

EFFICIENT MULTISCALE SIMULATION OF CIRCADIAN RHYTHMS USING AUTOMATED PHASE MACROMODELLING TECHNIQUES

SHATAM AGARWAL

Indian Institute of Technology, Kanpur
shatam@iitk.ac.in

JAIJEET ROYCHOWDHURY

University of Minnesota, Twin Cities
jr@umn.edu

Circadian rhythm mechanisms involve multi-scale interactions between endogenous DNA-transcription oscillators. We present the application of efficient, numerically well-conditioned algorithms for abstracting (potentially large) systems of differential equation models of circadian oscillators into compact, accurate phase-only macromodels. We apply and validate our auto-extracted phase macromodelling technique on mammalian and *Drosophila* circadian systems, obtaining speedups of 9 – 13 \times over conventional time-course simulation, with insignificant loss of accuracy, for single oscillators being synchronized by day/night light variations. Further, we apply the macromodels to simulate a system of 400 coupled circadian oscillators, achieving speedups of 240 \times and accurately reproducing synchronization and locking phenomena amongst the oscillators. We also present the use of parameterized phase macromodels for these circadian systems, and elucidate insights into circadian timing effects directly provided by our auto-extracted macromodels.

1. Introduction

Circadian rhythms are amongst the most fundamental of physiological processes. They are found in virtually all organisms, ranging from unicellular (*e.g.*, amoebae, bacteria) to complex multicellular higher organisms (*e.g.*, human beings). These daily rhythms, of period about 24 hours, are associated with periodic changes in hormones controlling sleep/wakefulness, body temperature, blood pressure, heart rate and other physiological variables. Importantly, circadian rhythms are *endogenous* or *autonomous*; however, they are typically influenced by external cues, such as light. Progress in quantitative biology has established that such rhythms stem fundamentally from the molecular level,^{1,2} involving complex chains of biochemical reactions featuring a number of key proteins/hormones (such as melatonin and melanopsin), whose levels rise and fall during the course of the day. These biochemical reactions, which take place both within individual cells and at an extracellular level, function as *biological oscillators* or *body clocks*.³

Quantitative understanding, simulation and control of circadian rhythms is of great practical importance. Applications include devising medical remedies for rhythm disorders (*e.g.*, insomnia, fatigue, jet lag, *etc.*), synthetic biology (where a goal is to “program” artificial rhythms that are biologically viable), artificially extending periods of wakefulness/alertness (*e.g.*, for military purposes), and so on. Improved understanding of circadian rhythm mechanisms has led to increased awareness of how pervasively they affect virtually every aspect of the life of an organism. Hence, their simulation/analysis is an important endeavour in the biological domain.^{1,2}

Although individual oscillators constitute the fundamental core of circadian rhythm mechanisms, the rich circadian functionality of multicellular organisms results from the *interactions of many oscillators* over multiple temporal and spatial scales. Observations of periodicity in behavior, metabolism, body temperature, *etc.*, indicate that coupling/coherence mechanisms play a key role. Hierarchical organization of the circadian system, from the fundamental DNA transcription/translation level to endocrine system levels, involves complex oscillator interactions. The complex connectivity and high dimensionality of such coupled oscillator networks, which lead to unique effects such as *synchronization* and *injection locking/pulling*,^{4,5} make them difficult to understand at the intuitive or analytical level, thus engendering the need for efficient and powerful simulation and analysis tools with multi-scale capabilities.

Several oscillatory mathematical models are available for circadian rhythms^{1,2} that capture the dynamics of the relevant molecular biochemical reactions (see Section 2 for details). These models are in the form of systems of differential-algebraic equations (DAEs) or ordinary differential equations (ODEs). The prevalent technique today for their simula-

tion is to run initial value simulations. While such “time-course integration” of ODEs/DAEs has the advantage of generality, it suffers from serious disadvantages for oscillators, which are inherently *marginally stable*.⁶ For initial-value simulations, marginally stable systems tend to require orders of magnitude more computation for a specified accuracy, particularly *phase/timing accuracy*, than stable systems; even for individual oscillators, very small timesteps (*e.g.*, many hundreds per oscillation cycle) are typically needed, leading to high computational cost. The situation worsens for coupled oscillator systems, which typically feature multiple time scales; *e.g.*, *envelopes*^{7,8} typically feature much longer time scales than individual oscillation cycles.

In electronic circuit design, *automated nonlinear phase macromodel extraction* techniques^{5,6,9} have proven effective in solving such oscillatory problems. Given any oscillator as a system of DAEs or ODEs (however complicated), efficient and well-conditioned numerical techniques extract a *scalar nonlinear differential equation*, the phase macromodel. This macromodel captures the dynamic response of the oscillator’s phase or timing characteristics to external influences. It has been shown that such “PPV” (Perturbation Projection Vector) phase macromodels are able to accurately capture the gamut of phase/frequency-related dynamics of oscillators; most importantly locking, synchronization and phase noise (timing jitter) effects.^{6,10} Using the PPV macromodel instead of the original DAEs/ODEs confers important advantages: large simulation speedups due to system size reduction, the ability to use larger timesteps than for the original system, abstraction to the phase or timing level, precise insight about timing influences without the need for simulation, *etc.*. These advantages are especially pronounced for systems of many coupled oscillators spanning different temporal and spatial scales.¹¹

In this work, we present the first application of PPV-based automated nonlinear time-shifted macromodelling methods to biological systems, focussing on circadian rhythms. We use PPV phase macromodels^{6,12} to model circadian oscillators and show that they are considerably more efficient than standard “time course” simulations. PPV models alleviate the lack of accuracy and general applicability of a widely used prior phase model (Kuramoto’s model, see below), while retaining its advantage of relative simplicity and computational efficiency. PPVs provide direct insights into the effects of external stimuli, such as slowing down/speeding up of circadian rhythms; for example, it is easy to determine when and how to apply a light pulse for greatest de-synchronization. Using PPV macromodels, we are able to efficiently produce plots of circadian lock range vs amplitude of external stimuli; this is valuable for guiding experiments, explaining observations, and designing new (“synthetic”) DNA/protein based biological clock networks. Furthermore, we also present the application of *parameterized* PPV macromodels,¹³ which directly incorporate *effects of parameter changes* into our phase-only models of circadian rhythms. Being able to directly and quantitatively predict the impact of parameter changes on phase, frequency and timing behaviour is of significant biological value.

Indeed, the PPV constitutes a rigorous, Floquet-theoretic generalization of Winfree’s seminal concept of timing maps and phase sensitivity functions,¹⁴ used within “phase-only models” of oscillators popular in computational biology. Kuramoto¹⁵ applied the theory of asymptotics to find *sinusoidal* expressions for Winfree’s phase sensitivity functions; these are widely used for phase-only models, since they are capable of capturing synchronization effects. However, the sinusoidal simplifications inherent in Kuramoto’s model lead to significant inaccuracies for non-sinusoidal oscillators (including circadian ones). These inaccuracies compound in large networks and often lead to qualitatively incorrect conclusions about, *e.g.*, collective synchronization.^{11,16} From a utilitarian point of view, the usefulness of PPV macromodels over Kuramoto is twofold: firstly, PPVs represent the correct (often highly non-sinusoidal) phase sensitivity functions of the oscillator; secondly, the PPV macromodel is generated via an *algorithmic computational procedure* from the oscillator’s DAE/ODE description, typically taking seconds or less to reduce systems 1000s of equations in size.

We apply PPV macromodels to two different circadian rhythm models: one for mammals^{1,17} and one for *Drosophila melanogaster* (the fruit fly, shown in Fig. 1).² We show that the PPV macromodels are significantly faster to simulate than the original equation systems even for single oscillators (9x speedup for the mammalian clock, and 13x speedup for the *Drosophila* clock). Modelling light as an external input impinging upon circadian oscillators, we confirm injection locking using PPV macromodels and obtain plots of lock range vs amplitude of the light signal. We comment on the biological significance of the shapes of important components of the PPV. We then use PPV macromodels to rapidly explore synchronization behaviour in a network of 400 coupled oscillators, obtaining speedups of about 240x over standard time-course simulation. Finally, using parameterized PPV macromodels, we predict the effects of varying a number of model parameters on oscillation frequency and lock range.



Fig. 1: Male *Drosophila* (fruit fly).

The remainder of the paper is organized as follows. In Section 2.1, we provide background on circadian rhythms and their mechanisms, followed by a review of mathematical models for circadian rhythms in Section 2.2. Oscillator and phase macromodels are then introduced in Section 3; a brief review of PPV macromodels, injection locking analysis and parameterized PPVs is provided. Finally, in Section 4, results and speedups are presented.

2. Background and Previous Work

2.1. Circadian rhythms

Circadian rhythms are generated by “clock genes”, which encode genetic instructions that produce certain proteins whose levels oscillate during the course of the day. These oscillating biochemical signals control various functions, such as sleep/waking cycles – in other words, they constitute our “internal biological clock”, which adapts to the daily cycle of day and night. However, the natural period of this internal clock is not exactly 24 hours; it is typically longer if the organism is kept isolated and away from external cues,¹⁸ most importantly light (these cues are called *Zeitgebers*). Therefore, the internal clock needs to be “reset” every day, in order to keep the organism’s bodily rhythms synchronized with the external world’s day/night cycle.

Higher organisms are often composed of billions of cells. The nucleus of each cell contains the genetic material DNA, a long chain-like linear molecule built up of many links. RNA, also a nucleic acid polymer, serves as a DNA template for the *translation* of genes into proteins. The process of formation of an RNA molecule from a particular DNA is called *transcription*. Unlike DNA, RNA is capable of leaving the nucleus and moving into the cytoplasm. There, with the help of enzymes, specific RNA strings get converted to specific proteins responsible for different bodily functions. Some of these proteins return to the nucleus, forming complexes by binding to other proteins, some of which inhibit the expression of their own genes, giving rise to oscillatory patterns in protein concentrations and hence, *circadian rhythms*.^{1,2,18}

In mammals, core clock genes include *Per*, *Cry*, *Bmal1* and *Clock* genes. Their proteins act by inhibiting or stimulating transcriptions of other core clock genes. The proteins of the *Bmal1* and *Clock* genes, namely BMAL1 and CLOCK, form a complex CLOCK-BMAL1 inside the nucleus. This complex activates the transcription of the *Per* and the *Cry* genes. In the cytoplasm, *Per* and *Cry* RNA translate to their respective proteins, PER and CRY. Some of these proteins dimerize to form the complex PER-CRY which returns to the nucleus where, by binding to the CLOCK-BMAL1 complex, it prevents the further transcription of *Per* and *Cry* genes. Thus a negative feedback loop is created; PER and CRY proteins blocking the transcription of their own genes. The above mechanism for the mammalian clock is illustrated in Fig. 2.

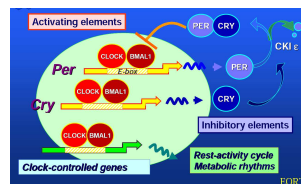


Fig. 2: Mammalian circadian clock mechanism (Francis Levi, EORTC Chronotherapy Group).

2.2. Models of *Drosophila* and mammal circadian rhythms

Computational models are available for circadian rhythms in *Drosophila*, *Neurospora*, cyanobacteria and mammalian systems.^{1,2} These models, useful for computing concentrations of core clock genes, take into account the processes of transcription, translation and phosphorylation.

The mammalian circadian clock model we use¹ consists of 16 variables (hence 16 differential equations) and 52 parameters. It incorporates the effect of negative autoregulation of *Per/Cry* gene expression by their own proteins. The *Drosophila* circadian model we use,² consisting of only 5 variables (5 equations) and 18 parameters, is small enough to be reproduced here:

$$\begin{aligned} \frac{dM_p}{dt} &= v_s \frac{K_I^n}{K_I^n + P_N^n} - v_m \frac{M_p}{K_m + M_p}, \\ \frac{dP_0}{dt} &= k_s M_p - v_1 \frac{P_0}{K_1 + P_0} + v_2 \frac{P_1}{K_2 + P_1}, \\ \frac{dP_1}{dt} &= v_1 \frac{P_0}{K_1 + P_0} - v_2 \frac{P_1}{K_2 + P_1} - v_3 \frac{P_1}{K_3 + P_1} + v_4 \frac{P_2}{K_4 + P_2}, \\ \frac{dP_2}{dt} &= v_3 \frac{P_1}{K_3 + P_1} - v_4 \frac{P_2}{K_4 + P_2} - v_d \frac{P_2}{K_d + P_2} - k_1 P_2 + k_2 P_N, \\ \frac{dP_N}{dt} &= k_1 P_2 - k_2 P_N. \end{aligned} \quad (1)$$

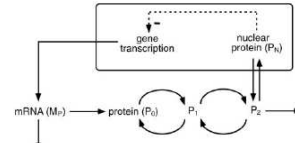


Fig. 3: *Drosophila* clock mechanism².

The variables in Eq. 1 are the same as those shown in Fig. 3; parameters that lead to endogenous oscillations are taken from Gonze/Goldbeter.² Note that Eq. 1 is in the canonical nonlinear ODE/DAE system form

$$\frac{d}{dt} \vec{q}(\vec{x}(t)) + \vec{f}(\vec{x}(t)) + \vec{b}(t) = 0, \quad (2)$$

with $\vec{q}(\vec{x}) \equiv \vec{x}$ and $\vec{b}(t) \equiv 0$. $\vec{b}(t)$ represents the influence of external inputs, such as light, which affects the transcription rate of the *Per* gene in both mammals and *Drosophila*. The effect of light is modelled by including a parameter in the rate equations of the *Per* gene (v_s in Eq. 1), which we recast in the form of Eq. 2 with $\vec{b}(t) \neq 0$ (details in Section 4).

3. Oscillators and PPV phase macromodels

The quantitative study and design of oscillators has a rich history in engineering, particularly in electronics: oscillators are fundamental components in virtually all electronic systems. For example, they are widely used in communication systems for frequency translation of information signals; phase locked loops (PLLs) for clock generation and frequency synthesis, etc.. As noted earlier, phase macromodelling techniques are widely used to improve simulation efficiency and accuracy^{5,19} in electronics. In particular, the Perturbation Projection Vector (PPV) phase macromodel^{6,9} is well established on account of its rigorous Floquet-theoretic underpinnings, broad applicability, effective numerical extraction procedures, large simulation speedups and extensive validation. We have already noted its advantages in Section 1; here, we summarize mathematical details of the model. For expositional convenience, we assume an ODE form for an oscillator under external perturbation:

$$\frac{d}{dt} \vec{x}(t) + \vec{f}(\vec{x}(t)) = \vec{b}(t). \quad (3)$$

$\vec{b}(t)$ is the vector of perturbations applied to the free running oscillator; $\vec{x}(t)$ and $\vec{f}(\vec{x}(t))$ have their usual meanings, as in Eq. 2. The solution of this perturbed oscillator can be shown⁶ to be in the form

$$\vec{x}_p(t) = \vec{x}_s(t + \alpha(t)) + \vec{y}(t + \alpha(t)), \quad (4)$$

where $\vec{x}_s(t)$ is the periodic, oscillatory solution of the unperturbed oscillator and $\alpha(t)$ is a phase deviation caused by the external perturbation $\vec{b}(t)$. $\vec{y}(t + \alpha(t))$ is an amplitude

variation; it is typically very small in circadian oscillators² and is therefore of secondary importance compared to the phase deviation $\alpha(t)$. Using a nonlinear extension of Floquet theory, Demir et al⁶ proved that $\alpha(t)$ is governed by the scalar, nonlinear, time-shifted differential equation

$$\dot{\alpha}(t) = v_1^T(t + \alpha(t)) \cdot b(t), \quad (5)$$

where $v_1^T(t)$ is a periodic vector known as the *perturbation projection vector* or PPV. Importantly, they also showed that the PPV can be calculated efficiently via simple post-processing steps following time- or frequency-domain steady-state computation.^{6,9} Each component of the PPV waveform represents the oscillator's "nonlinear phase sensitivity" to perturbations of that component. The PPV needs to be extracted only once from Eq. 1 (even if parameters change, see the description of parameterized PPVs below); once extracted, Eq. 5 is used for simulations.

3.1. Using the PPV macromodel for systems of coupled oscillators

By employing $\vec{b}(t)$ in Eq. 5 to capture coupling, PPV macromodels can be composed to represent systems of many coupled oscillators with different characteristics. For purposes of illustration, we outline the procedure for N identical oscillators coupling via only one component of $b(t)$. This results in the following set of governing equations for the coupled system:

$$\dot{\alpha}_i(t) = v^T(t + \alpha_i(t)) \cdot \gamma_i(t), \quad i \in 1, \dots, N, \quad (6)$$

where $\alpha_i(t)$ is the phase shift of oscillator i , $v(t)$ is the phase sensitivity of the node on which coupling occurs and $\gamma_i(t)$ is the perturbation resulting on oscillator i due to coupling from other oscillators. If the coupling $\gamma_i(t)$ and phase sensitivity $v(t)$ are purely sinusoidal, it is easy to show that Eq. 6 is equivalent to Kuramoto's model.¹⁵ In general, however, Eq. 6 is far more accurate since it considers all harmonics of the PPVs. We use the coupling function model given in To et al²⁰ as $\gamma_i(t)$ in Eq. 6, and solve for the phase dynamics of a 20×20 network of coupled oscillators.

3.2. Injection locking analysis

When an external signal of frequency f is injected into an oscillator with a central frequency f_o close to f , the oscillator can lock to the injected signal both in phase and frequency. This phenomenon is known as injection locking and can be very easily captured by the PPV macromodel of the oscillator.⁴ It has been shown⁵ that when injection locked, an oscillator's phase shift $\alpha(t)$ varies linearly with time as

$$\alpha(t) = \frac{\Delta\omega}{\omega_o} t + \frac{\theta(t)}{\omega_o}, \quad (7)$$

where ω_o is the natural frequency of the unperturbed oscillator and $\Delta\omega$ the difference between the frequencies of the injected signal and the unperturbed oscillator. $\theta(t)$ represents a *bounded, periodic* phase difference function, the exact form of which can be determined via time-course or steady-state simulation^{5,8} of Eq. 5. The presence of injection locking can therefore be detected by comparing the time-average of $\dot{\alpha}(t)$ with $\frac{\Delta\omega}{\omega_o}$.

3.3. Parameterized PPV macromodels

Circadian rhythm models typically involve large numbers of model parameters. For example, there are 18 parameters in the *Drosophila* clock model,² while the mammalian clock model¹ has 52 parameters. The values of these parameters are chosen so that the model's predictions best fit experimental observations. Leloup/Goldbeter¹⁷ have noted that circadian rhythm properties (particularly frequency) are sensitive to variations in several parameters. The conventional approach to assessing the effect of parameter variations involves brute-force time-course simulation of circadian models, a process that is not only expensive but can also generate numerical inaccuracies in phase.⁵

We, instead, use an extended form of Eq. 5 that directly incorporates parameter variations – we call this the *parameterized PPV macromodel*.¹³ The key advantage of the parameterized PPV macromodel is that it does not involve re-extracting the PPV when parameters change – this leads to huge speedups when, *e.g.*, many coupled oscillators with different parameters are involved. The parameterized PPV equation is given by

$$\dot{\alpha}(t) = v_1^T(t + \alpha(t)) \cdot (b(t) - S_p(t + \alpha(t))\Delta p), \quad (8)$$

where Δp is a vector containing parameter variation terms and $S_p(t)$ is a periodic, time-varying matrix function given by

$$S_p(t) = \frac{\partial f}{\partial p} \Big|_{x_s(t), p^*}. \quad (9)$$

In Eq. 9, $x_s(t)$ denotes the natural periodic solution of the unperturbed oscillator; p^* represents the vector containing nominal (basal) parameter values. This extra term captures phase deviations due to parameter variations, without having to re-extract the PPV when the parameters change. It also enables the study of the effects of multiple parameters varying at the same time.

4. Simulation of mammalian and *Drosophila melanogaster* circadian rhythms using PPV macromodels

In this section, we present results obtained by applying PPV macromodelling, described in Section 3, to mammalian and *Drosophila* circadian rhythm models.^{1,2} We first extract PPV macromodels for both circadian systems at nominal parameters⁹ and then simulate for phase deviation with external perturbation to demonstrate injection locking. We model the external perturbations as changes in external light intensity by first assigning a constant value to the light sensitive parameter v_s (signifying darkness) and then applying an external light signal of intensity

$$L(t) = A + A \sin(\omega t) W/m^2, \quad (10)$$

where $\omega = 2\pi f$, f being the frequency of the light/dark cycles, *i.e.*, corresponding to 1 cycle in 24 hours. Often, light is modelled as a step function for simulations in biological systems (*i.e.*, constant values for light and dark conditions respectively). However, to correspond more closely with continuously changing light intensities in reality and to illustrate the generality of the PPV model, we apply sinusoidal intensity waveform around an average value.²¹ (Note that any other shape, including step function shapes, can be handled equally easily). We assume the experimental setup used by Usui/Okazaki,²¹ where the illuminance of light is varied from 20 lux to 0.01 lux (*i.e.*, variation in light intensity from 0.15 W/m^2 to 0.00009 W/m^2), giving $A \sim 0.05 W/m^2$ in Eq. 10. Moreover, Eq. 10 multiplied by a constant gives the term $b(t)$ of Eq. 2, where the constant signifies the change in *Per* gene concentration for 1 W/m^2 of light intensity. In this paper, we assume the constant to be equal to 1 $nM/(W/m^2)$. The constant can be modelled accurately in future experiments.

We also extract parameterized PPV macromodels to study the effect of parameter variations in two cases – with and without external light variations. In the absence of external light variations, phase deviations from the parameterized PPV macromodel are useful for predicting changes in free-running frequency. When external light perturbations are included in parameter-varying PPV simulations, lock range information is also generated.

Finally, we put the above single-oscillator PPV macromodels together to model a locally-coupled 20×20 network of oscillators – a simple representation of a spatially multi-scale, coupled circadian system. We use this model to demonstrate synchronization behavior, obtaining speedups of about $240 \times$ over traditional time-course simulation.

4.1. Time-course simulations using full ODE models

For reference and validation, we first perform time-course simulations of the two ODE circadian rhythm models directly, to obtain concentration waveforms for all clock proteins and mRNAs in the model. The waveforms thus obtained are shown in Fig. 4(a) and Fig. 4(b). We observe an anti-phase relationship between the concentrations of the *Per/Cry* and *Bmall* mRNAs, as expected from theory.¹ The period of the oscillating waveforms is equal to 23.8 hrs for the mammalian clock and 22.4 hrs for the *Drosophila* clock.

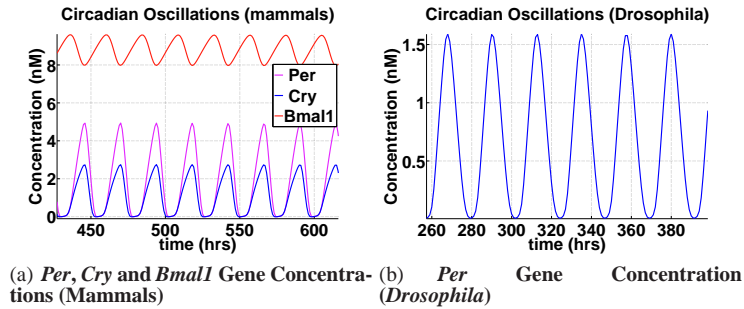


Fig. 4. (a) Plot of core clock gene concentrations (*Per*, *Cry* and *Bmal1*) in mammals vs. time. The concentrations are oscillatory and there is an antiphase relationship between the *Per/Cry* and *Bmal1* concentrations. (b) Plot of the *Per* gene concentration in *Drosophila* vs. time

4.2. Circadian PPV macromodels

In this section, we extract the PPV macromodel of the circadian oscillator for both models. Fig. 5(a) and Fig. 5(b) show the PPV waveforms of *Per* gene concentrations. This waveform gives the phase sensitivity of the concentration at each time instant and can be directly used to find the new concentration waveform under the effect of an external perturbation. It is equivalent to the phase response curve described by Winfree,¹⁴ with the only exception that PPV waveforms do not involve sinusoidal simplifications,¹⁵ implying greater accuracy, as already noted previously. By inspecting the phase sensitivity at each time instant, it becomes possible to determine the time at which light should be applied to shift the oscillator's time-keeping forward or backward. At zero crossings of the PPV phase sensitivity function, for example, a light pulse will have no effect on the phase/frequency characteristics of the oscillator.

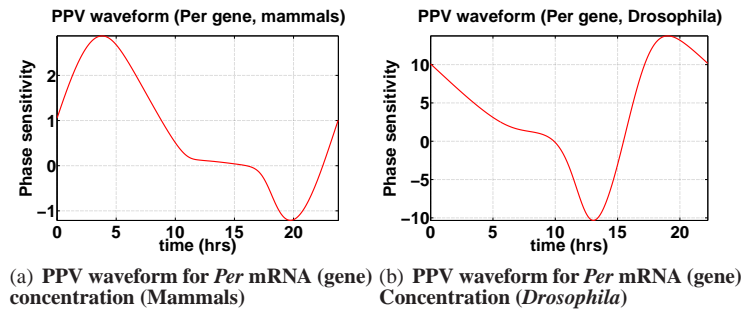


Fig. 5. (a) and (b) Plot of PPV phase sensitivities vs. time for the *Per* gene concentration.

Speedups: (a) For the mammalian clock model: full time course simulation using the Backward Euler (BE) integration method requires 18 seconds of computer time, while finding the free-running steady state via harmonic balance analysis takes about 6 seconds. This is followed by the PPV extraction algorithm, which takes around 1.5 seconds; the total time required for PPV extraction is about 7.5 seconds, representing a speedup of some $2.5 \times$. (b) For the *Drosophila* model: full time course simulation takes about 13 secs; harmonic balance analysis takes ~ 4 seconds, PPV extraction ~ 0.5 seconds; resulting in a speedup of $\sim 3 \times$.

To gauge the accuracy of the PPV macromodels, we plot concentration waveforms for the *Per* gene, obtained from time course and PPV simulations, in Fig. 6. Fig. 6(a) shows waveforms for an locked oscillator (distinct frequencies), while Fig. 6(b) shows waveforms

