko] Kosovichev AG, Solar oscillations, in Stellar Pulsation, AIP Conf Proc 1170 (2009) 547–559.

[L+] Lloyd JR, Duvenaud D, Grosse R, Tenenbaum JB, Ghahramani Z, Automatic construction and natural-language description of nonparametric regression models, arXiv:1402.4304 (2014)

[MBB] Machowski J, Bialek JW, Bumby JR, Power system dynamics (second ed, Wiley, 2008)

[M] MacKay DJC, Introduction to Gaussian Processes, NATO ASI series F Comp Sys Sci 168 (1998) 133–66.

[NG] http://www2.nationalgrid.com/Enhanced-Frequency-Response.aspx

[PLH] Papy JM, Latauwer L de, Huffel S van, Exponential data fitting using multilinear algebra: The single-channel and multi-channel case, Numer Lin Alg Appl 12 (2005) 809-26.

[PMPR] Phillips NE, Manning C, Papalopulu N, Rattray M, Identifying stochastic oscillations in single-cell live imaging time series using Gaussian processes, arXiv (2016)

[RW] Rasmussen CE, Williams CKI, Gaussian processes for machine learning (MIT press, 2006).

[RR] Reece S, Roberts S, An introduction to Gaussian processes for the Kalman filter expert,...

[RR+] Reece S, Roberts S, Ghosh S, Rogers A, Jennings N, Efficient state-space inference of periodic latent force models, arXiv 1310.6319v2 (2014)

[Ri] Riemann GFB, Über die Anzahl der Primzahlen unter einer gegebenen Grösse, Monat. der Königl. Preuss. Akad. der Wissen. zu Berlin aus der Jahre 1859 (1860) 671-680.

[RO+] Roberts S, Osborne M, Ebden M, Reece S, Gibson N, Aigrain S, Gaussian processes for time-series modelling, Phil Trans Roy Soc Lond A 371 (2013) 20110550.

[Rog] Roger G, Power system oscillations (Kluwer, 2000).

[Rom] Romer D, Advanced Macroeconomics (McGraw-Hill, 2001)

[SM] Susuki Y, Mezic I, Nonlinear Koopman modes and a precursor to power system swing instabilities, IEEE Trans Power Sys 27 (2012) 1182–91.

[SMH] Susuki Y, Mezic I, Hikihara T, Coherent swing instability of interconnected power grids and a mechanism of cascading failure, in: Control and optimisation methods for electric smart grids, eds Chakraborty A, Ilic MD (Springer, 2012) 185–202.

[TBT] Tobar F, Bui TD, Turner RE, Learning stationary time series using Gaussian Processes with nonparametric kernels, Adv Neural Info Processing Sys 27 (MIT Press, 2015)...

[TBP] Trip S, Bürger M, Persis C De, An internal model approach to (optimal) frequency regulation in power grids with time-varying voltages, arXiv: 1403.7019v3 (2015)

[T] Trudnowski DJ, Estimating electromechanical mode shape from synchrophasor measurements, IEEE Trans Power Sys 23 (2008) 1188–95.

[TR+] Turunen J, Renner H, Hung WW, Carter AM, Ashton PM, Haarla LC, Simulated and measured inter-area mode shapes and frequencies in the electrical power system of Great Britain, in IET International Conference on Resilience of Transmission and Distribution Networks (RTDN2015) (IET, 2015) 136–141.

[TT+] Turunen J, Thambirajah J, Larsson M, Pal BC, Thornhill NF, Haarla LC, Hung WW, Carter AM, Rauhala T, Comparison of three electromechanical oscillation damping estimation methods, IEEE Trans Power Sys 26 (2011) 2398–2407

[ZD] Zielinski TP, Duda K, Frequency and damping estimation methods - an overview, Metrol Meas Syst 18 (2011) 505–28.

Mathematics Institute and Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, U.K.

E-mail address: R.S.MacKay@warwick.ac.uk