Extending Phase Reduction to Excitable Media: Theory and Applications*

Dan Wilson[†]
Jeff Moehlis[†]

Abstract. Phase reduction methods have been tremendously useful for understanding the dynamics of nonlinear oscillators, but have been difficult to extend to systems with a stable fixed point, such as an excitable system. Using the notion of isostables, which measure the time it takes for a given initial condition in phase space to approach a stable fixed point, we present a general method for isostable reduction for excitable systems. We also devise an adjoint method for calculating infinitesimal isostable response curves, which are analogous to infinitesimal phase response curves for oscillatory systems. Through isostable reduction, we are able to implement sophisticated control strategies in a high-dimensional model of cardiac activity for the termination of alternans, a precursor to cardiac fibrillation.

1. Introduction. Phase reduction methods have been fruitfully applied to many physical, chemical, and biological systems [51], [52], [20], [28], [11], [22], [29], [2], [33], [50], [49]. Such methods are useful for understanding the dynamics of perturbed nonlinear oscillators because they allow for the reduction from a system

$$\dot{\mathbf{x}} = F(\mathbf{x}) + G(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^N, \tag{1.1}$$

where $\mathbf{x} \in \mathbb{R}^N$ is a vector of states and $G \in \mathbb{R}^N$ is the effect from an external stimulus, to a system with a single variable,

$$\dot{\theta} = \omega + Z(\theta)^T G(\mathbf{x}(\theta), t), \quad \theta \in \mathbb{S}^1. \tag{1.2}$$

Here $\theta \in [0, 2\pi)$ describes the oscillator's phase, ω is the oscillator's natural frequency, and $Z(\theta) \in \mathbb{R}^N$ is the oscillator's infinitesimal phase response curve (iPRC), which describes the change in the oscillator's phase from a small external stimulus. Phase reduction is particularly powerful with control frameworks such as Hamilton–Jacobi–Bellman, calculus of variations, or dynamic programming [27], since the computational effort of finding the desired control input grows exponentially with the number

†Department of Mechanical Engineering, University of California, Santa Barbara, CA 93106 (ddwilson@umail.ucsb.edu, moehlis@engineering.ucsb.edu).

of state variables. Phase reduction methods have also been useful for systems for which the state dynamics are not fully known. For example, in vitro experiments on biological tissue have successfully controlled spike timing in periodically firing neurons using inputs computed from phase models [44], [34]. The calculation of iPRCs is central to the utility of phase reduction. iPRCs can be calculated using a direct method [24], which entails perturbing the system at a given phase and measuring the resulting phase change, or numerically using a so-called "adjoint method" [11]; cf. [3]. Previously, the notions of phase and iPRCs were extended to excitable systems with a stable fixed point [23], [40], but those methods relied on the existence of slow-fast dynamics and required response curves to be calculated with a grid-based approach.

In this paper, we propose a more general method of phase reduction for excitable systems based on the notion of isostables [31]. Here, a system is said to be excitable if all initial conditions within a small neighborhood of a stable fixed point give trajectories that decay directly to it, but some initial conditions further away from the stable fixed point give trajectories that undergo large excursions before approaching the fixed point. Isostables of excitable systems are analogous to isochrons for asymptotically periodic systems. Isochrons can be understood as an extension of the phase of oscillation to the basin of attraction of the limit cycle [52]; they represent surfaces of constant phase. In this way, two locations in phase space which start on the same isochron will asymptotically approach the same location on a periodic orbit as time tends toward infinity. For excitable systems, it is not possible define a phase because there is no periodic orbit. Instead, isostables can be defined as sets of points in phase space that approach a fixed point together, in a well-defined sense described below. Figure 1.1 highlights the similarities and differences between isochrons for a two-dimensional system with a periodic orbit (left) and isostables for an excitable system with a stable fixed point (right). The left panel shows a system with equations found in [26], [37], and two initial conditions which start on the same isochron asymptotically converge to the same phase on the limit cycle. The right panel shows an excitable system (3.1); initial conditions which start on the same isostable approach the stable fixed point together.

The calculation of an isostable field, $\mathcal{I}(\mathbf{x})$, exploits the linear nature of the dynamics near a fixed point $\mathbf{x}_o$. For a linear system,

$$(1.3) \qquad\qquad \dot{\mathbf{x}} = A(\mathbf{x} - \mathbf{x_o}),$$

and solutions $\phi(t, \mathbf{x}(0))$ (also known as the flow) approach the fixed point as

$$(1.4) \qquad\qquad \phi(t, \mathbf{x}(0)) - \mathbf{x}_o = \sum_{j=1}^{n} s_j(\mathbf{x}(0)) \mathbf{v}_j e^{\lambda_j t},$$

where $s_j(\mathbf{x})$ are the coordinates of the vector $\mathbf{x}$ in the basis $\{\mathbf{v}_j, \; j = 1, \ldots, n\}$ of unit eigenvectors of A, with associated eigenvalues $\{\lambda_j, \; j = 1, \ldots, n\}$, sorted so that λ_1 corresponds to the unique slowest direction of the stable manifold, i.e., $\text{Re}(\lambda_j) < \lambda_1 < 0 \; \forall \; j > 1$. Here, we assume that λ_1 is real and unique and, as shown in [31], the magnitude of $s_1(x)$ determines the infinite time approach to the origin. In other words, hyperplanes of constant isostables, $\mathcal{I}_\tau \equiv \{\mathbf{x} \in \mathbb{R}^n | \mathcal{I}(\mathbf{x}) = \tau\}$, *near a fixed point* are parallel to the faster directions $\mathbf{v}_2, \ldots, \mathbf{v}_n$.

For nonlinear systems, the calculation of isostable fields within the fixed point's entire basin of attraction, $\mathcal{I}(\mathbf{x})$, can be calculated by monitoring the infinite time approach of $\phi(t, \mathbf{x})$ to the fixed point, $\mathbf{x}_o$, by computing

$$(1.5) \qquad\qquad \mathcal{I}(\mathbf{x}) = \lim_{t \to \infty} e^{-\lambda_1 t} ||\phi(t, \mathbf{x}) - \mathbf{x}_o||,$$
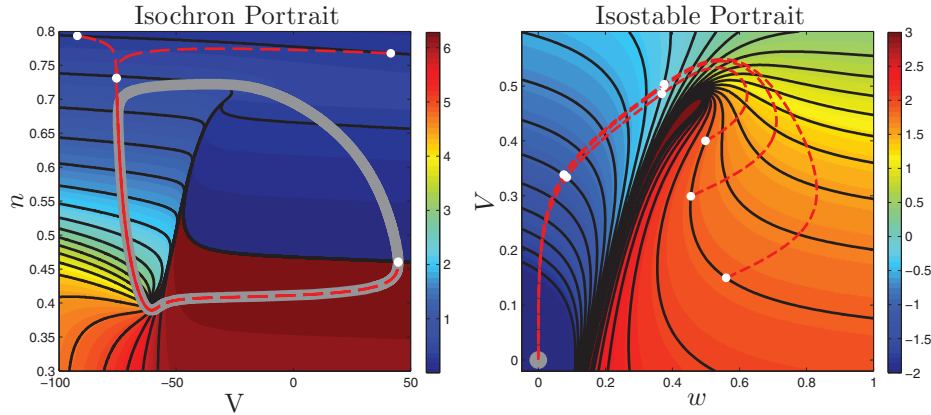
Fig. 1.1 *Isochrons and isostables in two different systems. Isochrons can be used to analyze a system with a periodic orbit, while isostables are appropriate to analyze systems with a stable fixed point. The left panel shows isochrons for a system with a stable periodic orbit shown as the gray curve, with the phase ranging from 0 to 2π. Black lines show isochrons equally spaced in time in the sense that the time required to go from one of the displayed isochrons to the next is always the same. Two trajectories shown as red, dashed lines start on the same isochron and approach the periodic orbit with the same phase. Snapshots are shown as white dots. The right panel shows an isostable field for a system with a stable fixed point shown as the gray dot, and black lines show isostables which are equally spaced in time. Three trajectories shown as red, dashed lines starting on the same isostable will approach the fixed point together. Snapshots are shown as white dots. Isostables are calculated on a grid of initial conditions using (1.5). Isochrons are calculated using a similar grid-based approach, but can also be calculated in finer detail using methods detailed in [37].*

where $|| \cdot ||$ can be any norm but for this paper we will be working with the 1-norm on $\mathbb{R}^n$. Intuitively, (1.5) compares the asymptotic approach to the fixed point along the slowest direction of the stable manifold, $\mathbf{v}_1$, to an exponential function governed by the associated eigenvalue, λ_1. We emphasize that (1.5) is valid for systems with a stable fixed point where λ_1 is real and unique. In other cases, such as when λ_1 is complex or the fixed point is unstable, isostables can still be calculated, and we refer the interested reader to [31] for a more complete discussion.

As mentioned previously, phase reduction is particularly useful to make complicated problems in nonlinear dynamics analytically and computationally tractable. In this paper, we choose a problem from cardiology to which we will apply the aforementioned isostable reduction, but first we give a brief overview of the heart from a dynamical systems point of view.

The heart beats roughly three billion times during a typical human lifetime to transport blood containing oxygen, nutrients, immune cells, regulatory molecules, and waste products to keep the body functioning. The heart is driven by a small area of specialized cardiac muscle cells known as the sinoatrial node, which acts as pacemaker, producing a propagating wave of electrical activity which spreads through the four chambers of the heart (right atrium, left atrium, right ventricle, and left ventricle) in a coordinated fashion. From a modeling perspective, the propagation of electrical activity throughout the heart is often viewed as a reaction-diffusion equation on a one-, two-, or three-dimensional representation of the *myocardium*, the muscular tissue of the heart. This cardiological medium is made up of individual cells, or *myocytes*, which are specialized muscle tissue that coordinate the action of a heart beat.

Myocytes are excitable cells: a large enough positive perturbation to the transmembrane voltage of a myocyte will elicit a depolarizing action potential and a contraction of the muscle tissue [52]. Cardiac action potentials last for a significant amount of time, allowing the elevated transmembrane voltage to spread to nearby cells and creating a wave of action potentials analogous to the way dominoes fall [13]. After this wave propagates through the myocardium, under normal conditions the cells repolarize and remain quiescent until the pacemaker elicits another pulse.

Ideally, this healthy pattern of coordinated activity occurs throughout the host's lifespan. However, under certain conditions, the myocardium can become susceptible to *cardiac fibrillation*, the uncoordinated contraction of the cardiac muscle, which can be deadly in an otherwise healthy heart if not treated within minutes of onset. Most researchers agree that atrial fibrillation is caused by the presence of unwanted spiral waves [52] within the myocardium which interfere with the normal sinoatrial rhythm [19], [12], [6].

Cardiac fibrillation can be the final step in an increasingly complex series of events that begins with a phenomenon known as cardiac *alternans*. Cardiac alternans is the beat-to-beat alternation of electrochemical cardiac dynamics at a constant rate of pacing and has been implicated as a precursor to fibrillation [43]. These alternations include the strength of cardiac muscle contraction, action potential duration (APD), and intracellular calcium concentration dynamics. In particular, the phenomenon known as *discordant alternans*, where cardiac tissue in different spatial locations exhibit alternans of APD of opposite phase, can produce particularly favorable conditions for initiating the spiral wave reentry required for fibrillation [38], [39], [46], [47], [18]. Spiral waves are more likely to be initiated, for instance, with a premature beat when there is a large degree of spatial heterogeneity in the refractory cells in the heart. For a review of alternans from a dynamical systems perspective, we refer the reader to [7].

In recent decades, researchers have looked into ways of suppressing alternans as a way of preventing cardiac fibrillation, bypassing the need for painful and damaging defibrillating shocks. Many of these methods, such as those described in [41] and [8], attempt to suppress alternans by monitoring the APD or the length of time that the tissue remains depolarized (i.e., at elevated voltage) and appropriately modifying the time at which the next action potential is elicited. These methods have been shown to be successful in real cardiac tissue [21], [9], [25].

Using isostable reduction techniques we are able to work in a more useful coordinate system to devise a novel and energy-efficient approach to terminating alternans. In section 2 we develop the mathematical framework required for an isostable reduction for excitable systems (analogous to phase reduction for periodic systems) and the corresponding infinitesimal isostable response curve (iIRC) (analogous to the iPRC for periodic systems). We also develop an "adjoint method" to calculate an iIRC, requiring only the knowledge of the trajectory along which the system approaches the fixed point. Section 3 illustrates the isostable reduction for a relatively simple, two-dimensional model of excitability. We devise an energy-optimal control technique to eliminate cardiac alternans in section 4, and illustrate this technique for the 13-dimensional Fox model of canine cardiac activity [15] in section 5. We find numerically in section 6 that we can eliminate alternans using many orders of magnitude less energy than with a pulsatile strategy, which may be important from a clinical perspective to maximize the battery life of an implantable device and limit tissue damage. We give concluding remarks in section 7.

2. Infinitesimal Isostable Response Curves. Following the derivation of the adjoint method for calculating iPRCs set forth in, for example, [3], we consider a general

N-dimensional differential equation

$$\dot{\mathbf{x}} = F(\mathbf{x}) + G(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^N, \tag{2.1}$$

where $F(\mathbf{x})$ is the vector field, $G(\mathbf{x}, t)$ is an external stimulus, and $x_1, x_2, \ldots, x_N$ are the state variables. For a given set of initial conditions, suppose that the system follows the known trajectory γ to the stable fixed point $\mathbf{x}_o$.

Our objective is to simplify (2.1) to a one-dimensional equation by defining scalar isostable coordinates, $\psi(\mathbf{x}) \in (-\infty, \infty]$, for all $\mathbf{x}$ in some neighborhood U of $\mathbf{x}_o$ within its basin of attraction. It will be convenient to take $\psi(\mathbf{x}) = -\log(\mathcal{I}(\mathbf{x}))$, where $\mathcal{I}(\mathbf{x})$ is defined in (1.5). Changing to isostable coordinates, from the chain rule we find

$$\frac{d\psi(\mathbf{x})}{dt} = \nabla\psi(\mathbf{x}) \cdot (F(\mathbf{x}) + G(\mathbf{x}, t)). \tag{2.2}$$

In order to simplify (2.2), first consider any trajectory $\phi(t, \mathbf{x})$ in the basin of attraction of the fixed point for which the external stimulus G is set to zero. Using (1.5), we have

$$\frac{d\psi(\mathbf{x})}{dt} = \lim_{\Delta t \to 0} \frac{\psi(\phi(t + \Delta t, \mathbf{x})) - \psi(\phi(t, \mathbf{x}))}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{-\log\left[\lim_{t \to \infty}\left(e^{-\lambda_1 t}(\phi(t+\Delta t, \mathbf{x}) - \mathbf{x}_o)\right)\right] + \log\left[\lim_{t \to \infty}\left(e^{-\lambda_1 t}(\phi(t, \mathbf{x}) - \mathbf{x}_o)\right)\right]}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{-\log\left[\vartheta e^{-\lambda_1 t} e^{\lambda_1 (t + \Delta t)}\right] + \log\left[\vartheta e^{-\lambda_1 t} e^{\lambda_1 t}\right]}{\Delta t}$$
$$= -\lambda_1,$$
$$\tag{2.3}$$

where ϑ is a positive constant determined by the direction of $\mathbf{v}_1$. In the second to last line, we use the fact that the trajectory will eventually approach the fixed point along the slowest direction of the stable manifold, and the last line is obtained solely through algebraic manipulation. Recall that (2.3) was derived for $G = \mathbf{0}$, which, using (2.2), implies that

$$\nabla\psi(\mathbf{x}) \cdot F(\mathbf{x}) = \omega \tag{2.4}$$

and thus

$$\frac{d\psi(\mathbf{x})}{dt} = \omega + \nabla\psi(\mathbf{x}) \cdot G(\mathbf{x}, t), \tag{2.5}$$

where $\omega = -\lambda_1$ (recall that λ_1 is negative) and is similar to the natural frequency term in the classic phase reduction (1.2). We note that the isostable field can be scaled by a constant if desired, which will yield a different value of ω. In the absence of external stimuli, $\frac{d\psi}{dt} = \omega$, i.e., $\psi(\mathbf{x})$ increases at a constant rate. By definition, $\psi(\mathbf{x}) = \infty$ corresponds to $\mathbf{x} = \mathbf{x}_o$, meaning that in the absence of external control, all trajectories in the domain of attraction of $\mathbf{x}_o$ approach the fixed point in infinite time.

Evaluating the vector field at $\mathbf{x}^\gamma(\psi)$, which we define as the intersection of the trajectory γ and the $\psi(\mathbf{x})$ level set (i.e., isostable), we have

$$\frac{d\psi(\mathbf{x})}{dt} = \omega + \nabla\psi(\mathbf{x}^\gamma(\psi)) \cdot G(\mathbf{x}^\gamma(\psi), t), \tag{2.6}$$

where we have dropped an error term of order $|G|^2$ (cf. [29]), so that (2.6) is valid for perturbations with small $|G|$. For phase reductions of systems with limit cycles, the analogue of the term $\nabla\psi(\mathbf{x}^\gamma(\psi))$ can be calculated experimentally using the "direct method" [24] as well as the numerically using the "adjoint method" [11]; cf. [3]. Here we detail an extension of the adjoint method for calculation of $\nabla\psi(\mathbf{x}^\gamma(\psi))$ for systems with a stable fixed point.

Suppose for the following analysis that $G = \mathbf{0}$ for $t > 0$. Consider an infinitesimal perturbation $\Delta\mathbf{x}$ to the trajectory $\mathbf{x}(t) \in \gamma$ at time $t = 0$. Let $\mathbf{x}_\epsilon(t) = \mathbf{x}^\gamma(t) + \Delta\mathbf{x}(t)$ be the trajectory starting at the perturbed initial condition. Then

$$(2.7) \qquad \frac{d\Delta\mathbf{x}(t)}{dt} = DF(\mathbf{x}(t))\Delta\mathbf{x}(t) + \mathcal{O}(||\Delta\mathbf{x}||^2),$$

where DF is the Jacobian matrix. Furthermore, for the associated isostable shift defined as $\Delta\psi = \psi(\mathbf{x}_\epsilon(t)) - \psi(\mathbf{x}(t))$, we have

$$(2.8) \qquad \Delta\psi = (\nabla_{\mathbf{x}(t)}\psi)^T \cdot \Delta\mathbf{x}(t) + \mathcal{O}(||\Delta\mathbf{x}||^2),$$

where $\nabla_{\mathbf{x}(t)}\psi$ is the gradient of ψ evaluated at $\mathbf{x}(t)$. Let $\langle\cdot,\cdot\rangle$ denote the standard Euclidean inner product on $\mathbb{R}^n$. Following the derivation presented in [3], after the initial perturbation at $t = 0$, $\Delta\psi$ is independent of time. Therefore, taking the time derivative of (2.8) yields, to lowest order in $||\Delta\mathbf{x}||$,

$$\left\langle \frac{d\nabla_{\mathbf{x}(t)}\psi}{dt}, \Delta\mathbf{x}(t) \right\rangle = -\left\langle \nabla_{\mathbf{x}(t)}\psi, \frac{d\Delta\mathbf{x}(t)}{dt} \right\rangle,$$

$$= -\langle \nabla_{\mathbf{x}(t)}\psi, DF(\mathbf{x}(t))\Delta\mathbf{x}(t) \rangle,$$

$$(2.9) \qquad = -\langle DF^T(\mathbf{x}(t))\nabla_{\mathbf{x}(t)}\psi, \Delta\mathbf{x}(t) \rangle.$$

The matrix $DF^T(\mathbf{x}(t))$ is the transpose, or adjoint, of the real-valued matrix $DF(\mathbf{x}(t))$. Equation (2.9) holds for arbitrary perturbations $\Delta\mathbf{x}(t)$, which gives

$$(2.10) \qquad \frac{d\nabla_{\mathbf{x}(t)}\psi}{dt} = -DF^T(\mathbf{x}(t))\nabla_{\mathbf{x}(t)}\psi.$$

In order to use (2.6) we need the isostable gradient at the intersection of γ and the $\psi(\mathbf{x})$ level set, but (2.10) gives the isostable gradient along γ. However, the one can readily be inferred from the other with knowledge of the map $\mathbf{x}(t) \rightarrow \psi(t)$.

The solution of (2.10) requires N initial conditions. Recalling that $G = \mathbf{0}$ for $t > 0$ in the present analysis, we note that, close to the fixed point, we can rewrite (2.1) as

$$(2.11) \qquad \dot{\mathbf{x}} = A(\mathbf{x} - \mathbf{x}_o) + \mathcal{O}((\mathbf{x} - \mathbf{x}_o)^2),$$

where $A = DF(\mathbf{x}_o)$. Neglecting higher order terms, (2.11) is linear. We assume that each eigenvalue $\lambda_i = \sigma_i + i\kappa_i$ of A has negative real part, geometric multiplicity of one, and corresponds to the right eigenvector $\mathbf{v}_i$. For convenience, we sort the eigenvalues so the real part of λ_1 has the smallest magnitude, i.e., $|\sigma_1| < |\sigma_2| \leq \cdots \leq |\sigma_N|$. Note that $|\sigma_1|$ must be strictly less than $|\sigma_2|$ and that $\frac{dx(t)}{dt} \cdot \mathbf{v}_1 \neq 0$ near the fixed point for the following analysis. From linear systems theory, we know that trajectories $\phi(t, \mathbf{x}(0))$ will approach the fixed point in a manner given previously by (1.4), with $s_j(\mathbf{x}(0)) = \langle \mathbf{x}(0), \mathbf{v}_j \rangle$. The infinite time approach to the fixed point will be governed
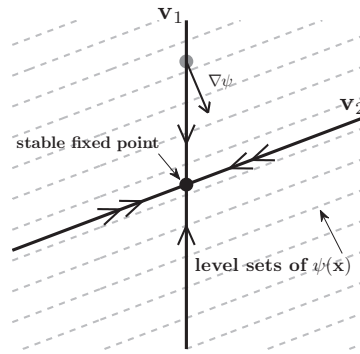
Fig. 2.1 *Level sets of $\psi(\mathbf{x})$ for a two-dimensional system are shown as dashed lines. The faster and slower directions of the stable manifold are shown with two and one arrows, respectively.*

by the value of $s_1(\mathbf{x}(0))$, and any perturbation to $s_j(\mathbf{x}(0))$ for $j \geq 2$ will not have any effect on the trajectory's infinite time approach to the origin. Thus,

$$(2.12) \qquad \langle \nabla_{\mathbf{x}(t)} \psi, \mathbf{v}_j \rangle = 0 \qquad \forall \, j \geq 2.$$

Equation (2.12) mandates that, close to the fixed point, $\nabla_{\mathbf{x}(t)} \psi$ must be orthogonal to the subspace spanned by $\{\mathbf{v}_2, \ldots, \mathbf{v}_n\}$; in other words, $\nabla_{\mathbf{x}(t)} \psi$ is orthogonal to hyperplanes of constant isostables. Because the geometric multiplicity for each eigenvector is one, this defines a unique direction for $\nabla_{\mathbf{x}(t)} \psi$ near the fixed point. The isostable portrait is summarized in Figure 2.1 for a two-dimensional system. The magnitude can be obtained by noting that (2.5) must hold at all points along $\mathbf{x}(t)$. Using (2.12) and (2.5), we can choose any point $\mathbf{x}_1$ along γ, close enough to the fixed point that the system is described accurately by its linearization, to serve as an appropriate end point condition to (2.10) and solve by integrating backward in time to determine $\nabla \psi$ at all points close to γ. For convenience, we will scale the isostable coordinates $\psi(\mathbf{x})$ so that $\psi(\mathbf{x}_1) = 1$, which means that ω in (2.6) is determined by the time it takes to reach $\psi = 1$ from $\psi = 0$, where $\psi = 0$ is chosen to be another location on γ as convenient. Intuitively, $\nabla \psi$ does not exist for $\psi = \infty$ since any perturbation from the fixed point will take an infinitely long time to reach the fixed point again. We can, however, evaluate $\nabla \psi$ arbitrarily close to the fixed point.

We note that the preceding analysis is valid for stable fixed points with strictly real λ_1. If λ_1 is complex, the iIRC can still be calculated by a similar analysis, but instead of $\nabla \psi$ being orthogonal to the slower directions of the stable manifold near the fixed point, as in (2.12), $\nabla \psi$ will be perpendicular to a cylindrical hypersurface of constant isostable. We refer the interested reader to [31] for a discussion of isostables for the case when λ_1 is complex.

3. Illustration for a Simple FitzHugh–Nagumo-Based Model. We first illustrate the calculation of $\nabla \psi(\mathbf{x}^\gamma(\psi))$, which we will refer to as the infinitesimal isostable response curve (iIRC), for FitzHugh–Nagumo-based equations as a prototypical model for excitability [42]:

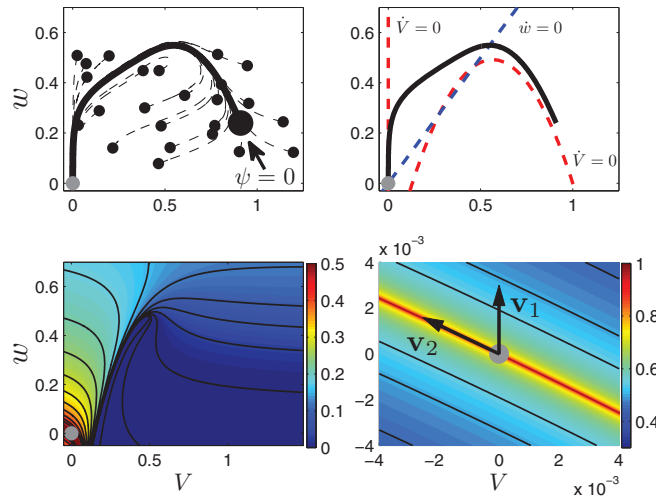$$(3.1) \qquad \begin{aligned} \dot{V} &= c_1 V(V-a)(1-V) - c_2 V w, \\ \dot{w} &= b(V - dw). \end{aligned}$$

Fig. 3.1   *In the top-left panel, trajectories of* (3.1) *follow a similar path toward the fixed point; this path is shown as a bold line. The top-right panel shows the V and w nullclines of* (3.1), *as dashed lines. The bottom-left panel gives a global isostable portrait for* (3.1) *with black lines showing constant isostables, equally spaced in time. The bottom-right panel shows a close-up near the fixed point where the system is well-approximated by a linear system and the isostables are parallel to $\mathbf{v}_2$, the fastest direction of the stable manifold. In all panels, the stable fixed point is shown as a gray dot.*

Nondimensionalized variables V and w represent, for example, voltage and gating variables, respectively, of an excitable cell membrane. We will take $a = 0.13$, $b = 0.013$, $c_1 = 0.26$, $c_2 = 0.1$, and $d = 1$. We can see from Figure 3.1 that trajectories follow a similar path as they approach the stable fixed point $(V, w) = (0, 0)$; [23] and [40] would call this a transient attractor. When the system is linearized around this fixed point, we find $|\lambda_1| = 0.0130$ and $|\lambda_2| = 0.0338$. The top-right panel of Figure 3.1 shows nullclines for this system. The bottom-left panel shows the global isostable portrait for (3.1), and the bottom-right panel gives a zoomed-in isostable portrait near the fixed point.

We choose $(V, w) \approx (0.91, 0.24)$ to correspond to $\psi = 0$ and integrate forward until the final condition is close enough to the origin for the linearization to be a good model. The initial condition is chosen because most trajectories that start around $(V, w) \approx (1, 0.1)$ pass close to this point, as shown in the top-left panel of Figure 3.1. For this model, integrating for 1000 time units is sufficient, and we take the resulting trajectory to be γ and the end point condition to correspond to $\psi = 1$. We use γ to calculate the iIRC for (3.1) using the adjoint method presented in section 2. For this choice of γ, $\omega = \frac{1}{1000}$. Results are shown in Figure 3.2. The top-right panel shows γ, with the bottom-right panels showing $V(\psi)$ and $w(\psi)$. The left panels show iIRCs for each variable at different levels of zoom. As might be expected, as the trajectory approaches the fixed point, the magnitude of the iIRC becomes very large, because even small perturbations will significantly affect the asymptotic approach to the fixed point. Bottom-left panels show zoomed-in plots of the iIRC as well as numerical verification at various points, for example, by perturbing by ΔV at a given value of ψ and measuring the resulting difference $\Delta \psi$. The iIRC can then be approximated by $\frac{\partial \psi}{\partial V} \approx \frac{\Delta \psi}{\Delta V}$.
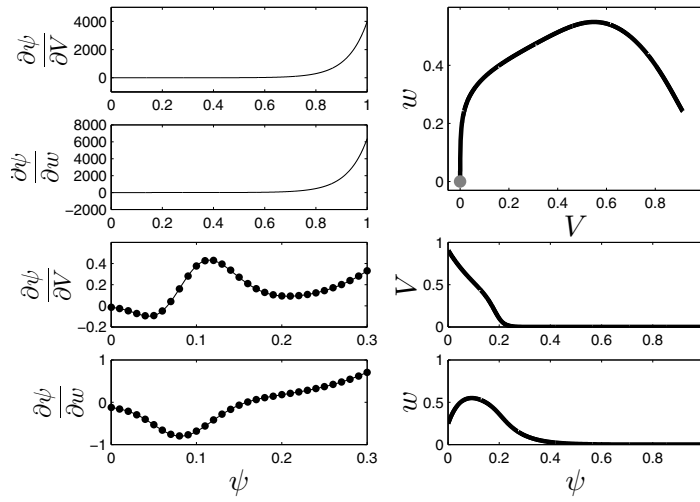
Fig. 3.2 *Calculation of the iIRC for the FitzHugh–Nagumo-based model using the adjoint method. The top-right panel shows γ, with the bottom-right panels showing $V(\psi)$ and $w(\psi)$. iIRCs calculated using the adjoint method are given in the left panels, with dots calculated according to $\frac{\Delta\psi}{\Delta V}$ and $\frac{\Delta\psi}{\Delta w}$ for a small but finite ΔV and Δw.*

4. An Energy-Optimal Control Strategy for Terminating Cardiac Alternans.

As mentioned in section 1, cardiac alternans, the beat-to-beat alternation between long and short action potentials, has been implicated as a precursor to cardiac fibrillation. In order to understand how alternans develop in cardiac tissue, one must examine the relationship between action potential duration (APD), diastolic interval (DI), and basic cycle length (BCL) in cardiac cells. The APD is defined as the amount of time an action potential lasts, i.e., the amount of time the cell remains depolarized (i.e., at elevated voltage) after it is excited, and the DI is the amount of time the cell remains quiescent before the next action potential. The BCL is the amount of time between successive action potentials and can be thought of as a constant "heart rate," for example, as generated by pacemaker cells. The APD of a cardiac cell is typically a monotonically increasing function of DI, and the left panel of Figure 4.1 shows the APD plotted against the DI (commonly referred to as the APD restitution curve) for the Fox model of cardiac activity [15], which we will consider further in sections 5 and 6; the solid curve uses nominal parameters, and the dashed curve is obtained when the value of the L-type Ca^{2+} channel permeability to Ca^{2+} is reduced by 20 percent. As originally shown in [36], when the APD restitution curve has a slope greater than 1, alternans can develop in cardiac tissue. To understand how alternans may develop, we first examine the dashed APD restitution curve, which has a slope that is strictly less than one. The plot in the bottom-right panel of Figure 4.1 uses this curve and can be interpreted as follows. Suppose, for a constant BCL of 185 ms, an action potential is elicited when the DI is around 42 ms. From the APD restitution curve we can see that the following APD will be around 115 ms. Because BCL = APD + DI, the next DI will be about 70 ms. The arrows show the result of this procedure over multiple iterations, and the successive iterations spiral toward an APD of about 125 ms. Conversely, the solid APD curve in the left panel of Figure 4.1 has a slope larger than one in places, and the top-right panel shows that for the same
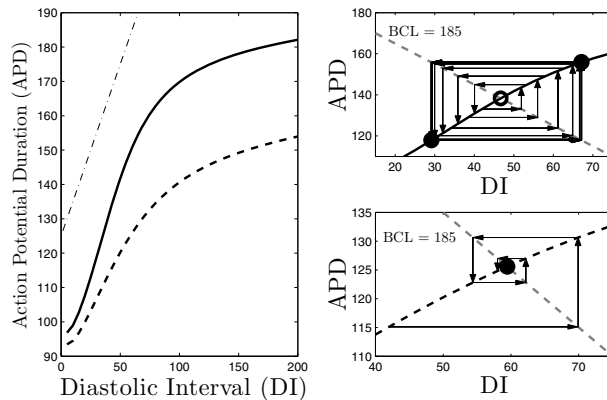
Fig. 4.1 *The left panel shows APD restitution curves for two different parameter sets for the Fox model and a dot-dashed line with a slope of one for reference. The solid APD restitution curve has slope greater than one in places, and the associated top-right panel shows how a stable period-2 orbit develops when the BCL is a constant 185 ms (represented by the gray dashed line plotting 185 = DI + APD). For the dashed APD restitution curve, with slope strictly less than one, the period-1 orbit is stable for a constant BCL of 185 ms as shown in the bottom-right panel.*

BCL of 185 ms, a stable orbit develops with APDs alternating between 118 and 156 ms; this is called a period-2 orbit because it takes two BCLs to repeat. Note that in the top-right panel, every fifth transition is shown for clarity; also note that there is an unstable period-1 orbit, meaning that it repeats every BCL, at the intersection of APD restitution curve and the dashed line, represented by an open circle.

As might be expected, the APD restitution curve is not strictly a function of the preceding DI. The APD can be affected by other things such as pacing history or interactions between neighboring cells [45], [5], [4]. Nevertheless, the goal of "flattening" the APD restitution curve by reducing its slope provides a basis for developing drug interventions that make it less likely for alternans and cardiac fibrillation to develop [16], [1], [30].

Other strategies to terminate alternans have focused on stabilizing the unstable period-1 rhythm through the application of external stimuli [8], [41] and have been useful in designing experimentally successful, model-independent control techniques [21], [9], [25]. Some other control techniques have attempted to control alternans using model-based feedback control [10], [17]. While model-independent control techniques are certainly more directly applicable to in vivo experiments, they must sacrifice knowledge of the underlying dynamics which makes them, most likely, not optimal from an energy perspective. On the other hand, model-based control strategies which use feedback control make use of the underlying dynamics of the system but can be hard to implement accurately in real time experiments, especially when the underlying models for the experiment are of high dimension. Here, we propose an energy-optimal, model-based control approach that only requires knowledge of a system's iIRC. Control techniques based on response curve reduction are useful in high order models because they do not require the full dynamics, only the one-dimensional dynamics near a trajectory. Furthermore, iPRCs have been successfully measured for neurons for control purposes [34], and we expect these methods to be readily adaptable to measuring iIRCs for cardiac systems.

In a common experiment to test control algorithms for alternans, in vitro cardiac tissue is excited by an artificial pacemaker at a prescribed BCL which is known to produce APD alternans. The APD of each beat is measured as the length of its action potential. Previous, model-independent methods to suppress alternans have operated by adjusting the pacemaker's cycle length according to the following control algorithm (cf. [9], [25]):

$$
(4.1) \qquad T_n = \begin{cases} \mathrm{BCL} + \Delta T_n & \text{if } \Delta T_n < 0, \\ \mathrm{BCL} & \text{if } \Delta T_n \geq 0, \end{cases}
$$

where

$$
(4.2) \qquad \Delta T_n = (\beta/2)(\mathrm{APD}_n - \mathrm{APD}_{n-1}),
$$

T_n is the control-perturbed (external) pacemaker cycle length, β is the feedback gain, and n is the beat number. In this article, we use $\beta = 1$. Typically, as the DI increases, so will the following APD, and, intuitively, this control strategy works by giving a premature pulse after the short APD, reducing the DI of the upcoming long APD and subsequently increasing the DI of the next short APD. See panels C and E of Figure 6.1 for a numerical example of the control strategy from (4.1) and (4.2) applied to a single cardiac cell.

We now devise a new alternans suppression methodology in the context of the 13-dimensional Fox model for canine cardiac activity [15]. We emphasize that the following results are not limited to this model, and we expect that this control strategy could be applied successfully to other models which display alternans. For a BCL of 175 ms, cardiac cells from the Fox model will display beat-to-beat alternans with successive APD differences of approximately 36 ms. Figure 4.2 shows the associated stable (solid line) and unstable (dashed line) periodic orbits in the $V - [\mathrm{Ca}^{2+}]_i$ plane,
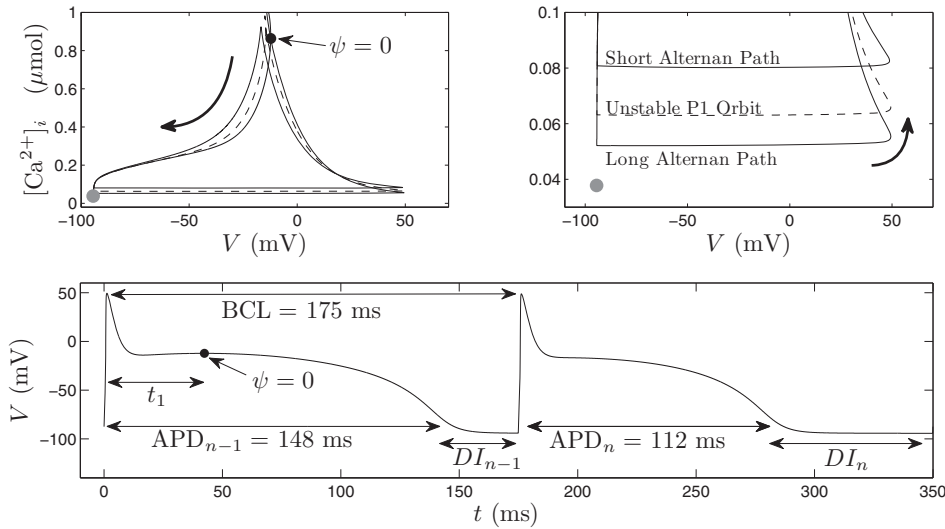

Fig. 4.2 *Unstable period-1 orbit (dashed line) and stable period-2 orbit (solid line) in the $V-[\mathrm{Ca}^{2+}]_i$ plane for the Fox model. In the top panels, curved arrows show the direction of the orbits and gray dots show the location of the stable fixed point.*

where V is the transmembrane voltage and $[\text{Ca}^{2+}]_i$ is the intracellular Ca^{2+} concentration. We note that the control strategy from (4.1) is sufficient to suppress alternans in this model, but our goal is to devise a control strategy that accomplishes the same objective while using less energy and without artificially modifying T_n. To this end, following [33] we adapt the energy-optimal methodology proposed for spike-timing control in neurons for use in cardiac cells.

Consider the isostable model for a cardiac cell

$$\dot{\psi} = \omega + \nabla \psi \cdot G(t). \tag{4.3}$$

We assume that we only have control over the voltage variable, which we take to be the first of the 13 components of $\mathbf{x}$, so that $G(t) = [u(t), 0, \ldots, 0]^T$, where $u(t) = i(t)/C$ represents the control input, $i(t)$ is the applied current, and C is the cell capacitance. Equation (4.3) becomes

$$\dot{\psi} = \omega + \frac{\partial \psi}{\partial V}(\psi)u(t) = \omega + I(\psi)u(t), \tag{4.4}$$

where $I \equiv \frac{\partial \psi}{\partial V}$. Note that explicit $\mathbf{x}$ dependence has been dropped for notational simplicity. While cardiac alternans is caused by a period-2 rhythm in the presence of periodic forcing, the iIRC must be calculated by removing the forcing and letting the state approach a fixed point. Recall from section 2 that the values $\psi = 0$ and $\psi = 1$ can be chosen as is convenient to define the iIRC. In a numerical setting, we can take $\psi = 1$ as sufficiently close to the fixed point, but in an experimental setting, this may not be possible. Therefore, we take $I(\psi) = 0 \ \forall \ \psi > 1$.

Now that the iIRC has been found, we return to the case of periodic forcing. We define $t = 0$ to be the time that the present action potential started and t_1 to be the time when $\psi(\mathbf{x}(t_1))$ next reaches 0. Suppose that, at a time determined by the BCL (when the next action potential will start), for all stimuli $u(t)$ which evolve $\psi(t_1) = 0$ to $\psi(\text{BCL}) = \alpha > 0$ for some phase α which will be chosen later, we want to find the stimulus $u(t)$ which minimizes the cost function $\mathcal{M}[u(t)]^2 = \int_{t_1}^{\text{BCL}} u^2(t) \ dt$, that is, the power associated with the stimulus. (Other costs including the derivative of the current and Faradaic charge transfer [32] can also be handled by other suitable cost functions.) This is now in the form of an optimal control problem. We apply calculus of variations to minimize [14],

$$\mathcal{C}[u(t)] = \int_{t_1}^{\text{BCL}} \underbrace{\left[u^2(t) + \lambda_1 \left(\frac{d\psi}{dt} - \omega - I(\psi)u(t) \right) \right]}_{\mathcal{M}[u(t)]} dt, \tag{4.5}$$

where λ_1 is a Lagrange multiplier, forcing the dynamics to satisfy (4.4). Associated Euler–Lagrange equations are

$$u(t) = \lambda_1(t)I(\psi)/2, \tag{4.6}$$

$$\dot{\psi} = \omega + (\lambda_1 I^2(\psi))/2, \tag{4.7}$$

$$\dot{\lambda_1} = -\lambda_1^2 I(\psi)I'(\psi)/2, \tag{4.8}$$

where $' = d/d\psi$. The optimal control $u(t)$ can be found by solving (4.7) and (4.8) subject to $\psi(t_1) = 0$, $\psi(\text{BCL}) = \alpha$. This requires finding the initial condition $\lambda_1(t_1)$ which satisfies the boundary conditions, e.g., numerically with a shooting method. To

examine the uniqueness of such an optimal control, we first note that the Hamiltonian for this system,

$$(4.9) \qquad H_o = \lambda_1^2 I^2(\psi)/4 + \lambda_1 \omega,$$

is conserved along trajectories of (4.7) and (4.8). With this in mind, we find that

$$\psi(\mathrm{BCL}) = \int_{t_1}^{\mathrm{BCL}} \dot\psi \, dt = \int_{t_1}^{\mathrm{BCL}} \frac{\omega + \lambda_1 I^2(\psi(t))}{2} dt = \int_{t_1}^{\mathrm{BCL}} \sqrt{\omega^2 + I^2(\psi)H_o}\, dt,$$

$$(4.10)$$

$$\frac{\partial \psi(\mathrm{BCL})}{\partial H_o} = \int_{t_1}^{\mathrm{BCL}} \frac{I^2(\psi(t))}{2\sqrt{\omega^2 + I^2(\psi)H_o}} dt.$$

Note that in order for $\psi(\mathrm{BCL})$ to be real-valued, $H_o > -\omega^2/\sup[I^2(\psi)]$. Therefore, for all valid choices of H_o, $\frac{\partial \psi(\mathrm{BCL})}{\partial H_o} > 0$. Thus $\psi(\mathrm{BCL})$ increases monotonically with H_o and implies that if a solution to (4.7) and (4.8) exists, it must be unique.

To implement the control strategy, we choose α to be

$$(4.11) \qquad \alpha^* \equiv \omega(\mathrm{BCL} - t_1) - \omega(\mathrm{APD}_n - \mathrm{APD}_{n-1})/2.$$

This particular choice for α^* is explained in Figure 4.3, which shows zoomed-in orbits from Figure 4.2. We note that our proposed control strategy is functionally identical to the strategy in (4.1) and (4.2), but the means of achieving the goal are different. In (4.1) and (4.2), the time of the next pulse is modulated, which requires a large amount of energy to elicit a premature pulse, while our new control strategy modifies the cell's
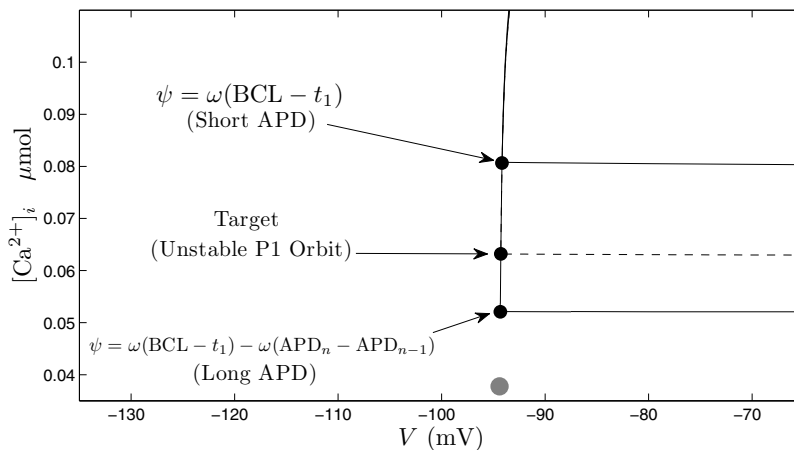

Fig. 4.3 *Illustration of the alternans termination control strategy showing zoomed-in orbits from Figure 4.2. Assuming the next action potential will be a long APD, the trajectory will reach the short APD curve at exactly $t = \mathrm{BCL}$, and the ensuing APD will be short. Advancing the phase by an extra $-\omega(\mathrm{APD}_n - \mathrm{APD}_{n-1})$ will cause the ensuing APD to be long again. The control strategy is only valid until the cell's next action potential at $t = \mathrm{BCL}$. At this point, the time can be reset to $t = 0$ and, if alternans persist, the control can be reapplied in order to gradually drive the system to the unstable period-1 orbit. For reference, the stable fixed point is shown as a gray dot.*

phase when the next pulse arrives. Assume the next action potential will be a long APD; in the absence of control, the trajectory will reach the location corresponding to the short APD at precisely $t = \text{BCL}$, and a short APD will occur afterwards. If we wanted to drive the trajectory to another long APD, we would need to advance ψ by an extra $-\omega(\text{APD}_n - \text{APD}_{n-1})$ with an external control (recall that before a long action potential, $\text{APD}_n - \text{APD}_{n-1} < 0$). Note that the above control strategy is only valid until the cell's next action potential at $t = \text{BCL}$. At this point, the time can be reset to $t = 0$ and, if alternans persist, the control can be reapplied in order to gradually drive the system to the unstable period-1 orbit. We emphasize that it is implicit in the preceding analysis that the state of the system remains close to the trajectory γ from which the iIRC was derived. If the applied control creates a large excursion from γ, it will degrade the performance of the controller and, in extreme cases, can elicit a premature excitation.

5. Optimal Control Illustrated for the Fox Model of Cardiac Activity. We now illustrate the optimal control strategy for terminating cardiac alternans for the 13-dimensional Fox model [15]. The Fox model is particularly attractive because it considers ionic currents within canine cardiac cells and produces action potentials that are consistent with experimental observations. Furthermore, it produces stable electrical alternans at fast BCLs, making it ideal to test our optimal control strategy. This model has one slowly changing variable, the sarcoplasmic reticulum Ca^{2+} concentration, $[\text{Ca}^{2+}]_{\text{SR}}$, which does not change much on a beat-to-beat basis for BCLs in which the model exhibits alternans. This allows for the reduction of the dimensionality through averaging [48], assuming that $[\text{Ca}^{2+}]_{\text{SR}} = 318\,\mu\text{mol}$. For reference, when the reduced system is linearized around its stable fixed point, $|\lambda_1| = 0.0015$ and $|\lambda_2| = 0.0043$.

The Fox model exhibits alternans for BCLs between 150 and 210 ms. Using a Poincaré map, shown in Figure 5.1, we visualize the approach γ toward the fixed point for many action potentials. We stimulate the reduced model for BCLs of 150, 160, 170, 180, 190, 200, and 210 and record the state during repolarization when V crosses -40,
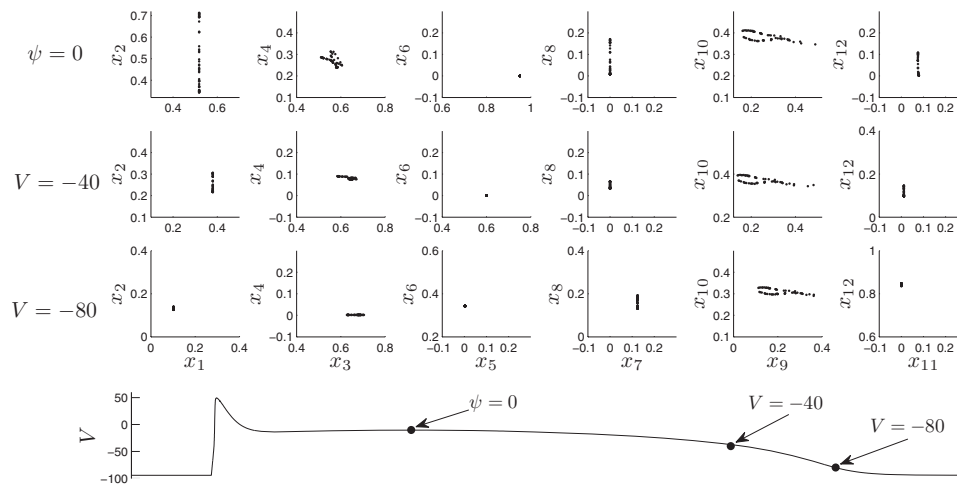

Fig. 5.1 *State variables in the Fox model at various points during the long alternan action potential, stimulated using multiple BCLs that produce alternans. Trajectories become closer as the transmembrane voltage approaches the resting potential of about $-94\ mV$.*

and -80, as well as when $\frac{dV}{dt} = 0$. Variables $(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12})$ correspond to $(V, [\text{Ca}^{2+}]_i, f, d, m, h, j, f_{Ca}, X_{\text{Kr}}, X_{\text{Ks}}, X_{\text{to}}, Y_{\text{to}})$ from the reduced model. Note that each variable has been normalized as the fraction of its maximum value during stimulation; see [15] for the definitions of these variables and equations for their dynamics.

In order to implement the control strategy from section 4, we restrict our attention to the relaxation to the fixed point. The reason for this is two-fold: first, as shown in Figure 4.3, the state variables follow a similar trajectory toward the fixed point during repolarization so that the iIRCs will be expected to be similar on a beat-to-beat basis. Second, the initial upstroke of the action potential occurs on a rapid time scale, which would make it difficult to apply the control at the precise time it is required. During the last moments of repolarization as the transmembrane potential approaches its resting potential of about -94 mV, the beat-to-beat trajectories tend to follow a similar path, but it is not as clearly defined as it is for the FitzHugh–Nagumo-based model (3.1). There is some natural variance in the preferred approach in the Fox model, and for this reason we calculate the iIRC for multiple action potentials, which in the absence of an external depolarizing stimulus take different paths toward the fixed point. We define the state at which V is largest during its final approach to the fixed point on the long alternan as $\psi = 0$ (i.e., $\mathbf{x}^\gamma(0) \equiv \mathbf{x} \mid \frac{dV}{dt} = 0$ and $\frac{dV}{dt} < 0 \ \forall \ \psi \geq 0$) and simulate for 2000 ms to obtain $\mathbf{x}^\gamma$ for each approach toward the fixed point. We emphasize that there are multiple possible trajectories toward the stable fixed point, and the curves $\mathbf{x}^\gamma$ between trials are different. Results are shown in Figure 5.2. The black line gives the mean iIRC for the voltage variable, and the gray band gives an interval in which 90 percent of calculated iIRCs reside. We find that perturbing later in the cycle will have a much smaller effect than perturbing earlier. Note that the iIRCs for the other 11 variables were calculated but are not shown.

Using average values of $\frac{\partial \psi}{\partial V}$ from Figure 5.2, we employ the control strategy on the phase reduction (4.4) and illustrate for a BCL of 175 ms the cycle length at which alternans are most pronounced. Also, we only apply control on the long alternan because the iIRC from Figure 5.2 was calculated for long alternan trajectories. Results are shown in Figure 5.3. The top panel shows the voltage trace of the cell during a
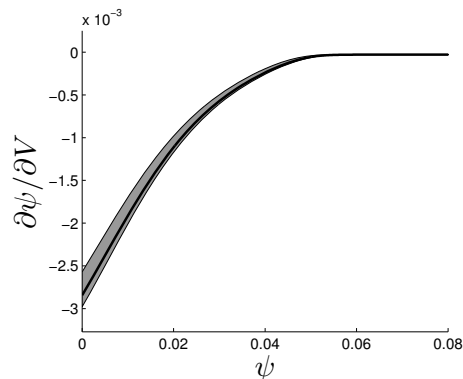

Fig. 5.2 *iIRC for the transmembrane voltage in the Fox model reduction using the adjoint method. Ninety percent of values lie within the gray band, with the mean value plotted as a thick solid line. The transmembrane voltage is within 95 percent of its resting potential at $\psi \approx 0.025$.*
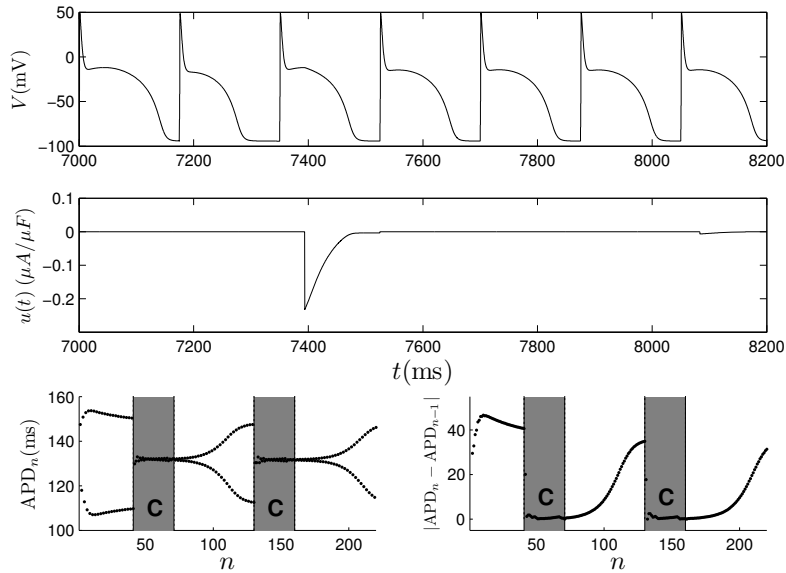
Fig. 5.3 *Control of alternans in the Fox model. The top and middle panels show the voltage traces for the cell and the applied control, respectively. The bottom two panels show plots of APD and $|\mathrm{APD}_n - \mathrm{APD}_{n-1}|$, with the shaded sections showing when the controller is on. The control strategy is quickly able to eliminate alternans in the model. Note that the control applied at $t \approx 8100$ is much smaller in magnitude than the first control application. Also, the action potential that begins at $t \approx 7350$ corresponds to APD_{42}.*

time when the controller is turned on. We see that the control eliminates alternans by the next action potential. We note that once the alternans are eliminated, it takes relatively little control effort compared to the first control application to keep the cell in an alternan-free trajectory. The bottom two panels give plots of the beat-to-beat action potentials, as well as $|\mathrm{APD}_n - \mathrm{APD}_{n-1}|$. Times when the controller is turned on are denoted by shaded regions marked with a "C." When the controller is on, the cell remains alternan-free. When the controller is turned off, the cell gradually displays alternans that become more and more pronounced, which are quickly alleviated once the controller is turned back on.

6. Alternan Control with iIRCs Determined by the Direct Method. A marked advantage of this control strategy over other model-based strategies is that it only requires knowledge of the iIRC of the system. In neuroscience, methods that rely on analogous iPRCs are attractive because the iPRC can be approximated experimentally using the direct method [24], [35], [34]. We note that the direct method can be used to measure the iPRC along the entire periodic orbit in neuroscience applications, because a neural spike, which can be measured as a sudden rise in the transmembrane voltage, is a clear marker of a particular phase. In cardiological applications, the primary observable (transmembrane voltage) only changes during repolarization, after which another observable would be necessary to obtain a full iIRC. Nevertheless, here we devise a feasible means of using the direct method to obtain at least part of a cardiac cell's iIRC for control purposes.

First, $\psi = 0$ is defined to be some marker voltage during repolarization; we will illustrate this using the Fox model and, as above, we take the marker $\psi = 0$ to

be the final local maximum during repolarization, i.e., at $\mathbf{x}$ for which $\frac{dV}{dt} = 0$ and $\frac{dV}{dt} < 0 \ \forall \ \psi \geq 0$. Note that the phase resets upon each APD. Further, we define $\psi = 1$ to be when the transmembrane voltage reaches 95 percent of its resting potential. The value ω can be obtained by determining $t_{0 \to 1}$, the average time it takes a cell to travel from $\psi = 0$ to $\psi = 1$. For the Fox model, we average over BCLs ranging from 150 to 210 to find $t_{0 \to 1} = 105.06$ ms, and we calculate $\omega = 1/t_{0 \to 1}$. In order to obtain one data point for the iIRC, a short-duration pulse of current can be applied to a cardiac cell at a random isostable ψ and the resulting isostable change is measured by observing the change in the time it takes to reach $\psi = 1$. The resulting value $I(\psi)$ is $\frac{\Delta \psi}{Q_p/C}$, where Q_p is the total charge injected through the cell membrane and $\Delta \psi$ is the change in isostable. An experimentally reasonable sampling size of 300 data points was calculated according to $I(\psi) = \frac{\Delta \psi}{Q_p/C} + 0.025 \mathcal{N}(0,1)$, where uncertainty is included to mimic the undesired, but unavoidable noise in experimental measurements. The data was fit to a sixth order polynomial, as is customary in calculation of iPRCs for neurons [35], [34], and constrained so that $I(1) = 0$ and $\frac{dI}{d\psi}(0) = 0$. Note that these constraints are consistent with the qualities of the iIRC shown in Figure 5.2. Recall that we assume $I(\psi) = 0$ for $\psi > 1$, and we enforce these extra constraints so that the iIRC is smooth in at least the first derivative at $\psi = 1$.

In an experimental setting, an additional concern is that a control strategy that must calculate a new optimal control on every application would be difficult to implement, because by the time the optimal control is calculated, the window to implement it would have already passed. Indeed, a MATLAB routine on a good workstation takes on the order of one second to solve the boundary value problem given by (4.7) and (4.8), and the processing power of an implantable device will certainly be less than that of a desktop computer. For this reason, we seek an approximate solution to the boundary value problem when $\alpha \approx \omega(\text{BCL} - t_1)$. First, notice that the line $\lambda_1 = 0$ is invariant for (4.7)–(4.8) and corresponds to $\frac{d\psi}{dt} = \omega$, thus $\alpha = \omega(\text{BCL} - t_1)$. For $\alpha \approx \omega(\text{BCL} - t_1)$, we Taylor expand $\psi(\text{BCL})$ with respect to initial condition $\lambda_1(0)$, yielding $\alpha = \omega(\text{BCL} - t_1) + \left(\frac{\partial \alpha}{\partial \lambda_1(0)}\big|_{\lambda_1(0)=0}\right) \lambda_1(0)$ to lowest order in $\alpha - \omega(\text{BCL} - t_1)$. Recalling that $\alpha = \psi(\text{BCL})$ and using α^* from (4.11) for α, we have

$$(6.1) \qquad \lambda_1(0) = \frac{\alpha - \omega(\text{BCL} - t_1)}{\frac{\partial \alpha}{\partial \lambda_1(0)}\big|_{\lambda_1(0)=0}} = \frac{-\omega(APD_n - APD_{n-1})}{2 \frac{\partial \psi(\text{BCL})}{\partial \lambda_1(0)}\big|_{\lambda_1(0)=0}}.$$

Let $APD_n - APD_{n-1} = \mathcal{O}(\epsilon)$. Asymptotically expanding $\lambda_1(t)$ and $\psi(t)$ yields

$$(6.2) \qquad \lambda_1(t) = \lambda_1^{(0)}(t) + \epsilon \lambda_1^{(1)}(t) + \epsilon^2 \lambda_1^{(2)}(t) + \cdots,$$

$$(6.3) \qquad \psi(t) = \psi^{(0)}(t) + \epsilon \psi^{(1)}(t) + \epsilon^2 \psi^{(2)}(t) + \cdots.$$

From (4.7) and (4.8), we see that $\lambda_1^{(0)}(t) = 0$, $\lambda_1^{(1)}(t) = \lambda_1(0)/\epsilon + \mathcal{O}(\epsilon)$, and $\psi^{(0)}(t) = \omega(t - t_1)$. In order to further simplify (6.1), we use the asymptotic expansion from (6.2) and recall that $\lambda_1(0)$ is of order ϵ to yield

$$(6.4)$$

$$\frac{\partial \psi(\text{BCL})}{\partial \lambda_1(0)}\bigg|_{\lambda_1(0)=0} = \lim_{\lambda_1(0) \to 0} \frac{\int_{t_1}^{\text{BCL}} [\lambda_1(0) + \mathcal{O}(\epsilon^2)](I^2(\psi)/2)dt}{\lambda_1(0)} = \int_{t_1}^{\text{BCL}} \frac{I^2(\psi)}{2} dt.$$

Using (6.1), this implies that $\lambda_1(0) = \frac{-\omega(APD_n - APD_{n-1})}{\int_{t_1}^{\text{BCL}} I^2(\psi)dt}$. Finally, substituting (6.2)

and (6.3) into (4.6), we find

$$u(t) = \frac{1}{2}\lambda_1(0)I(\psi^{(0)}) + \mathcal{O}((\text{APD}_n - \text{APD}_{n-1})^2)$$

$$(6.5) \qquad = \frac{-\omega(\text{APD}_n - \text{APD}_{n-1})I(\omega(t - t_1))}{2\int_{t_1}^{\text{BCL}} I^2(\omega(\tau - t_1))d\tau} + \mathcal{O}((\text{APD}_n - \text{APD}_{n-1})^2).$$

Using the iIRC computed using the direct method on noisy data, we illustrate the use of the *approximation* to the optimal control strategy (6.5) and compare to the method from (4.1). For the pulsed control strategy, as in [15] we use a $-80\ \mu A/\mu F$ pulse of duration 1 ms to elicit a premature action potential and note that for pulse intensities much below this, the pulse will not always elicit an action potential. Further, we take the threshold for the control applications to be $|\text{APD}_n - \text{APD}_{n-1}| \geq 5$. We note that both control strategies work equally well with a smaller threshold, but we choose 5 in order to better illustrate small differences between strategies. We choose a BCL of 175 ms and begin applying control after $n = 30$ action potentials for a total of 150 action potentials. Results are shown in Figure 6.1. Panel A shows each data point in the numerical calculation of the iIRC, as well as the sixth order polynomial fit (black line), and the iIRC calculated from the adjoint method (gray dashed line), scaled by $\omega_{dir}/\omega_{adj}$, where ω_{dir} and ω_{adj} are the natural frequency for the iIRC calculated for the direct method and adjoint method, respectively. Panels B and D show a portion of the applied optimal control and the associated APD history, respectively. In panel B, the first applied control is the largest, with the following control applications being smaller in magnitude, because required alterations to the
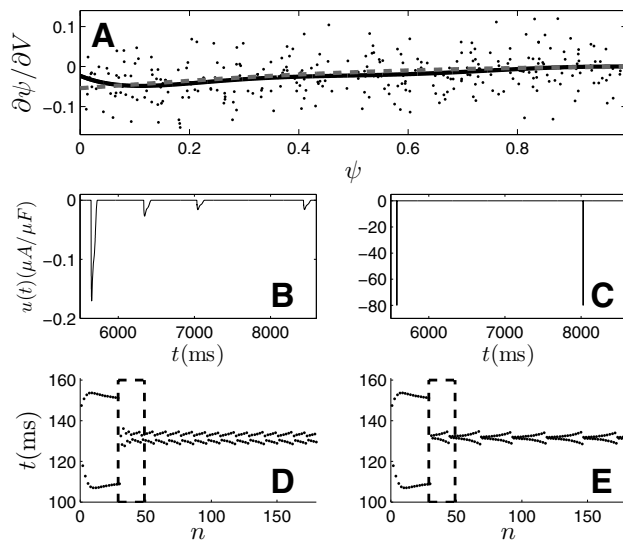

Fig. 6.1 *Comparison of a pulsed versus approximately optimal control strategy. Panel A shows the polynomial fit iIRC (black line) along with an appropriately scaled iIRC calculated from the adjoint method (gray dashed line). Panel B (resp., C) gives a portion of the applied optimal (resp., pulsatile) control associated with the APD history in the dashed rectangular box from Panel D (resp., E). Panel D (resp., E) gives the APD history using the optimal (resp., pulsatile) control strategy.*

action potential are much smaller. While the approximately optimal control strategy does not perform quite as well as before, because we are using both a polynomial fit for $\frac{\partial \psi}{\partial V}$ and an approximation of the optimal control, it is still sufficient to terminate alternans. Panels C and E show a portion of the pulsed control and the associated APD history, respectively. The most notable difference between these control strategies is the energy consumption as measured by $\int u(t)^2 dt$. The energy-optimal control strategy uses approximately 0.7097 units of energy, while the pulsed control strategy uses 44800 units. We emphasize that this is not actual experimental data and that we are neglecting the power consumption of the processor required to implement our control logic. Nevertheless, our numerical simulations suggest that the optimal control strategy may be able terminate alternans at a cellular level using *several orders of magnitude less energy* than a pulsed method.

7. Conclusion. This paper extends the notion of phase reduction to systems with a stable fixed point through the development of an adjoint method for the efficient calculation of infinitesimal isostable response curves (iIRCs). This method of reduction is particularly useful for excitable systems that take similar trajectories toward their fixed point, so that the iIRCs between trajectories are similar. This novel method of phase reduction has been illustrated for both a simple, two-dimensional model of excitability and for a more complicated, higher-dimensional model. Sections 4–6 of this paper display the strength of this new reduction method, as it allows us to work with a complicated, high-dimensional model in a more useful coordinate system.

With only the knowledge of a system's iIRC, one can design a control strategy to eliminate cardiac alternans, which has been implicated as a precursor to fibrillation, that outperforms a pulsatile, model-independent strategy by several orders of magnitude in terms of energy. Even though the processor for a complicated control strategy will likely use more energy than a processor for a simple controller, there is still the potential for significant energy savings compared to a pulsatile strategy. While model-independent control strategies are attractive from an experimental standpoint, forfeiture of the state dynamics almost certainly comes at the price of using a nonoptimal control strategy in terms of overall energy consumption, as we have shown here. Furthermore, the proposed control strategy does not require premature excitation of cardiac cells to eliminate alternans, but rather gently perturbs cells during an action potential to eliminate alternans.

Certainly there are other considerations that need to be addressed before this control strategy can be used in real experiments. For instance, it takes a nontrivial amount of time to calculate the optimal control for a given sequence of action potentials, and the processor must be reasonably quick at calculating the control and determining the precise time at which it should be applied. However, some of these challenges could be addressed by using the approximation to the optimal control derived in section 6. Also, it may be challenging to both stimulate cardiac tissue and record APDs simultaneously. Nevertheless, we believe that this control strategy holds promise for successful implementation in real cardiac systems.

Phase reduction is particularly attractive when the full dynamics of a system are unknown, as is the case for biological tissue in a experimental setting. Control strategies based on phase reduction have been successful in the control of in vitro neurons [34], [44], and we expect that similar techniques to experimentally measure iIRCs will carry over to excitable biological systems. We imagine that this control strategy could ultimately be applied to hearts in vivo with a grid of independent pairs of sensors and actuators, each using this control strategy to locally eliminate alternans.

While not as elegant as a strategy that eliminates alternans from a single point source control, a control strategy that employs reduction using iIRCs to view cardiac action potentials in a more natural coordinate system allows for lower-amplitude, lower-energy perturbations, and thus may be desirable from a clinical standpoint in terms of maximizing the battery life of an implantable device and mitigating the risk of tissue damage.

Phase reduction has proven to be a powerful tool for the analysis of complex systems with periodic orbits and has led to many interesting and important results. This new method of isostable reduction could be of similar use by allowing for the study of complicated excitable systems in a more natural and convenient coordinate system.

REFERENCES

[1] K. E. BRACK, V. H. PATEL, J. H. COOTE, AND G. A. NG, *Nitric oxide mediates the vagal protective effect on ventricular fibrillation via effects on action potential duration restitution in the rabbit heart*, J. Physiol., 583 (2007), pp. 695–704.

[2] E. BROWN, P. HOLMES, AND J. MOEHLIS, *Globally coupled oscillator networks*, in Perspectives and Problems in Nonlinear Science, E. Kaplan, J. C. Marsden, and R. Sreenivasan, eds., Springer, New York, 2003, pp. 183–215.

[3] E. BROWN, J. MOEHLIS, AND P. HOLMES, *On the phase reduction and response dynamics of neural oscillator populations*, Neural Comput., 16 (2004), pp. 673–715.

[4] E. M. CHERRY AND S. J. EVANS, *Properties of two human atrial cell models in tissue: Restitution, memory, propagation, and reentry*, J. Theoret. Biol., 254 (2008), pp. 674–690.

[5] E. M. CHERRY AND F. H. FENTON, *Suppression of alternans and conduction blocks despite steep APD restitution: Electrotonic, memory, and conduction velocity restitution effects*, Amer. J. Physiol.-Heart Circulatory Physiol., 55 (2004), pp. H2332–H2341.

[6] E. M. CHERRY AND F. H. FENTON, *Visualization of spiral and scroll waves in simulated and experimental cardiac tissue*, New J. Phys., 10 (2008), 125016.

[7] E. M. CHERRY, F. H. FENTON, AND R. F. GILMOUR, JR., *Mechanisms of ventricular arrhythmias: A dynamical systems-based perspective*, Amer. J. Physiol.-Heart Circulatory Physiol., 302 (2012), pp. H2451–H2463.

[8] D. J. CHRISTINI AND J. J. COLLINS, *Using chaos control and tracking to suppress a pathological nonchaotic rhythm in a cardiac model*, Phys. Rev. E, 53 (1996), pp. R49–R52.

[9] D. J. CHRISTINI, M. L. RICCIO, C. A. CULIANU, J. J. FOX, A. KARMA, AND R. F. GILMOUR, JR., *Control of electrical alternans in canine cardiac Purkinje fibers*, Phys. Rev. Lett., 96 (2006), 104101.

[10] S. DUBLJEVIC, *Optimal boundary control of cardiac alternans*, Internat. J. Robust Nonlinear Control, 19 (2009), pp. 135–150.

[11] G. B. ERMENTROUT AND N. KOPELL, *Multiple pulse interactions and averaging in systems of coupled neural oscillators*, J. Math. Biol., 29 (1991), pp. 195–217.

[12] F. FENTON, E. CHERRY, H. HASTINGS, AND S. EVANS, *Multiple mechanisms of spiral wave breakup in a model of cardiac electrical activity*, Chaos, 12 (2002), pp. 852–892.

[13] F. H. FENTON, E. M. CHERRY, AND L. GLASS, *Cardiac arrhythmia*, Scholarpedia, 3 (2008), p. 1665.

[14] D. B. FORGER AND D. PAYDARFAR, *Starting, stopping, and resetting biological oscillators: In search of optimal perturbations*, J. Theoret. Biol., 230 (2004), pp. 521–532.

[15] J. J. FOX, J. L. MCHARG, AND R. F. GILMOUR, JR., *Ionic mechanism of electrical alternans*, Amer. J. Physiol.-Heart Circulatory Physiol., 282 (2002), pp. H516–H530.

[16] A. GARFINKEL, Y. H. KIM, O. VOROSHILOVSKY, Z. QU, J. R. KIL, M. H. LEE, H. S. KARAGUEUZIAN, J. N. WEISS, AND P. S. CHEN, *Preventing ventricular fibrillation by flattening cardiac restitution*, Proc. Natl. Acad. Sci. USA, 97 (2000), pp. 6061–6066.

[17] A. GARZÓN, R. O. GRIGORIEV, AND F. H. FENTON, *Model-based control of cardiac alternans in Purkinje fibers*, Phys. Rev. E, 84 (2011), 041927.

[18] A. GIZZI, E. M. CHERRY, R. F. GILMOUR, JR., S. LUTHER, S. FILIPPI, AND F. H. FENTON, *Effects of pacing site and stimulation history on alternans dynamics and the development of complex spatiotemporal patterns in cardiac tissue*, Frontiers Physiol., 4 (2013), 71.

[19] R. A. GRAY, A. M. PERTSOV, AND J. JALIFE, *Spatial and temporal organization during cardiac fibrillation*, Nature, 392 (1998), pp. 75–78.

[20] J. GUCKENHEIMER, *Isochrons and phaseless sets*, J. Math. Biol., 1 (1975), pp. 259–273.
[21] K. HALL, D. J. CHRISTINI, M. TREMBLAY, J. J. COLLINS, L. GLASS, AND J. BILLETTE, *Dynamic control of cardiac alternans*, Phys. Rev. Lett., 78 (1997), pp. 4518–4521.
[22] F. C. HOPPENSTEADT AND E. M. IZHIKEVICH, *Weakly Connected Neural Networks*, Springer, New York, 1997.
[23] N. ICHINOSE, K. AIHARA, AND K. JUDD, *Extending the concept of isochrons from oscillatory to excitable systems for modeling an excitable neuron*, Internat. J. Bifur. Chaos, 8 (1998), pp. 2375–2385.
[24] E. M. IZHIKEVICH, *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*, MIT Press, Cambridge, MA, 2006.
[25] U. B. KANU, S. IRAVANIAN, R. F. GILMOUR, JR., AND E. J. CHRISTINI, *Control of action potential duration alternans in canine cardiac ventricular tissue*, IEEE Trans. Biomed. Engrg., 58 (2011), pp. 894–904.
[26] J. KEENER AND J. SNEYD, *Mathematical Physiology*, Springer-Verlag, New York, 1998.
[27] D. KIRK, *Optimal Control Theory*, Dover, New York, 1998.
[28] Y. KURAMOTO, *Chemical Oscillations, Waves, and Turbulence*, Springer-Verlag, Berlin, 1984.
[29] Y. KURAMOTO, *Phase-and center-manifold reductions for large populations of coupled oscillators with application to non-locally coupled systems*, Internat. J. Bifur. Chaos, 7 (1997), pp. 789–805.
[30] A. MAHAJAN, D. SATO, Y. SHIFERAW, A. BAHER, L. H. XIE, R. PERALTA, R. OLCESE, A. GARFINKEL, Z. QU, AND J. N. WEISS, *Modifying L-type calcium current kinetics: Consequences for cardiac excitation and arrhythmia dynamics*, Biophys. J., 94 (2008), pp. 411–423.
[31] A. MAUROY, I. MEZIĆ, AND J. MOEHLIS, *Isostables, isochrons, and Koopman spectrum for the action-angle representation of stable fixed point dynamics*, Phys. D, 261 (2013), pp. 19–30.
[32] D. MERILL, M. BIKSON, AND J. JEFFERYS, *Electrical stimulation of excitable tissue: Design of efficacious and safe protocols*, J. Neurosci. Methods, 141(2) (2005), pp. 171–198.
[33] J. MOEHLIS, E. SHEA-BROWN, AND H. RABITZ, *Optimal inputs for phase models of spiking neurons*, ASME J. Comput. Nonlinear Dynam., 1 (2006), pp. 358–367.
[34] A. NABI, T. STIGEN, J. MOEHLIS, AND T. NETOFF, *Minimum energy control for* in vitro *neurons*, J. Neural Engrg., 10 (2013), 036005.
[35] T. NETOFF, M. A. SCHWEMMER, AND T. J. LEWIS, *Experimentally estimating phase response curves of neurons: Theoretical and practical issues*, in Phase Response Curves in Neuroscience, N. W. Schultheiss, A. A. Prinz, and R. J. Butera, eds., Springer, New York, 2012, pp. 95–129.
[36] J. B. NOLASCO AND R. W. DAHLEN, *A graphic method for the study of alternation in cardiac action potentials.*, J. Appl. Physiol., 25 (1968), pp. 191–196.
[37] H. M. OSINGA AND J. MOEHLIS, *Continuation-based computation of global isochrons*, SIAM J. Appl. Dyn. Syst., 9 (2010), pp. 1201–1228.
[38] J. M. PASTORE, S. D. GIROUARD, K. R. LAURITA, F. G. AKAR, AND D. S. ROSENBAUM, *Mechanism linking T-wave alternans to the genesis of cardiac fibrillation*, Circulation, 99 (1999), pp. 1385–1394.
[39] Z. QU, A. GARFINKEL, P. CHEN, AND J. N. WEISS, *Mechanisms of discordant alternans and induction of reentry in simulated cardiac tissue*, Circulation, 102 (2000), pp. 1664–1670.
[40] A. RABINOVITCH AND I. ROGACHEVSKII, *Threshold, excitability and isochrones in the Bonhoeffer–van der Pol system*, Chaos, 9 (1999), pp. 880–886.
[41] W. RAPPEL, F. FENTON, AND A. KARMA, *Spatiotemporal control of wave instabilities in cardiac tissue*, Phys. Rev. Lett., 83 (1999), pp. 456–459.
[42] J. ROGERS AND A. MCCULLOCH, *A collocation-Galerkin finite element model of cardiac action potential propagation*, IEEE Trans. Biomed. Engrg., 41 (1994), pp. 743–757.
[43] D. S. ROSENBAUM, L. E. JACKSON, J. M. SMITH, H. GARAN, J. N. RUSKIN, AND R. J. COHEN, *Electrical alternans and vulnerability to ventricular arrhythmias*, New England J. Medicine, 330 (1994), pp. 235–241.
[44] T. STIGEN, P. DANZL, J. MOEHLIS, AND T. NETOFF, *Controlling spike timing and synchrony in oscillatory neurons*, J. Neurophysiol., 105 (2011), pp. 2074–2082.
[45] E. G. TOLKACHEVA, D. G. SCHAEFFER, D. J. GAUTHIER, AND W. KRASSOWSKA, *Condition for alternans and stability of the* 1:1 *response pattern in a "memory" model of paced cardiac dynamics*, Phys. Rev. E, 67 (2003), 031904.
[46] M. A. WATANABE, F. H. FENTON, S. J. EVANS, H. M. HASTINGS, AND A. KARMA, *Mechanisms for discordant alternans*, J. Cardiovascular Electrophysiol., 12 (2001), pp. 196–206.
[47] J. N. WEISS, A. KARMA, Y. SHIFERAW, P. CHEN, A. GARFINKEL, AND Z. QU, *From pulsus to pulseless: The saga of cardiac alternans*, Circulation Res., 98 (2006), pp. 1244–1253.
[48] S. WIGGINS, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Texts Appl. Math. 2, Springer, New York, 2003.

[49] D. WILSON AND J. MOEHLIS, *Locally optimal extracellular stimulation for chaotic desynchronization of neural populations*, J. Comput. Neurosci., 37 (2014), pp. 243–257.

[50] D. WILSON AND J. MOEHLIS, *Optimal chaotic desynchronization for neural populations*, SIAM J. Appl. Dyn. Syst., 13 (2014), pp. 276–305.

[51] A. T. WINFREE, *Patterns of phase compromise in biological cycles*, J. Math. Biol., 1 (1974), pp. 73–93.

[52] A. T. WINFREE, *The Geometry of Biological Time*, Interdiscip. Appl. Math. 12, Springer, New York, 2001.