JOURNAL OF THE ROYAL SOCIETY Interface

Design principles underlying circadian clocks

D. A. Rand, B. V. Shulgin, D. Salazar and A. J. Millar

J. R. Soc. Interface 2004 **1**, 119-130
doi: 10.1098/rsif.2004.0014

THE ROYAL SOCIETY

J. R. Soc. Interface (2004) **1**, 119–130
DOI: 10.1098/rsif.2004.0014

Design principles underlying circadian clocks

D. A. Rand[1,†], B. V. Shulgin[1], D. Salazar[1,2] and A. J. Millar[2]

[1] *Interdisciplinary Programme in Cellular Regulation & Mathematics Institute,*
and
[2] *Interdisciplinary Programme in Cellular Regulation & Department of Biological Sciences,*
University of Warwick, Coventry CV4 7AL, UK

A fundamental problem for regulatory networks is to understand the relation between form and function: to uncover the underlying design principles of the network. Circadian clocks present a particularly interesting instance, as recent work has shown that they have complex structures involving multiple interconnected feedback loops with both positive and negative feedback. While several authors have speculated on the reasons for this, a convincing explanation is still lacking. We analyse both the flexibility of clock networks and the relationships between various desirable properties such as robust entrainment, temperature compensation, and stability to environmental variations and parameter fluctuations. We use this to argue that the complexity provides the flexibility necessary to simultaneously attain multiple key properties of circadian clocks. As part of our analysis we show how to quantify the key evolutionary aims using infinitesimal response curves, a tool that we believe will be of general utility in the analysis of regulatory networks. Our results suggest that regulatory and signalling networks might be much less flexible and of lower dimension than their apparent complexity would suggest.

Keywords: circadian clocks; gene expression; feedback loops; oscillations; mathematical models; flexibility

1. INTRODUCTION

Current descriptions of the molecular circadian clock have a negative feedback loop with delay at their heart (Goldbeter 2002; Roenneberg & Merrow 2003; Young & Kay 2001). Indeed, such a single feedback loop with a very simple structure will produce robust oscillations (Leloup *et al.* 1999). It is, therefore, pertinent to ask why current understanding of the regulatory networks of these clocks suggests that they almost universally have a much more complicated structure, with multiple interlocking feedback loops with both negative and positive feedback. Several authors have speculated on the reasons for this complexity (Albrecht *et al.* 2001; Cheng *et al.* 2001; Cyran *et al.* 2003; Daan *et al.* 2001; Glossop *et al.* 1999; Johnson *et al.* 2003; Lee *et al.* 2000; Preitner *et al.* 2002; Reddy *et al.* 2002; Roenneberg & Merrow 2003; Smolen *et al.* 2001; Ueda *et al.* 2001). The primary reason (although not the only one) suggested is robustness, either to parameter perturbations or to stochastic fluctuations. However, it has not been shown that the observed structure leads to robustness or, indeed, why one would expect this. Of course, it is likely that some of the complexity arises from specific needs of the organism in question. Nevertheless, it is important to consider whether there are general principles behind the form of the structures observed, especially in view of the fact that the disparate clock mechanisms maintain biological rhythms in a very similar fashion in all organisms, an apparent instance of convergent evolution (Young & Kay 2001).

In this paper we address this problem by using differential equation models of the clock. We show (i) that clocks involving a single loop are inflexible in a precise sense (defined below) and (ii) that the degree of flexibility of a clock network is related to the complexity of the loop structure. Then we analyse why this flexibility is important for the functioning of the clock.

Circadian oscillators are entrained by the daily cycles of light and temperature (Johnson *et al.* 2003). Entrainment by light is generally considered to work by modulating a small number of particular parameters of the regulatory network such as certain degradation or expression rates. For temperature the mechanism is less clear. It is generally assumed that temperature fluctuations affect many more rates and it is unclear how these combine (Rensing & Ruoff 2002). For entrainment by light or temperature to work, the clock must be sensitive to fluctuations in either of these environmental factors. On the other hand, an important property of many clocks is that key characteristics such as period are not sensitive to sustained changes in, for example, temperature (Johnson *et al.* 2003). To analyse the relations between these goals that are *prima facae* at odds

---

[†] Author for correspondence (dar@maths.warwick.ac.uk).

Received 30 July 2004
Accepted 1 October 2004

Table 1. Flexibility dimension and relevant singular spectrum of various published models. n and s are, respectively, the number of dynamical variables and parameters. The four values given for d are, respectively, the values of the flexibility dimension d when $\varepsilon^2 = 0.05$, 0.01, 0.005 and 0.001, so that the first d principal components capture approximately 95, 99, 99.5 and 99.9% of the variance. Thus the third of these is the number of i with $\log_{10} \sigma_i/\sigma_1$ greater than $(\log_{10}\ 0.005)/2 \sim 1.15$. The values of d given are for changes that are measured relative to the size of the corresponding parameter values and therefore the singular values are those for $M^*\Delta$ (see the text). It is usually more relevant to consider these than the values for absolute changes although, in fact, the values of d obtained in this case are almost the same. Almost universally s is bigger than d by an order of magnitude and d grows roughly linearly with the loop complexity. The ratio s/d is lower for the last model, but in this case the number of parameters s has been kept artificially low by assuming that several different parameters are identical. It should be noted that we have not included the Goodwin model (Goodwin 1965) or the model of Tyson *et al.* (1999). The Goodwin model was used to predict a variety of *Neurospora* circadian clock properties which later were found experimentally (Ruoff & Rensing 1996; Ruoff *et al.* 2001). The model of Tyson *et al.* was one of the earliest models for the *Drosophila* clock. The methods of this paper can be applied to these models, but we have not included them here because in the published versions the light input is not included. Both have very low flexibility dimensions.

Model	n	s	d	$\log_{10} \sigma_j/\sigma_1$ with $\sigma_j/\sigma_1 > 10^{-2}$
Leloup *et al.* (1999) *Neurospora*	3	10	1, 2, 3, 4	0, −0.97, −1.15, −1.26, −1.62, −1.75, −1.81
Leloup *et al.* (1999) *Drosophila*	10	38	2, 3, 3, 6	0, −0.19, −0.85, −1.32, −1.38, −1.43, −1.59, −1.73, −1.78, −1.91, −1.93
Ueda *et al.* (2001) *Drosophila*	10	55	1, 1, 1, 2	0, −1.33, −1.52, −1.67, −1.84, −1.93
Leloup–Goldbeter (2003) mammalian (Leloup & Goldbeter 2003)	16	53	1, 2, 2, 5	0, −0.71, −1.24, −1.4, −1.5, −1.71, −1.74, −1.78, −1.88
Mammalian with 2nd *Per* loop (see Electronic Appendix)	27	93	2, 3, 3, 6	0, −0.64, −0.74, −1.24, −1.27, −1.42, −1.52, −1.73, −1.77, −1.8, −1.82, −1.9, −1.95
Forger–Peskin (2003) mammalian	73	36	5, 7, 9, 10	0, −0.14, −0.37, −0.58, −0.61, −0.67, −0.82, −1.07, −1.1, −1.32, −1.53, −1.61, −1.7, −1.83, −1.84, −1.99

we introduce a set of tools, infinitesimal response curves (IRCs), which allow us to characterize the stability and entrainment properties of the clock.

The phenotype of the clock is largely determined by the set of characteristics describing, for example, how it is entrained by light and temperature, the phase relationships between the protein products, the coordination of output pathways and which phases it can track (e.g. dawn and dusk). These characteristics need to be robustly persistent in response to both sustained and stochastic changes in environmental variables such as temperature and pH, to changes in nutrition or growth conditions and to internal fluctuations of the molecular environment of the cell. They are largely determined by the network structure of the clock and the values of the various parameters (such as rate constants) describing the interactions quantitatively. Thus, we can regard evolution as acting on both the network structure and the parameters by small changes, as revealed by the natural genetic variation (Johnson *et al.* 2003) in the mouse, *Drosophila*, and *Arabidopsis*. Although larger changes occur, such as the deletion of core genes, they do not seem to be maintained and thus we do not consider them here.

Evolution will act to simultaneously tune the multiple and possibly conflicting characteristics. To understand how this can be achieved one has to address two problems:

(a) one must understand the relationships between the various characteristics (e.g. to what extent they can be tuned independently and which of them are strongly related);
(b) one must ascertain how flexible the clock's structure is, i.e. how easy it is to simultaneously tune for the multiple goals.

To address the first of these (a), we show that the key goals can be expressed in terms of certain system variables so that each of them corresponds to tuning one or more of these system variables to certain prescribed values. These system variables can be simply expressed in terms of the IRCs introduced below. Our analysis shows that there are a significant number of effectively independent quantities that evolution has to act on and one is able to determine which combinations of parameters are to be tuned in order to move towards the realization of a specific circadian characteristic.

To address problem (b) we introduce below a measure of the flexibility. Its importance is that it tells us how many key system variables evolution is able to simultaneously tune and in how many dimensions evolution can move the system. We produce evidence that this flexibility is much lower than one would expect from the fact that all clock systems depend upon a very large number, s, of parameters such as rate or coupling constants.

For the published models considered in table 1, s ranges from nine to over 50 and in reality the number of parameters is likely to be higher. On the face of it, the large number of tunable parameters might suggest that there is a huge flexibility to explore and optimize key characteristics of the clock. However, it turns out that our analysis shows that this is not the case and that the maximum number of goals that can be simultaneously tuned is much smaller than s, usually by an order of magnitude, and is roughly proportional to the loop complexity of the network. In order to do this we introduce a measure of the maximum number of goals that can be simultaneously tuned which we call the *flexibility dimension*. We argue that:

- there is a significant number, q, of key evolutionary targets that are largely independent of each other;
- the realization of these is effectively impossible if the flexibility dimension d is less than q;
- therefore, there is a selective advantage in increasing d, which generates selective pressure for increased loop complexity.

The literature now contains a broad array of explicitly molecular models for circadian clocks that are described by differential equations (Goldbeter 2002). This provides the most appropriate framework in which to address the above problems. The species covered include *Neurospora*, *Drosophila* and the mouse. These models were constructed using the current molecular information to define the regulatory network structure. The parameters were tuned to obtain as good a match as possible with experimental results, such as measured phase relationships and phase response curves, and the results of mutation analysis. In hardly any cases have the parameter values been measured experimentally, although broad physiological ranges are known for many. Our analysis is not specific to any of these models but in table 1 we list a number of them that are used to illustrate the points we make. All these models display sustained oscillations in the appropriate parameter regimes and are entrained by light–dark cycles of appropriate intensity. This sustained oscillation is described by a (stable) limit cycle.

Throughout we denote the parameters of the system under consideration by k_i and collect them into a vector $k=(k_1,\dots,k_s)$. These parameters, for example, determine the functional descriptions of transcription, translation, (de)phosphorylation, degradation and substrate binding. When the parameters are changed (usually by small amounts), the variation is denoted by $\delta k=(\delta k_1,\dots,\delta k_s)$ (so that the new parameter values are given by $k+\delta k=(k_1+\delta k_1,\dots,k_s+\delta k_s)$) and the change caused in an output such as Q_j is denoted by δQ_j. Outputs Q_j will include quantities such as period, the phases of the maxima and minima of mRNA and proteins, the amplitude of these maxima and minima, and the levels of mRNA and protein at prescribed phases. They can also be functions rather than numbers; for instance a phase response curve (PRC). All of the outputs Q_j we consider are functions of the limit cycle, i.e. assuming that we know the differential equations describing the clock, the value of the output Q_j is determined by the limit cycle.

2. CHARACTERIZING EVOLUTIONARY GOALS: IRCs

We now turn to problem (a): to what extent can the key characteristics be tuned independently and which of them are strongly related. In particular, we wish to describe which combinations of parameters can be tuned in order to produce a specific circadian characteristic. The aim is to characterize the key evolutionary goals so that they are given by transparent and comparable mathematical conditions. We use the fact that the effect of small parameter changes can be well approximated using perturbation theory. The quality of approximation can be determined by the calculation of higher order terms.

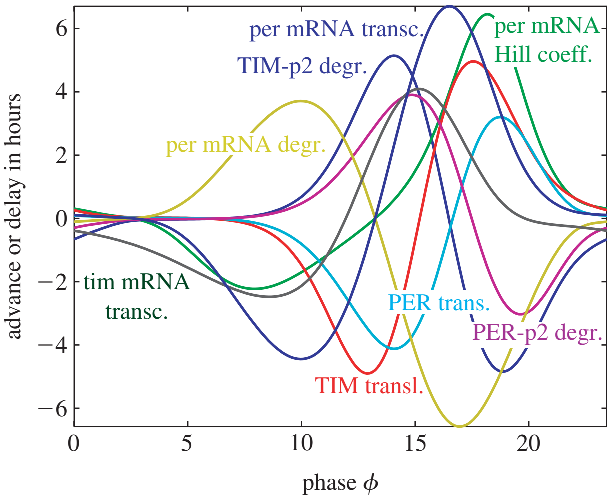

Figure 1. All the large amplitude IRCs $f_{k_i,\mathrm{period}}(\phi)$ for period for the *Drosophila* model of Leloup *et al.* (1999).

2.1. IRCs

Suppose that the oscillator has a stable limit cycle of period p given by $x=g(\phi)$ with ϕ representing the phase, i.e. time t mod p. The key insight is that for each parameter k_i and output Q_j there is a function $f_{k_i,Q_j}(\phi)$ of the phase ϕ with the following property: if one changes a number of parameters $k_1,\dots,k_s$ to $k_1+\delta k_1,\dots,k_s+\delta k_s$ only when the phase ϕ is between ϕ_1 and ϕ_2, then the linear approximation to the change δQ_j in an output variable Q_j is of the form

$$\delta Q_j=\sum_{i=1}^{s}\delta k_i\left(\int_{\phi_1}^{\phi_2} f_{k_i,Q_j}(\phi)\,\mathrm{d}\phi\right). \tag{2.1}$$

This relation holds for all choices of ϕ_1 and ϕ_2 and, therefore, uniquely defines the function $f_{k_i,Q_j}(\phi)$ that we call the IRC for parameter k_i on output variable Q_j. Since these curves can readily be computed from their analytical expressions, for a given oscillator one can rapidly numerically calculate all IRCs for all parameters k_i and all relevant output variables Q_j (see figure 1 and Electronic Appendix).

2.2. Relationship to PRCs

When Q_j is the period of the cycle,

$$f(\phi)=-\delta k_i\int_{\sigma_1}^{\sigma_2} f_{k_i,Q_j}(\phi)\,\mathrm{d}\phi \tag{2.2}$$

is the PRC of a small perturbation δk_i in the parameter k_i applied between the phases ϕ_1 and ϕ_2. If this change in k_i is caused by light, then this is the usual PRC for a light pulse applied between the phases ϕ_1 and ϕ_2. Thus, since for the models under consideration these curves provide excellent approximations to the usual PRCs even when the perturbations are not particularly small (see figure 2), the analytical expressions allow

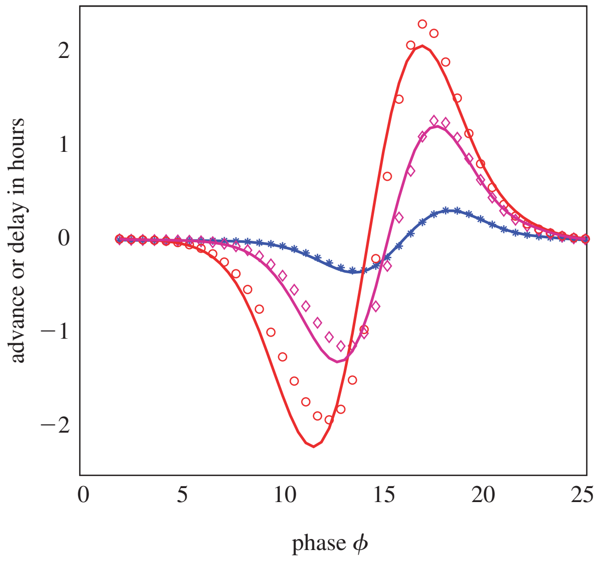

Figure 2. PRCs for the *Drosophila* model of Leloup *et al.* (1999) for light pulses of duration 1 (blue stars), 2 (purple triangles) and 4 hours (red circles) together with an approximation obtained by integrating the corresponding IRC (as given in equation (2.4)) from dawn to dusk (corresponding colour full curve). The PRCs are calculated directly by numerically solving the differential equation which has been perturbed by an appropriate light pulse and measuring the phase change. In this model light acts by increasing the degradation of TIM-p2 and thus the IRC used in the integral is $-f_{\nu_{dt},\text{period}}(\phi)$ because the parameter ν_{dT} is changed when light is on. The effect of the light is as in the original publication (ν_{dT} is changed from 3 to $6\,\text{nM}\,\text{h}^{-1}$), but we very slightly changed the light profile by smoothing out the discontinuities.

rapid computation of all possible PRCs without having to simulate the system and calculate PRCs in the usual way. Later we will see that the fact that IRCs can approximate PRCs allows us to analyse such things as the relation between entrainment and temperature compensation. Figure 1 shows the largest amplitude IRCs for a model of the *Drosophila* clock when Q_j is period. It should be noted that only type 1 PRCs can be approximated in this way because type 0 PRCs require a large perturbation that changes the topology of the PRC.

Using the results of an analysis such as that shown in figure 1, in the next section we show that one can estimate the effects of any hypothetical input pathway on the phase and strength of entrainment. Moreover, by the linearity of equation (2.1) it follows that one can also estimate the effect of combinations of different pathways. For example, if light of intensity I has the effect of changing k_i by $\delta k_i = \delta k_i(I)$ and k_ℓ by $\delta k_\ell = \delta k_\ell(I)$, then the corresponding PRC is given by

$$f(\phi) = -(\delta k_i f_{k_i,\text{period}}(\phi) + \delta k_\ell f_{k_\ell,\text{period}}(\phi)).$$

2.3. Relating entrainment and IRCs

The simplest (but only approximate) way to relate IRCs to stable entrainment is as follows. Suppose that the effect of light of intensity I is to change the parameter k_i by an amount $\delta k_i = \delta k_i(I)$ to $k_i + \delta k_i(I)$. We assume that the light acts periodically with period L (usually 24 h), the length of a day. Let us consider the phase ϕ_n of the oscillator at dawn on day n. A simple argument (see Electronic Appendix) shows that

$$\phi_{n+1} = F(\phi_n) = \phi_n + V(\phi_n) + \Delta L, \qquad (2.3)$$

where ΔL (the *circadian correction*) is the difference between L and the period p of the unforced system, and

$$V(\phi) = -\delta k_i(I) \int_{\phi}^{\phi_2-\phi_1+\phi} f_{k_i,\text{period}}(\sigma)\,\mathrm{d}\sigma \qquad (2.4)$$

is the phase response curve (given by equation (2.2)) for light of intensity I acting from dawn $\phi = \phi_1$ to dusk $\phi = \phi_2$. For (2.3) to be a good approximation, equation (2.4) must give a good approximation of the PRC (as in figure 2) and, for the unforced system, perturbed trajectories must rapidly return to the limit cycle. Equation (2.3) determines a dynamical system so that if ϕ_0 is the phase at dawn on day 0 then the phases on days 1, 2, 3, 4 and so on are given by $\phi_1 = F(\phi_0)$, $\phi_2 = F(\phi_1) = F(F(\phi_0)) = F^2(\phi_0)$, $\phi_3 = F^3(\phi_0)$, $\phi_4 = F^4(\phi_0), \ldots$. We call this dynamical system F the *phase response mapping* and in figure 3 we show this mapping for the *Drosophila* model of Leloup *et al.* (1999).

Note that if $V(\phi)$ is regarded as a PRC then F is related to the so-called phase transition curve (PTC). The PTC $\phi' = Q(\phi)$ is defined to be the final phase reached at the end of the stimulus that began at phase ϕ. Thus $F(\phi) = Q(\phi) + \Delta L$.

Entrainment corresponds to the existence of a stable fixed point ϕ_* of the map F given by (2.3) because then, for almost all starting conditions ϕ_0, the system eventually settles down to a state where ϕ_n is approximately constant at $\phi_*(\phi_n \to \phi_*$ as $n \to \infty)$. A fixed point ϕ_* satisfies the equation $V(\phi_*) = -\Delta L$. In other words, if the starting phase is this fixed point value ϕ_* the phase shift resets the clock by an amount ΔT so that the next starting phase is the same and so the clock is entrained to period L. If the graph of (2.3) is drawn (as in figure 3), the fixed points correspond to the intersections between the graph and the diagonal given by $\phi_{n+1} = \phi_n$. The fixed point (and hence entrainment) is stable provided that $-2 < V(\phi_*) < 0$ because then $|F'(\phi_*)| < 1$.

Robustness of the entrainment means that the stable fixed point of (2.3) persists under reasonable environmental, physiological and other perturbations. Thus entrainment requires that the amplitude of V is greater than the circadian correction ΔL. Otherwise, there will be no intersection between the graph of the map F and the diagonal given by $\phi_{n+1} = \phi_n$. However, it follows that, since V is given by (2.4), entrainment can only occur if $\delta k_i(I) f_{k_i,\text{period}}$ has sufficient amplitude to produce a phase shift of ΔL (Johnson *et al.* 2003). The required amplitude as a function of ΔL can be estimated from equation (2.3). Moreover, the fixed point ϕ_* determines the phase of entrainment and this can therefore be determined from $f_{k_i,\text{period}}$ in the same manner as from a PRC.

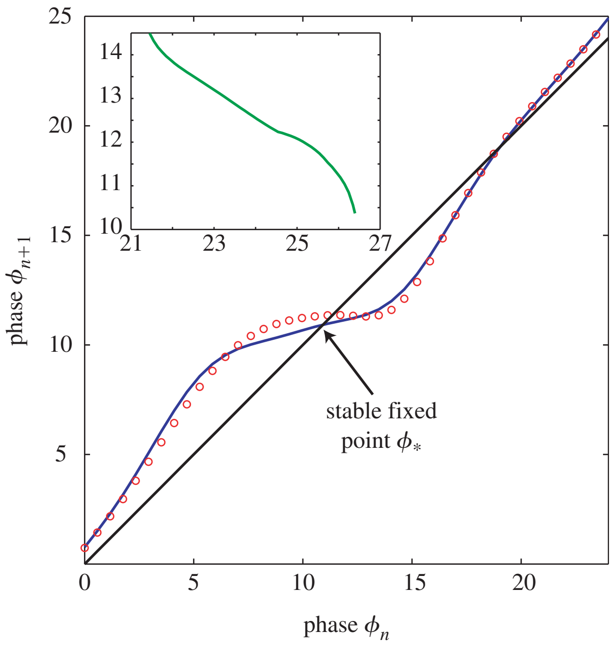

Figure 3. Phase return mapping for the *Drosophila* model of Leloup *et al.* (1999) when the day length is 9 h. The blue full curve shows the mapping as calculated using the IRCs as in equations (2.3) and (2.4) and the red circles show what is obtained if instead of using the IRC to calculate the PRC V in (2.4) one calculates the PRC numerically as in figure 2. The line $\phi_{n+1} = \phi_n$ is also shown in black. The intersection of the blue curve with this gives the set of fixed points and one of these fixed points ϕ_* is stable. This has the property that if the initial phase is not a fixed point then the phase ϕ_n on day n converges to ϕ_* as n increases. The embedded graph in green shows the phase of entrainment as a function of the period p of the unforced system (i.e. the free-running system in darkness) calculated using the IRCs as above (full line) and numerically as above (dots). The period p was changed using a software tool that we have developed that allows one to move parameters so as to change a particular output (here p) without changing the other key outputs. This tool is available on our website.

2.4. Temperature compensation

Circadian clocks maintain period under a wide range of environmental conditions. In the case of temperature, this property is referred to as *temperature compensation* (Rensing & Ruoff 2002). Temperature T will presumably affect a number of parameters k_i that are functions of T: $k_i = k_i(T)$. When temperature changes from T to $T + \delta T$, then the change in the parameter k_i will be approximated by $k_i'(T)\delta T$ where $k_i'(T)$ is the derivative of $k_i(T)$. Thus, by the linearity of equation (2.1) we can define the IRC for temperature (at T) acting on the output variable Q_j by $f_{T,Q_j}(\phi) = \sum_i k_i'(T) f_{k_i,Q_j}(\phi)$. Then temperature compensation (for period) holds around T provided that the integral over a complete cycle of $f_{T,Q_j}(\sigma)$ is close to zero when Q_j is period (Ruoff *et al.* 2000), i.e.

$$\int_0^p f_{T,\mathrm{period}}(\phi)\,\mathrm{d}\phi \approx 0. \qquad (2.5)$$

The chief advantages of IRCs in this context are (a) that the effects of multiple parameter changes are easy to compute and (b) that we can use $f_{T,Q_j}(s)$ (in a similar fashion to $f_{k_i,\mathrm{period}}$ in §2.3) to also study the entrainment properties of temperature. To study entrainment by temperature one can, for example, use a map F as in (2.3), but where $f_{k_i,\mathrm{period}}$ is replaced by $f_{T,\mathrm{period}}$ in the definition of V. Then we see that for robust entrainment by temperature we require that $f_{T,\mathrm{period}}$ has a large enough amplitude, and that to have temperature compensation as well, $f_{T,\mathrm{period}}$ must be balanced in the sense that it must satisfy (2.5). In figure 4 we show an example of a temperature compensated system that satisfies (2.5) and discuss the difference between local and global temperature compensation.

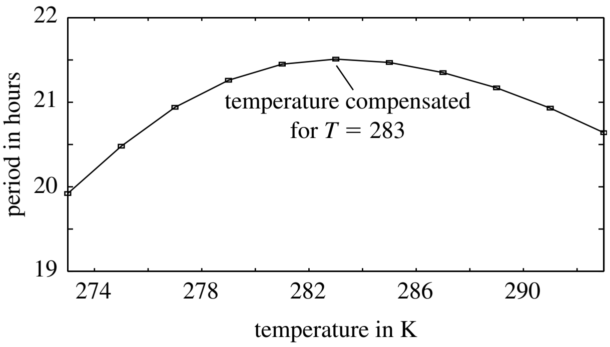

Figure 4. Period (in hours) as a function of temperature (in K) for the *Neurospora* model of Leloup *et al.* (1999) when the dependence of parameters is given by the Arrhenius relation. A commonly considered functional relationship is given by the Arrhenius relation $k_i = A_i \exp(-E_i/RT)$ where T is temperature, A_i is the so-called pre-exponential factor, E_i is the activation energy of the appropriate reaction and R is the gas constant. We have chosen the activation energies E_i so that the model for *Neurospora* of Leloup *et al.* (1999) is temperature compensated at $T = 283$. We see that near to $T = 283$ the period changes slowly with T, but further away it changes more rapidly. Biological oscillators have a more global temperature compensation where the period is roughly constant over a more substantial temperature range. Thus we refer to this as *global* and the sort of temperature compensation seen in the figure as *local*. We will discuss how to achieve global compensation in a future paper.

The preservation of the unforced D-D period p under temperature changes is only a secondary phenomenon for evolution in that the exact value of the period in D-D does not appear to have a direct selective value. One may ask why temperature compensation has occurred. A possible reason is suggested by our discussion of robust entrainment which presumably does have selective advantage. We showed above that the condition for preservation of robust entrainment involves changes in both the function V defined there and the circadian correction $\Delta L = L - p$. Therefore, under temperature changes, if the changes caused in V are small, we get persistent robust entrainment (with a robust phase) provided that p is stable under temperature changes.

2.5. Parameter stability

If we require that the period or other output variable Q_j is relatively stable to (sustained) perturbations of the parameter k_i, then by equation (2.1) we require that the integral over a complete cycle in equation (2.1) of the IRC f_{k_i,Q_j} is close to zero:

$$\int_0^p f_{k_i,Q_j}(\phi)\,\mathrm{d}\phi \approx 0. \tag{2.6}$$

By the linearity of equation (2.1) we can apply this to study mixed simultaneous perturbations of many parameters. Robustness to sustained parameter perturbations requires that all IRCs must satisfy (2.6). We can ignore those IRCs that have small amplitude because they will therefore have small integral. However, by the above discussion, some IRCs must have large amplitude for robust entrainment to be possible. Additionally, for some other parameters k_i and output quantities Q_j, the IRCs will have large amplitude because of the nature of the k_i, which impact directly upon the output Q_j under consideration (e.g. for the effect of phosphorylation and degradation rates on period). It is widely thought that there is a strong evolutionary advantage to robustness, so selection may be expected to change the system so as to balance the large amplitude IRCs so that their integral is small. This provides a significant number of evolutionary goals.

2.6. Robustly tuned output pathways

Let us suppose that a particular output pathway is driven by the molecular species whose level is given by $x_i(t)$. Presumably there will be strong selective pressure on the output pathway to ensure the appropriate phase and amplitude relationships. The change in the level of $x_i(t)$ at $t=t_0$ produced by a small change in the parameters can be calculated directly from the IRCs $f_{k_i,Q_j}(\phi)$ where $Q_j=x_i(t_0)$ via equation (2.1). Moreover, we can also use these same IRCs to calculate the corresponding change in phase as explained in the Electronic Appendix.

3. TRACKING MULTIPLE PHASES

There is a clear advantage if the clock is able to control genes using independent phases (Daan *et al.* 2001; Pittendrigh & Daan 1976). For example, for many organisms there will be a need to trigger behaviours or to express genes at times that are determined by either dawn or dusk. This regulation has been demonstrated in some cases (Pittendrigh & Daan 1976). Since day length varies during the year this cannot be handled by a pathway that introduces a fixed time delay starting from just one phase such as dawn. The obvious way to deal with this is to use either two oscillators or, as we consider here, a single oscillator that has two or more loops, one of which is locked to dawn and while the other is locked to dusk (Daan *et al.* 2001; Pittendrigh & Daan 1976). When we say a loop is locked to dawn we mean that if the timing of dawn is changed by a small amount δ_{dawn} and dusk is kept fixed then the dominant change in all the time-series of the products in the given loop is a simple phase change by approximately δ_{dawn} (and similarly for dusk). Recent evidence suggests that the pairs of mammalian clock genes (*Per1* and *Per2*, *Cry1* and *Cry2*) exhibit some of the properties expected in such systems (Albrecht *et al.* 2001; Oster *et al.* 2002; Reddy *et al.* 2002; Zheng *et al.* 2001). However, to date there are no models of molecular systems that have been shown to have the required properties. Indeed, although they almost all have multiple loops, the published oscillators in table 1 only track a single phase for the parameter values used.

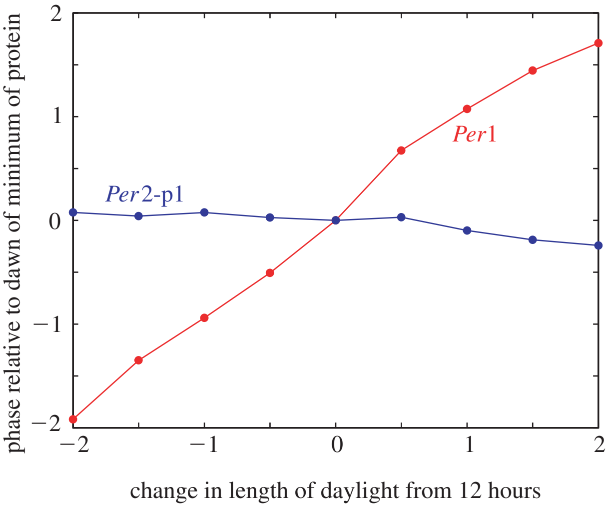

Figure 5. The two *Per* loops of the model detailed in the Electronic Appendix separately track dawn and dusk. The figure shows the phase relative to dawn of the minimum of *Per1* and the phosphorylated *Per2* levels as a function of the change in daylength from 12 h. We obtain the same plot if we replace *Per1* or phosphorylated *Per2* by another product in the same loop. We see that *Per1* tracks dusk and *Per2* tracks dawn. Note that because the model is entrained, if we measure phase relative to dawn (respectively dusk) then the change in phase is just a function of the length of the period between dawn and dusk. Thus, the extent to which a variable tracks dawn is completely determined by how it tracks dusk and *vice versa*.

In general a single loop can only track a single phase. A change of dawn or dusk causes the same phase change to all products in the loop. This is because a small change in either dawn or dusk causes a small perturbation that is localized around the single phase where a change has been made, say for phases in $[\phi, \phi+\delta\phi]$. As argued below, the dominant change to the limit cycle $x=g(\phi)$ at these phases is a simple phase change involving all products. Since for phases outside of the interval $[\phi, \phi+\delta\phi]$ the system has the same light conditions as before, it follows that the change to the limit cycle as a whole is a simple phase change $x=g(\phi+\alpha)$ and the phase relation of all products to one another is unchanged.

However, the theory of dynamical systems tells us that oscillators that can track independent phases can certainly be constructed. This follows from the theory of normal hyperbolicity (Guckenheimer & Holmes 1983) that allows us to couple systems together while preserving such key flexibilities. We have used this theory to construct such a system. It can be shown that the mammalian clock model of Leloup & Goldbeter (2003) tracks dusk. It contains a single *Per* gene and we have added a second *Per* gene loop to this model in such a way that (a) without light the system stably oscillates in a periodic fashion and (b) with light the new loop tracks dawn while the old loops tracks dusk. This model is included in table 1 and the details are given in the Electronic Appendix. Figure 5 shows how it tracks dawn and dusk.

The theory of dynamical systems and our arguments about flexibility and phase changes imply that there are the following restrictions on such systems:

- a single loop can only track one phase even if it has multiple input pathways for light; and
- to track m independent phases one needs at least m loops with at least one input pathway for light into each loop; although loops can share an input pathway this pathway must modulate a different parameter in each loop.

Note that the discussion here suggests an interesting prediction: in many cases (and all where both dawn and dusk are tracked) one should find multiple input pathways for light. It has been pointed out to us by an anonymous referee that evidence for this already exists (Gau *et al.* 2002; Oster *et al.* 2003).

4. FLEXIBILITY AND EVOLUTIONARY ACCESSIBILITY

We now address part (b) of the fundamental problem discussed in §1, namely, how easy it is to simultaneously tune for multiple evolutionary goals.

Each variation $\delta k = (\delta k_1, \ldots, \delta k_s)$ of the parameters will cause the limit cycle to vary and this in turn changes the vector of output characteristics $Q = (Q_j)$ by an amount $\delta Q = (\delta Q_j)$. In principle, since there are s independent parameters it is reasonable to assume that if these are freely varied, then the resulting changes δQ in the output will fill out a s-dimensional set. However, even when the parameter variations δk can freely explore all the parameter combinations, the direction of the resultant vector δQ of outputs may be highly non-uniform with the result that while some specific small output changes δQ can be achieved by small parameter changes (the accessible directions), many others can only be reached by excessively large parameter changes. Significant movement in the latter inaccessible directions is also extremely rare if the parameter changes are random with most movement being in the accessible directions: the outputs δQ resulting from these random variations may be dominated by just a few directions with the result that the outputs only explore a small fraction of the available δQ-space.

Since evolutionary adaption will find it very difficult to move the system in the inaccessible directions it is important to understand them and, in particular, the dimension of the set of the accessible directions. This dimension (which we call the *flexibility dimension*) is an important quantity because it tells us how many of the key output variables evolution is able to tune at any time and in how many dimensions evolution can move the system. It turns out that the definition and calculation of this is a little sophisticated mathematically but, nevertheless, very natural. We now turn to this.

4.1. Definition of accessibility and flexibility

It is possible to understand and measure flexibility using mathematical analysis and thus to obtain universally applicable principles. If the variations δk are small, the relationship between δk and δQ is approximately linear and given by a matrix $M = (M_{ij})$ where $\delta Q_j = \sum_i M_{ij} \delta k_i$ is the change to Q_j caused by a change in k_i of δk_i for $i = 1, \ldots, s$. To characterize accessibility we define what we mean for a output vector to be *accessible* and calculate the dimensionality d of the set of accessible vectors. The accessible vectors can be described either as those directions in δQ-space that are easier to move in using parameter variations or those directions that account for most of the variance of the outputs produced by random variation of the parameters.

In the first approach, for a desired change $\delta Q = (\delta Q_j)$ that can be achieved by a parameter variation $\delta k = (\delta k_1, \ldots, \delta k_s)$, we consider the ratio $R_{\delta Q}$ of the length of the vector of changes (δQ_j) to that of the vector $(\delta k_i / k_i)$. (If there are multiple δk achieving δQ we take the smallest by length.) Because δQ and δk are linearly related, this ratio does not depend upon the size of δQ, but only on its direction.

The reason we take $(\delta k_i / k_i)$ instead of δk_i in the definition of $R_{\delta Q}$ is that the values of the parameters k_i can vary over more than one order of magnitude and it is therefore usually more appropriate to consider the proportional variation, which has the added advantage of being dimensionless. We do not do this scaling for δQ_j because (a) we assume that they have been chosen so that they are dimensionless and so that their sizes are all of the same order of magnitude and (b) later we will consider the case where δQ is the actual variation in the limit cycle, in which case the scaling makes no sense.

To define accessibility we fix a small number ε and say that the direction δQ is *accessible* if $R_{\delta Q}$ is greater than εR^* where $R^* = \sup_{\delta Q} R_{\delta Q}$, i.e. $R_{\delta Q}$ is not too small compared with the maximum possible value it can take. For example, if we choose $\varepsilon = 0.01$, a variation δQ is inaccessible if it requires a percentage parameter change that is more than 100 times the percentage change of δQ.

A simple example may help the reader here. Suppose that the components δQ_j of δQ represent changes in normalized (dimensionless) period (= period/daylength $= p/L$) and the (suitably normalized) phases and amplitudes mentioned at the end of §1. We can ask whether the vector $\delta Q = (0.1, 0, \ldots, 0)$, representing a change in period of $0.1L$ and no change in any of the phases or amplitudes, is accessible or not. If we choose $\varepsilon = 0.01$ then it is inaccessible if the change it represents requires a tenfold change in parameters or more.

In the second approach using random variations, a direction δQ is regarded as being accessible if the variance of the size of the projection of the output variations onto the given direction δQ as a proportion of the total variance of the output variations is greater than ε^2. This is elaborated in Appendix A.

Both approaches lead to essentially the same thing. For both of them the set of accessible vectors forms a cone C that is d-dimensional in the sense that it contains a d-dimensional vector space and no vector space of higher dimension. If W is such a vector space then all the vectors in this cone (i.e. all accessible vectors) are of the form $w+v$ where w is in W and where, compared with w, v is extremely small (of order ε). Thus, neglecting these very small adjustments v, the set of accessible vectors is d-dimensional. We call the number d the *flexibility dimension* of the vector of outputs Q_j under consideration.

The reader should note that these notions of accessibility are applied in the linear regime. The changes δQ and the parameter δk needed to achieve them are small even though in the inaccessible case the size of δk may be much greater than δQ. (Recall that $R_{\delta Q}$ does not depend on the size of δQ, but only on its direction.) We are not considering the case where δQ is so large (and inaccessible) that δk must be so large that the perturbation is outside the region where the linear approximation is accurate.

As defined above, the flexibility dimension for a given vector δQ of outputs depended on the set of output characteristics we have chosen to focus on. However, there is a natural way to define a flexibility dimension of the full system, which accounts for the complete set of possible outputs in one go. This is because all changes in the output characteristics of importance are determined by the change in the limit cycle γ and its period p. A small change δk in the parameters will cause a change $(\delta\gamma, \delta p)$ to the limit cycle γ and its period p, and the changes in all output variables can be calculated from this. Since, the period $p=p(k)$ can vary with the parameters it is necessary to normalize the time-parametrization of γ and replace it by $\bar{\gamma}(t)=\gamma(p(k)t)$. Then, as k varies, $\bar{\gamma}$ remains periodic of period 1 and therefore variations $\delta\bar{\gamma}$ in $\bar{\gamma}$ are always of period 1 and can be compared with each other. Moreover, γ and p are determined by $\bar{\gamma}$ and p and *vice versa*. Thus we consider $\delta k \rightarrow (\delta\bar{\gamma}, \delta p)$ rather than the correspondence $\delta k \rightarrow (\delta\gamma, \delta p)$.

We therefore define the *flexibility dimension* d to be the flexibility dimension of the output vector $\delta Q = (\delta\bar{\gamma}, \delta p)$. Note that if we are dealing with an entrained system, the period p does not change when the parameters are varied by a small amount. Therefore, in this case we can ignore the variations δp.

The linearized relationship $\delta k \rightarrow (\delta\bar{\gamma}, \delta p)$ is described by a linear operator that can easily be well approximated by a matrix M^* and d can be calculated from this matrix as explained in Appendix A and the Electronic Appendix.

4.2. Inflexibility of circadian clock models

In Appendix A we explain how one can estimate the flexibility dimension d. This uses singular value decomposition (Press *et al.* 1988). We have estimated d directly for a range of systems and the results are given in table 1. We find that all of these systems are relatively inflexible in the sense that for small values of ε^2 of the order 10^{-3}–10^{-2} the flexibility dimension d is smaller than the number s of parameters by an order of magnitude. On the other hand, we see that increasing the loop complexity generally causes d to increase in proportion to the loop complexity.

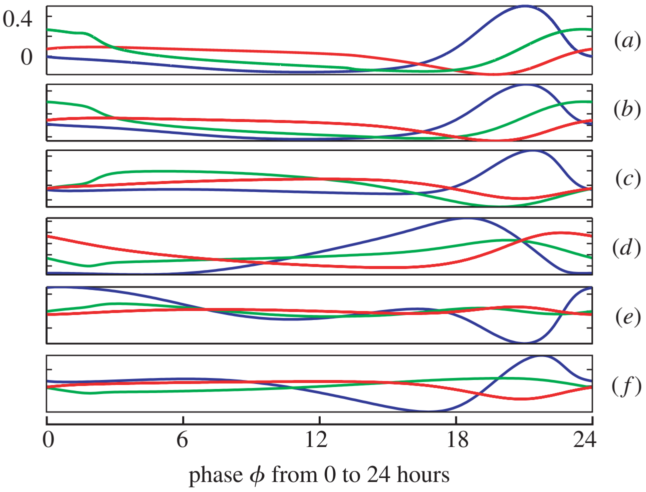

Figure 6. (a) The derivative with respect to time of the limit cycle for the *Drosophila* model of Leloup *et al.* (1999). (b)–(f) The principal components with largest singular values for the matrix M^* for the same model. The units of the ordinate are hours and those of the abscissa are largely irrelevant in that it is the shape of the curves that matter and not their scale. Only the *Per* mRNA, phosphorylated TIM and nuclear *Per*:TIM complex levels are shown. The corresponding singular values are: (b) 1561.7; (c) 318.1; (d) 178.7; (e) 32.6; (f) 17.2. Note that the similarity of the curves in (a) and (b) shows that the dominant principal component is an infinitesimal phase change.

4.3. For entrained systems the dominant principal component is a phase change

If we apply a number of random parameter changes $\delta k^{(\ell)}$ as described in Appendix A we obtain the set of corresponding output change vectors $\delta Q^{(\ell)}$. We can perform a principal component analysis of the set of output vectors $\{\delta Q^{(\ell)}\}$. These principal components can be calculated from the matrix M^* described above in §4.1 (also see the Electronic Appendix).

In each of the entrained systems we find that the dominant principal component is almost exactly given by the derivative $g'(t)$ of the limit cycle $x=g(t)$ (see figures 6a and 6b). Such a principal component $e(t)$ is an infinitesimal phase change, i.e. it is the derivation at $\alpha=0$ of the change in $g(t)$ produced by a change of phase α, i.e. $e(t)=\mathrm{d}/\mathrm{d}t|_{\alpha=0}\, g(t+\alpha)=g'(t)$.

This can be understood as follows. A general insight from the theory of dynamical systems is that those directions that a limit cycle moves most in when parameters are varied are correlated with the directions that are softest with respect to perturbations of the initial conditions, i.e. those directions with the property that a perturbation of the dynamical variables away from the limit cycle is least rapidly damped. These correspond to the direction associated with the Floquet multipliers (Guckenheimer & Holmes 1983) of maximum modulus. For the systems in table 1, a perturbation of the dynamical variables causes a deviation away from the limit cycle, which is then corrected back to zero.

The correction has a rapid phase in which the shape of the limit cycle is recovered (so that $x = g(\phi + \alpha)$) and a slow phase where the phase shift α is corrected. The way that this latter relaxation takes place will typically be described by one or two multipliers: one when the coupling to light is relatively weak and the phase adjusts monotonely and two (as in model 1) when the phase correction overshoots and the multipliers are complex conjugates. Thus, parameter changes easily result in a change of phase while other characteristics (such as phase relationships) are harder to change with the difficulty being greatest in systems with lower loop complexity.

The local structure of the dynamics near to the limit cycle are largely determined by the Floquet multipliers. These are associated with the different rates of contraction onto the limit cycle (Guckenheimer & Holmes 1983). For the clock systems considered here: (i) one multiplier is 1; (ii) all others have modulus less than 1 (i.e. the limit cycle is attracting); and (iii) almost all of them have a very small modulus (corresponding to directions with very fast contraction onto the limit cycle). The inflexibility is due to (iii) because the flexibility dimension d is related to the number of Floquet multipliers λ for which $1/|\log \lambda|$ is small (of order ε).

The fact (iii) is due to the loop structure of the clocks and the nature of protein degradation and modification because these determine the Floquet multipliers. It is therefore expected to be a general feature of regulatory networks. Although it is difficult to prove general results about how rapidly the multipliers decrease, it is clear that a large class of regulatory systems will have this property. What is important for this are the following characteristics: (i) one of more of the protein products in each loop has a degradation rate whose time average is not too small and (ii) the forward and backward rates $k_i^+, k_{i-1}^- : P_i \leftrightarrow P_{i+1}$ (typically corresponding to (de)phosphorylation) also have a time average that is $\mathrm{O}(1)$ on a time scale of hours. This fact ensures that the product of all the multipliers is of the order $\Pi = \exp(-rL)$, where L is the period of the oscillator and r is the number of products in the loop. This product Π is therefore extremely small. A more detailed calculation is needed to show that only very few of the multipliers are larger than $\mathrm{O}(\varepsilon)$.

4.4. Sources of flexibility

The relevant Floquet multipliers with larger modulus (i.e. those giving flexibility) are often associated with specific structural or dynamical aspects. For example, as discussed above, the largest multipliers of the models considered here are usually associated with phase. It appears that a system consisting of a single loop can only possess other relevant multipliers if it has a topology and rate constants that ensure that under reasonable starting conditions the mean time before a protein is degraded is large (for example, because it typically has to go through a series of modifications (and their reverses) before it is in a state where it is targeted for degradation). Otherwise, all but one of the Floquet multipliers have very small modulus.

An interesting consequence of this observation is that the need for flexibility constrains the molecular structure. For example, it seems to imply a selective advantage for selective degradation. The more products in a loop that are degraded, the smaller the modulus of the Floquet multipliers and hence the less flexible the system. Since some degradation has to occur, this suggests that as few as possible of the protein products in a loop should be degraded at as low a rate as possible.

The number of these relevant multipliers is multiplied when loops are coupled and, in addition, new ones result from the coupling. To see this consider the case when the coupling is weak. There will typically be a multiplier associated with the way in which perturbations of the relative phase of the two loops die away. As the coupling is increased this multiplier may become smaller or complex, but for some range of coupling strength it will remain relevant. This suggests that the flexibility dimension is much less than and roughly proportional to the loop complexity.

4.5. Robustness to stochastic perturbations

It is worthwhile considering the relationship between flexibility and robustness to stochastic perturbations in the context of the large Ω limit, where Ω is the number of molecules involved in the clock (Gonze *et al.* 2002). In this limit the fluctuations about the limit cycle are normally distributed with zero mean. The projections of these fluctuations onto the direction of the limit cycle after a time period of Np (p is the period) have a variance that is $\sigma^2 = N/\alpha\Omega$, where α is calculated according to the prescription in Gonze *et al.* (2002). The projections onto the eigendirections that are transverse to the limit cycle have a variance that is $\mathrm{O}(|\lambda_j^2|)$, where λ_j is the corresponding Floquet multiplier. It follows from this that one can often reduce the effect of stochastic perturbations by reducing the multipliers. However, this also reduces the flexibility of the system. Therefore, evolution must trade off these two effects or come up with different ways of counteracting this stochasticity.

5. DISCUSSION: THE NEED FOR FLEXIBILITY IN THE EVOLUTIONARY TUNING OF MULTIPLE TARGETS

We have presented an analysis of key evolutionary aims of circadian clocks such as robustness to perturbations, temperature and pH compensation, appropriate period, robust entrainment by environmental signals, and correct phase relationships for output pathways. It is widely believed that there is selective pressure for each of these targets. We have discussed the characterization of each of these in terms of IRCs (see table 2). These IRCs are relatively easy to calculate and we have provided software to do this on our website (http://www.maths.warwick.ac.uk/ipcr/).

The IRCs for period give accurate approximations of type 1 PRCs for environmental pulses with a broad range of durations and intensities via equation (2.2). This can be used to study entrainment by environmental signals such as light or temperature. We find that robust entrainment requires that the appropriate

Table 2. Summary of how different properties are characterized by IRCs. Each evolutionary aim can be described as tuning one or more particular system variables. Note that in the case of entrainment and PRCs the system variable is actually a function rather than just a number.

Evolutionary aim	Mathematical characterization in terms of IRCs
setting period	change in period due to change δk_i is $\approx \delta k_i \int_0^p f_{k_i,\mathrm{period}}(\phi)\,\mathrm{d}\phi$
robust entrainment	stability and phase determined by $V(\phi) = -\delta k_i(I) \int_{\phi}^{\phi+\mathrm{dusk}-\mathrm{dawn}} f_{k_i,\mathrm{period}}(\psi)\,\mathrm{d}\psi$ when input pathway modulates k_i by $\delta k_i(I)$
PRCs for short disturbances of duration $\mathrm{d}\phi$	PRC approximated by $-\delta k_i f_{k_i,\mathrm{period}}(\phi)\,\mathrm{d}\phi$ when input pathway modulates k_i by δk_i
PRCS for disturbances lasting from ϕ_1 to ϕ_2	PRC approximated by $f(\phi) = -\delta k_i \int_{s}^{\phi+\phi_2-\phi_1} f_{k_i,\mathrm{period}}(\psi)\,\mathrm{d}\psi$
temperature compensation	$\int_0^p f_{T,\mathrm{period}}(\phi)\,\mathrm{d}\phi \approx 0$
pH compensation	$\int_0^p f_{\mathrm{pH},\mathrm{period}}(\phi)\,\mathrm{d}\phi \approx 0$
rough parameter stability for k_i	$\int_0^p f_{k_i,Q_j}(\phi)\,\mathrm{d}\phi \approx 0$
robust phases for output pathways	uses linear combination of $f_{k_i,x_\ell(t_0)}(\phi)$

combination of IRCs for the pathways involved in the environmental input has large amplitude and that the circadian correction $\Delta L = L - p$ (§2.3) is appropriate to ensure a robust stable fixed point of the mapping F given by (2.3). On the other hand, properties such as temperature compensation require that the IRC for temperature, which is a linear combination of basic IRCs $f_{k_i,Q_j}(\phi)$, is balanced in the sense of equation (2.5). A similar argument applies to compensation for sustained variations of other environmental components such as pH. Thus we see that the conditions for robust entrainment by an environmental variable such as temperature and compensation for that variable are independent and perfectly compatible. Stability of key output variables with respect to parameter changes is also characterized by equation (2.6). This suggests that for key outputs Q_j it will be necessary for evolution to roughly balance (in the sense of equation (2.6)) those IRCs $f_{k_i,Q_j}(\phi)$ that have large amplitude because otherwise these outputs will be unstable to variations in the parameters. For example, if Q_j is period and $f_{k_i,Q_j}(\phi)$ has large amplitude and is far from balanced, then sustained variations in k_i are likely to change the circadian correction sufficiently to destroy entrainment (see figure 3 and equation (2.3)). The conditions for correctly tuned output pathways and the robustness of this correct tuning also involve combinations of the IRCs.

Since we can express the various evolutionary aims in terms of IRCs, we can determine to what extent they are independent of each other. This requires an analysis of which IRCs are involved, the nature of the condition on the IRC and the extent to which the IRCs are linearly independent of each other. However, it is not difficult to see that, in general, a large number of the most important evolutionary goals discussed here are independent. It therefore emerges that there are multiple independent characteristics that we can expect to confer a selective advantage and that moreover should be accessible to a process of small random perturbations and selection provided the flexibility of the clock is large enough. It seems reasonable that, there will be more than five or six such characteristics that are of key importance and, therefore, that the required flexibility will at least require a loop complexity equivalent to the most loop complex models in table 1.

In order to simultaneously tune q of the characteristics that are of key importance it is necessary to be able to move the appropriate output vector Q independently in q dimensions by using small changes in the parameters. However, even if one can freely move lots of parameters it does not follow that by doing this one can freely move the output Q with the same dimensionality. Movement of Q in certain directions is highly resistant for the clock systems studied. Seen another way: if the parameter changes are random and uncorrelated, then the movement produced in Q will tend to be highly correlated with the changes strongly concentrated in just a few dimensions. The number d of these dimensions is given by the flexibility dimension that we have introduced and we have provided strong evidence that it is smaller than the number of parameters by an order of magnitude and roughly proportional to the loop complexity of the system.

Thus, evolution will only be effective in reaching multiple independent targets if the flexibility dimension d of the system is as large as the number of targets. If the system is constrained so that it can only reasonably move in a small number of dimensions then it will only be able to tune a small number of targets.

It follows that there is likely to be a strong selective advantage in increasing loop complexity and strong selection for mechanisms that enable this such as gene duplication and protein variation. Mechanisms and divergence that may increase complexity are found in the circadian clocks of *Neurospora* (Garceau *et al.* 1997), *Drosophila*, *Arabidopsis* (Eriksson *et al.* 2003) and the mouse (Daan *et al.* 2001; Oster *et al.* 2002; Zheng *et al.* 2001). In addition, we have argued that the selective degradation of protein products also aids flexibility and therefore that we would expect to find that not all protein products are degraded at the same rate but that degradation is concentrated on selected products in certain modified states. Finally, we have shown that increased loop complexity is implied by the need to track both dawn and dusk (and perhaps other phases). A single loop can only track one phase even if it has multiple input pathways for light; to track multiple independent phases one needs at least the same number of loops with at least one input pathway for light into each loop.

Since one can understand the lack of flexibility in terms of properties of dynamical systems one can make some estimate of the range of applicability of these ideas. It therefore seems rather clear that the ideas discussed will apply to a broad range of dynamical processes of such regulatory networks and not just to oscillating systems. For example, similar ideas should apply to the propagation of perturbations along pathways, multistable networks that act, for example, as switches and networks of transcription factors that determine spatial patterning.

We are grateful to Sanyi Tang and Isabelle Carré for useful discussions on these topics and the BBSRC for funding. We thank Hugo van den Berg for a very helpful critical reading of the manuscript.

APPENDIX A

The mathematical aspects of this paper depend on the theory concerned with the linearization of a dynamical system around a limit cycle (Guckenheimer & Holmes 1983). We have developed a software tool that will calculate most of the concepts that are introduced in this paper. In particular, it will calculate all IRCs and the flexibility dimension. It can also be used to help steer in parameter space so as to achieve a particular goal. This tool is available from our website (http://www.maths.warwick.ac.uk/ipcr/).

A.1. Flexibility, accessibility and random variations

We can also define the accessible cone by considering random variations such as when the outputs $\delta Q^{(\ell)}$ arise from variations $\delta k_j^{(\ell)} = \eta_j^{(\ell)} k_j$, where the proportions $\eta_j^{(\ell)}$ are identically distributed independent random variables. This is a natural approach as it can be thought of as modeling the process of evolutionary mutation. We compare the variance of the size of the projection of the $\delta Q^{(\ell)}$ onto a given direction δQ in the output space with the variance of the sizes $\|\delta Q^{(\ell)}\|$:

$$(R'_{\delta Q})^2 = \frac{\sum_\ell \delta Q^{(\ell)} \delta Q}{\|\delta Q\|^2 \sum_\ell \|\delta Q^{(\ell)}\|^2}.$$

Thus if random perturbations of the parameters produce output changes that tend to have a significant component in a particular direction δQ the $R'_{\delta Q}$ will be significant. On the other hand, if they tend to have little or no projection onto this direction then $R'_{\delta Q}$ will be small.

We define the accessible vectors δQ to be those for which $R'_{\delta Q} > \varepsilon$. If we choose ε so that $\varepsilon^2 = 0.001$, then we can capture 99.9% of the variance by projecting the output changes $\delta Q^{(\ell)}$ onto a subspace within the set of accessible vectors. Randomly chosen perturbations are extremely unlikely to result in Q being changed in directions outside the accessible cone.

A.2. Measurement of flexibility

If the relationship between small parameter changes δk and the changes $(\delta\bar{\gamma}, \delta p)$ caused to the reparametrized limit cycle $\bar{\gamma}$ and its period p is given (or approximated) by the matrix M^* (i.e. $(\delta\bar{\gamma}, \delta p) = M^* \delta k$), then the relationship between the scaled parameters $\eta = (\eta_i = \delta k_i / k_i)$ and $\delta\bar{\gamma}$ is given by $(\delta\bar{\gamma}, \delta p) = M^* \Delta \eta$, where $\Delta = \mathrm{diag}(k_1, \ldots, k_s)$.

The estimation of the flexibility dimension d uses the singular value decomposition (Press *et al.* 1988) of $M^*\Delta$. It can be shown (see Electronic Appendix) that for a given choice of ε the flexibility dimension d is given by the number of singular values σ_j of $M^*\Delta$ with σ_j/σ_1 larger than ε (σ_1 is the largest singular value). Moreover, it is relatively easy to numerically estimate M^* and its singular values (see Electronic Appendix). This is how we estimate d.

Associated with each singular value σ_j is a direction e_j in the output space (see §4.3). These directions are called the *principal components*. For random perturbations $\delta k^{(\ell)}$ of the form considered above (with outputs $\delta Q^{(\ell)}$), the variance of the projection of the $\delta Q^{(\ell)}$ onto the e_j with $\sigma_j > \varepsilon$ is proportional to ε^2. Thus, we choose $\varepsilon \approx 0.001$ and we capture approximately 99.9% of the variance if we approximate the outputs by their projection onto the first d principal components e_j. It is through the characterization of accessibility by singular values that the equivalence of the two approaches to flexibility can be established.

The fact that d is so small compared with s follows from the rapid decrease in the singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_s$ (at most s are non-zero). The way in which they fall off can be determined from the so-called characteristic (Floquet) exponents (Guckenheimer & Holmes 1983) of the limit cycle. Roughly speaking, a perturbation of the parameters induces a change whose magnitude in the direction of the eigenspace of an exponent χ_j is of order $1/|\log \lambda_j|$. This allows one to relate the singular values to dynamical properties of the network.

REFERENCES

Albrecht, U., Zheng, B., Larkin, D., Sun, Z. S. & Lee, C. C. 2001 MPer1 and mper2 are essential for normal resetting of the circadian clock. *J. Biol. Rhythms* **16**, 100–104.

Cheng, P., Yang, Y. H. & Liu, Y. 2001 Interlocked feedback loops contribute to the robustness of the Neurospora circadian clock. *Proc. Natl Acad. Sci. USA* **98**, 7408–7413.

Cyran, S. A., Buchsbaum, A. M., Reddy, K. L., Lin, M. C., Glossop, N. R., Hardin, P. E., Young, M. W., Storti, R. V. & Blau, J. 2003 vrille, Pdp1, and dClock form a second feedback loop in the *Drosophila* circadian clock. *Cell* **112**, 329–341.

Daan, S., Albrecht, U., van der Horst, G. T. J., Illnerova, H., Roenneberg, T., Wehr, T. A. & Schwartz, W. J. 2001 Assembling a clock for all seasons: are there M and E oscillators in the genes? *J. Biol. Rhythms* **16**, 105–116.

Eriksson, M. E., Hanano, S., Southern, M. M., Hall, A. & Millar, A. J. 2003 Response regulator homologues have complementary, light-dependent functions in the *Arabidopsis* circadian clock. *Planta* **218**, 159–162.

Forger, D. B. & Peskin, C. S. 2003 A detailed predictive model of the mammalian circadian clock. *Proc. Natl Acad. Sci. USA* **100**, 14 806–14 811.

Garceau, N. Y., Liu, Y., Loros, J. J. & Dunlap, J. C. 1997 Alternative initiation of translation and time-specific phosphorylation yield multiple forms of the essential clock protein FREQUENCY. *Cell* **89**, 469–476.

Gau, D., Lemberger, T., von Gall, C., Kretz, O., Minh, N. L., Gass, P., Schmid, W., Schibler, U., Korf, H. W. & Schutz, G. 2002 Phosphorylation of CREB Ser142 regulates light-induced phase shifts of the circadian clock. *Neuron* **34**, 245–252.

Glossop, N. R. J., Lyons, L. C. & Hardin, P. E. 1999 Interlocked feedback loops within the *Drosophila* circadian oscillator. *Science* **286**, 766–768.

Goldbeter, A. 2002 Computational approaches to cellular rhythms. *Nature* **420**, 238–245.

Gonze, D., Halloy, J. & Gaspard, P. 2002 Biochemical clocks and molecular noise: theoretical study of robustness factors. *J. Chem. Phys.* **116**, 10 997–11 010.

Goodwin, B. C. 1965 Oscillating behaviour in enzymatic control processes. In *Advances in enzyme regulation* (ed. G. Weber), vol. 3, pp. 435–438. Oxford: Pergamon Press.

Guckenheimer, J. & Holmes, P. 1983 *Nonlinear oscillations, dynamical systems and bifurcations of vector fields.* Applied Mathematical Sciences. New York: Springer.

Johnson, C. H., Elliott, J. A. & Foster, R. 2003 Entrainment of circadian programs. *Chronobiol. Int.* **20**, 741–774.

Lee, K., Loros, J. J. & Dunlap, J. C. 2000 Interconnected feedback loops in the Neurospora circadian system. *Science* **289**, 107–110.

Leloup, J. C. & Goldbeter, A. 2003 Toward a detailed computational model for the mammalian circadian clock. *Proc. Natl Acad. Sci. USA* **100**, 7051–7056.

Leloup, J. C., Gonze, D. & Goldbeter, A. 1999 Limit cycle models for circadian rhythms based on transcriptional regulation in Drosophila and Neurospora. *J. Biol. Rhythms* **14**, 433–448.

Oster, H., Werner, C., Magnone, M. C., Mayser, H., Feil, R., Seeliger, M. W., Hofmann, F. & Albrecht, U. 2003 cGMP-dependent protein kinase II modulates mPer1 and mPer2 gene induction and influences phase shifts of the circadian clock. *Current Biol.* **13**, 725–733.

Oster, H., Yasui, A., van der Horst, G. T. & Albrecht, U. 2002 Disruption of mCry2 restores circadian rhythmicity in mPer2 mutant mice. *Genes Dev.* **16**, 2633–2638.

Pittendrigh, C. S. & Daan, S. 1976 A functional analysis of circadian pacemakers in nocturnal rodents. V. A clock for all seasons. *J. Comput. Physiol.* A **106**, 333–355.

Preitner, N., Damiola, F., Lopez-Molina, L., Zakany, J., Duboule, D., Albrecht, U. & Schibler, U. 2002 The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251–260.

Press, W. H., Flanner, B. P., Teukolsky, S. A. & Vetterling, W. T. 1988 *Numerical recipes in C.* Cambridge: Cambridge University Press.

Reddy, A. B., Field, M. D., Maywood, E. S. & Hastings, M. H. 2002 Differential resynchronisation of circadian clock gene expression within the suprachiasmatic nuclei of mice subjected to experimental jet lag. *J. Neurosci.* **22**, 7326–7330.

Rensing, L. & Ruoff, P. 2002 Temperature effect on entrainment, phase shifting, and amplitude of circadian clocks and its molecular bases. *Chronobiol. Int.* **19**, 807–864.

Roenneberg, T. & Merrow, M. 2003 The network of time: understanding the molecular circadian system. *Current Biol.* **13**, R198–R207.

Ruoff, P. & Rensing, L. 1996 The temperature-compensated Goodwin model simulates many circadian clock properties. *J. Theor. Biol.* **179**, 275–285.

Ruoff, P., Vinsjevik, M., Monnerjahn, C. & Rensing, L. 2001 The Goodwin model: simulating the effect of light pulses on the circadian sporulation rhythm of *Neurospora Crassa. J. Theor. Biol.* **209**, 29–42.

Ruoff, P., Vinsjevik, P. M. & Rensing, L. 2000 Temperature compensation in biological oscillators: a challenge for joint experimental and theoretical analysis. *Comm. Theor. Biol.* **5**, 361–382.

Smolen, P., Baxter, D. A. & Byrne, J. H. 2001 Modeling circadian oscillations with interlocking positive and negative feedback loops. *J. Neurosci.* **21**, 6644–6656.

Tyson, J. J., Hong, C. I., Thron, C. D. & Novak, B. 1999 A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM. *Biophys. J.* **77**, 2411–2417.

Ueda, H. R., Hagiwara, M. & Kitano, H. 2001 Robust oscillations within the interlocked feedback model of *Drosophila* circadian rhythm. *J. Theor. Biol.* **210**, 401–406.

Young, M. W. & Kay, S. A. 2001 Time zones: a comparative genetics of circadian clocks. *Nature Rev. Genetics* **2**, 702–715.

Zheng, B., Albrecht, U., Kaasik, K., Sage, M., Lu, W., Vaishnav, S., Li, Q., Sun, Z. S., Eichele, G., Bradley, A. & Lee, C. C. 2001 Nonredundant roles of the mPer1 and mPer2 genes in the mammalian circadian clock. *Cell* **105**, 683–694.

The supplementary Electronic Appendix is available at http://dx.doi.org/10.1098/rsif.2004.0014 or via http://www.journals.royalsoc.ac.uk.