



# The effect of spatial scale of resistive inhomogeneity on defibrillation of cardiac tissue

James P. Keener<sup>a,\*</sup>, Eric Cytrynbaum<sup>b</sup>

<sup>a</sup>Department of Mathematics, University of Utah, 155 South 1400 East, Salt Lake City, UT 84112, USA

<sup>b</sup>Institute of Theoretical Dynamics, University of California, Davis, CA 95616, USA

## Abstract

Defibrillation of cardiac tissue can be viewed in the context of dynamical systems theory as the attempt to move a dynamical system from the basin of attraction of one attractor (fibrillation) to another (the uniform rest state) by applying a stimulus whose form is physically constrained. Here we give an introduction to the physical mechanism of cardiac defibrillation from this dynamical perspective and examine the role of resistive inhomogeneity on defibrillation efficacy. Using numerical simulations with rotating waves on a one-dimensional periodic ring, we study the role of the spatial scale of resistive inhomogeneity on defibrillation.

For a rotating wave on a periodic ring there are three stable attractors, namely the uniform rest state, a wave traveling clockwise and a wave traveling counterclockwise. As a result, the application of a stimulus has the potential for three different outcomes, namely elimination of the wave, phase resetting of the wave, and reversal of the wave.

The results presented here show that with resistive inhomogeneities of large spatial scale, all three of these transitions are possible with large amplitude shocks, so that the probability of defibrillation is bounded well below one, independent of stimulus amplitude. On the other hand, resistive inhomogeneities of small spatial scale produce a defibrillation threshold that is qualitatively consistent with that found experimentally, namely the probability of defibrillation success is an increasing function that approaches one for large enough stimulus amplitude.

Extending these results to higher dimensions, we describe conditions for successful defibrillation of functional reentry with large scale spatial inhomogeneity, but find that elimination of anatomical reentry is quite difficult. With small spatial scale inhomogeneity, there are no similar restrictions.

© 2003 Elsevier Science Ltd. All rights reserved.

**Keywords:** Defibrillation; Defibrillation threshold; Resistive inhomogeneities

## 1. Introduction

Fibrillation is generally thought to be a highly disorganized pattern of electrical activation of the heart consisting of many small reentrant “spirals” that are continually created and destroyed (Gray et al., 1995; Panfilov, 1998, 1999; Choi et al., 2002). Ventricular fibrillation is usually self-sustained and unless there is a successful intervention, death is certain. Atrial fibrillation is a similar condition that occurs on the atria but which is not fatal. In both situations, however, it is highly desirable to eliminate the reentrant behavior and restore the normal pattern of activation.

Defibrillation with a large current shock is the process by which fibrillation is usually eliminated. In the typical situation, two conducting pads are placed on the chest (or in the case of open heart surgery or with implantable defibrillators, directly to the surface of the heart) and a short (10 ms) discharge of current is triggered. When applied to the body surface, the energy is of the order of 150 J, which explains why this is called a shock. For implantable defibrillators, the required energy is on the order of 15–20 J, which is still considerable.

When it works, a defibrillating shock eliminates the reentrant activity and the heart approaches its rest state, awaiting normal activation. When it fails, the reentrant activity is temporarily disturbed, but spontaneously returns (Ideker et al., 1991).

While the defibrillation threshold (DFT) is described as a threshold, this is not precisely correct. Rather, the DFT is defined as the shock strength at which 50% of

\*Corresponding author. Tel.: +1-801-581-6089; fax: +1-801-585-1640.

E-mail address: [keener@math.utah.edu](mailto:keener@math.utah.edu) (J.P. Keener).

the attempts are successful at eliminating fibrillation. At higher shock strengths, however, the probability of success increases. For example, the enhanced DFT (DFT++) is 90% effective and for implantable defibrillators the probability of success rises to approximately 99% with an additional increase of 4–6 J (Gold et al., 2002).

The theory to explain the mechanism of defibrillation is not completely resolved. According to the “critical mass hypothesis”, a sufficiently large portion of the tissue must receive a stimulus of sufficient strength for defibrillation to succeed. For example, an oft-mentioned criterion is that 90% of the tissue must have an extracellular field strength of at least 5 V/cm. This criterion does not explain the mechanism of defibrillation since it is not extracellular current, but the distribution of transmembrane current, that is responsible for stimulating cell membrane.

As we discuss below, inhomogeneities of resistance are necessary to drive transmembrane current, and homogeneous ventricles could not be defibrillated. Fortunately for us, there is no such thing as a resistively homogeneous heart, as there are numerous sources of inhomogeneities of resistance. For example, at the cellular level, cells are connected by gap junctions and surrounded by extracellular space that contains capillaries, collagen fiber, connective tissue, etc. all of which contribute inhomogeneity of conductance. In addition, myocytes are assembled in distinct layers, with extensive clefts between these layers (Caulfield and Borg, 1979; Robinson et al., 1983). At a larger space scale, cells are organized into fibers, there is fiber branching and tapering, and the fiber orientation changes both in the longitudinal and in the transverse directions.

The question that arises is which, if any, of these inhomogeneities is primarily responsible for the success of defibrillation. The answer that we have favored is small-scale spatial inhomogeneities, and this hypothesis has been explored in several previous papers (Fishler, 1998; Fishler and Vepa, 1998; Keener, 1996, 1998; Keener and Lewis, 1999; Keener and Panfilov, 1996; Krinsky and Pumar, 1998). Because the largest contribution to small scale inhomogeneities was thought to be gap junctional resistance, and because gap junctional resistance should lead to “sawtooth” profiles of transmembrane potential, this hypothesis is sometimes referred to as the sawtooth hypothesis (Krassowska et al., 1987, 1990). At present, the sawtooth hypothesis is not accepted by many workers in the field, largely because of experimental data suggesting that the amplitude of the sawtooth is too small to be the source of defibrillating stimuli (Gillis et al., 1996; Zhou et al., 1998). However, a new proposal that deserves consideration is that interlaminal clefts provide an adequate small scale resistive inhomogeneity (Hooks et al., 2002).

The alternate hypothesis is that large spatial scale inhomogeneities are primarily responsible for defibrillation success. There is no difficulty whatever to see the effects of large scale inhomogeneities of resistance in experiments (Fast et al., 1998; White et al., 1998). In fact, regions of membrane depolarization and hyperpolarization are easily seen and are referred to as virtual anodes and cathodes, virtual because they may occur at large distances from the stimulating electrode (Wikswa et al., 1994, 1995; Efimov et al., 2000a, b; Knisely et al., 1994).

The purpose of this paper is to describe the mechanism by which each of these sources of inhomogeneity lead to successful defibrillation and to describe the conditions for their success. We find that with large scale spatial inhomogeneity, defibrillation succeeds for functional reentry only if the virtual electrodes nearly cover the domain, and rarely succeeds in eliminating anatomical reentry. There are no similar restrictions for small scale inhomogeneities.

The outline of this paper is as follows. In the next section we describe a model that can be used to study the effect of externally applied stimuli. In the following section we describe the different dynamical responses that result from stimuli of different amplitude and phase, and different scale inhomogeneity on a one-dimensional ring. Then, we discuss the implications of these observations for higher dimensional tissue, concluding that small spatial scale inhomogeneity can account for much of the experimental data.

## 2. Modeling defibrillation

Models of cardiac activity typically combine two ingredients, a model of cellular behavior with a model of spatial coupling. For this paper we use simple models of cellular behavior, and couple them with the bidomain model for cardiac tissue. For an understanding of how externally applied currents affect cardiac tissue, we begin with the bidomain description of cardiac tissue (Henriquez, 1993; Keener and Sneyd, 1998; Neu and Krassowska, 1993). For the usual bidomain model, cardiac tissue is assumed to be a two-phase medium, with comingled intracellular and extracellular domains. At each point of the cardiac domain, denoted  $\Gamma$ , there are potentials  $\phi_e$  and  $\phi_i$ , the extracellular and intracellular potentials, respectively, and the transmembrane potential,  $v = \phi_i - \phi_e$ . These potentials drive currents,

$$i_e = -\sigma_e \nabla \phi_e, \quad i_i = -\sigma_i \nabla \phi_i, \quad (1)$$

and a transmembrane current across the cell membrane dividing the two (comingled) regions. The conductivities of the two media are represented by the conductivity

tensors,  $\sigma_i$  and  $\sigma_e$ . Kirchoff's laws imply that

$$\chi \left( C_m \frac{\partial v}{\partial t} + I_{ion} \right) = \nabla \cdot (\sigma_i \nabla \phi_i), \quad (2)$$

$$\nabla \cdot (\sigma_i \nabla \phi_i + \sigma_e \nabla \phi_e) = 0. \quad (3)$$

The first of these equations implies that current can leave the intracellular space only as a transmembrane current, and that the transmembrane current has two components, namely the capacitive current and the ionic current  $I_{ion}$ . The second equation states that the intracellular and extracellular currents can be redistributed but charge is conserved (there are no intracellular current sources). In Eq. (2),  $C_m$  is the membrane capacitance, and  $\chi$  is the ratio of cell surface to total volume.

The boundary conditions for the bidomain model are that current flows only across the boundary of the extracellular space, while there is no current across the boundary of the intracellular space,

$$\mathbf{n} \cdot (\sigma_e \nabla \phi_e) = I(t) f_I(x), \quad \mathbf{n} \cdot (\sigma_i \nabla \phi_i) = 0, \quad \text{on } \partial \Gamma, \quad (4)$$

where  $\mathbf{n}$  is the outward normal unit vector to the boundary  $\partial \Gamma$ . It is also required that the total injected current be zero,

$$\int_{\partial \Gamma} f_I(x) dx = 0. \quad (5)$$

It should be noted that while the bidomain model is adequate for the case of large-scale inhomogeneities, it is not appropriate for small-scale inhomogeneities, since the bidomain model is derived using a homogenization argument in which small scale spatial inhomogeneities are smoothed or averaged out. Thus, inclusion of small scale inhomogeneity requires a different model (Keener and Panfilov, 1996).

We can get some idea of the behavior of this model by examining the case of a one dimensional cable. In this case, the current conservation equation (3) can be integrated once to obtain (with  $|f_I| = 1$ )

$$\sigma_i \frac{\partial \phi_i}{\partial x} + \sigma_e \frac{\partial \phi_e}{\partial x} = I(t), \quad (6)$$

and we learn that

$$\sigma_i \frac{\partial \phi_i}{\partial x} = \frac{\sigma_i}{\sigma_i + \sigma_e} \left( \sigma_e \frac{\partial v}{\partial x} + I(t) \right), \quad (7)$$

so that Eq. (2) for the transmembrane potential becomes

$$\chi \left( C_m \frac{\partial v}{\partial t} + I_{ion} \right) = \frac{\partial}{\partial x} \left( \sigma \frac{\partial v}{\partial x} + \frac{\sigma_i}{\sigma_i + \sigma_e} I(t) \right), \quad (8)$$

where

$$\sigma = \frac{\sigma_i \sigma_e}{\sigma_i + \sigma_e}, \quad (9)$$

with boundary condition

$$\frac{\partial v}{\partial x} = -\frac{I}{\sigma_e} \quad \text{at } x = 0, L, \quad (10)$$

where  $L$  is the length of the cable.

It is easy to see that the response of the transmembrane potential to current stimuli is different for a homogeneous or inhomogeneous cable. For a homogeneous cable, the current source has influence only at the boundary, and the interior source term

$$\frac{\partial}{\partial x} \left( \frac{\sigma_i}{\sigma_i + \sigma_e} I(t) \right) \quad (11)$$

is identically zero. In a simple (unphysiological) situation in which the ionic current is linear  $I_{ion} = v/R_m$  and there is a steady applied current  $I$ , the steady solution is

$$v(x) = -\frac{I}{\sigma_e} \frac{\sinh((x - L/2)/\Lambda)}{\cosh(L/2\Lambda)}, \quad (12)$$

where  $\Lambda^2 = R_m \sigma / \chi$ . For a domain that is large compared to the space constant  $\Lambda$ , this solution exhibits exponential decay away from each boundary, and is essentially zero in the interior of the domain. This corresponds to the well known fact that the response to a stimulus is depolarization close to one boundary and hyperpolarization close to the opposite boundary, with little effect in the interior of the domain. Thus, without resistive inhomogeneity, there is effectively no transmembrane current generated by the stimulus in the interior of the tissue (several space constants from the stimulus source).

On the other hand, if  $\sigma_i/(\sigma_i + \sigma_e)$  is not constant (there is resistive inhomogeneity), the inhomogeneity provides additional sources and sinks of transmembrane current at points throughout the interior of the medium. It is the distribution of these sources and sinks that is responsible for defibrillation.

To get some insight into how this works in higher dimensions, we calculate that

$$\begin{aligned} & \nabla \cdot (\sigma_i \nabla \phi_i) - \nabla \cdot (\sigma_i (\sigma_i + \sigma_e)^{-1} \sigma_e \nabla v) \\ &= \nabla \cdot (\sigma_i \nabla v + \sigma_i \nabla \phi_e) - \nabla \cdot (\sigma_i (\sigma_i + \sigma_e)^{-1} \sigma_e \nabla v) \\ &= \nabla \cdot (\sigma_i (\sigma_i + \sigma_e)^{-1} \sigma_i \nabla v + \sigma_i \nabla \phi_e) \\ &= \nabla \cdot \sigma_i (\sigma_i + \sigma_e)^{-1} (\sigma_i \nabla v + (\sigma_i + \sigma_e) \nabla \phi_e) \\ &= \nabla \cdot (\sigma_i (\sigma_i + \sigma_e)^{-1} I_t), \end{aligned}$$

where  $I_t$  is the total current,  $I_t = \sigma_i \nabla v + (\sigma_i + \sigma_e) \nabla \phi_e$ . Eq. (2) becomes

$$\begin{aligned} \chi \left( C_m \frac{\partial v}{\partial t} + I_{ion} \right) &= \nabla \cdot (\sigma_i (\sigma_i + \sigma_e)^{-1} \sigma_e \nabla v) \\ &+ \nabla \cdot (\sigma_i (\sigma_i + \sigma_e)^{-1} I_t). \end{aligned} \quad (13)$$

From this we see that if  $\sigma_i (\sigma_i + \sigma_e)^{-1}$  is inhomogeneous in space, then when  $I_t$  is non-zero, there are sources and sinks of transmembrane current in the interior of the tissue. It should be noted that if  $\sigma_i (\sigma_i + \sigma_e)^{-1}$  is proportional to the identity matrix, there are no

virtual anodes and cathodes, and defibrillation is impossible. However, this occurs only if  $\sigma_i = c\sigma_e$  (equal anisotropy ratios), and it is well known that this does not hold for cardiac tissue.

There is no question that these source terms exist. In fact, virtual electrodes have been found to be induced by unequal anisotropy of intracellular and extracellular spaces (Sepulveda and Wikswo, 1987) myofiber curvature (Trayanova and Skouibine, 1998), fiber narrowing (Sobie et al., 1997), spatial inhomogeneity of intracellular volume fraction (Trayanova, 1999), discontinuity associated with gap junctions, and intercellular clefts (Fast et al., 1998). The primary issue of concern in this paper is how different spatial distributions of sources and sinks affects the outcome of a defibrillatory shock.

For much of the discussion that follows, to describe the ionic dynamics, we will use a two variable model of FitzHugh–Nagumo type (Keener and Sneyd, 1998). A useful (but not physiological) specific example of these takes the form

$$I_{ion} = -f(v) + \alpha w, \quad \frac{\partial w}{\partial t} = \varepsilon(v - \gamma w), \quad (14)$$

where  $f(v) = Av(1-v)(v-a)$ , with  $\alpha = a = 0.1$ ,  $\varepsilon = 0.01$ ,  $A = 1$ , and  $\gamma = \frac{1}{3}$ . In this model, the rest potential  $v = 0$  corresponds to the polarized membrane state. A positive stimulus, leading to an increase of  $v$  is a depolarizing stimulus, and a negative stimulus, leading to a decrease of  $v$  is a hyperpolarizing stimulus.

From time to time we will refer to more detailed ionic models such as the Beeler–Reuter model (Beeler and Reuter, 1977). However, since the phenomena we wish to explore are more readily described using two variable caricatures, much of our discussion will focus on FitzHugh–Nagumo kinetics.

Two variable models have the advantage that they can be viewed in the phase plane. For example, in Fig. 1 is shown a typical periodic traveling wave solution on a ring, moving from right to left. The upper panel shows the variables  $v$  and  $w$  plotted as functions of  $x$  for fixed  $t$ , while the lower panel shows the phase plane projection of these same trajectories. In the lower panel, the dashed curves are the  $v$  and  $w$  nullclines, found by setting  $I_{ion} = 0$  (the  $v$  nullcline, cubic shaped) and  $dw/dt = 0$  (the  $w$  nullcline, a monotone increasing function of  $v$ ).

In the phase plane, sharp transitions are seen as curve segments that connect the left and right branches of the cubic nullcline while keeping  $w$  relatively unchanged. These transitions are identified as fronts or backs if their wavespeed is positive (for fronts) or negative (for backs) (Keener and Sneyd, 1998). The wavespeed is defined as that number  $c$  for which there is a monotone increasing, heteroclinic trajectory of the equation

$$\frac{d^2v}{d\xi^2} - c \frac{dv}{d\xi} - I_{ion} = 0 \quad (15)$$

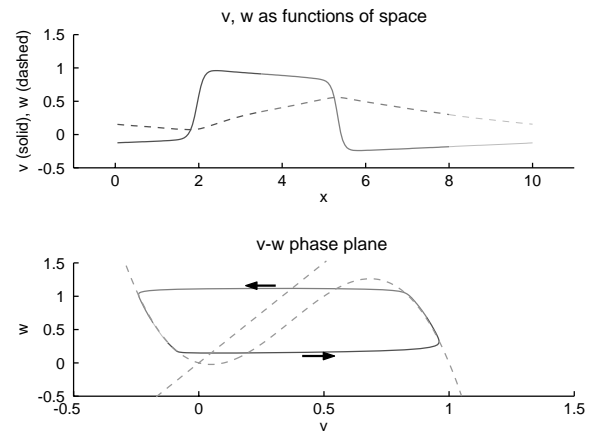


Fig. 1. Periodic traveling wave, traveling from right to left, with  $v$  shown as the solid curve, and  $w$  shown as dashed, in the upper panel. In the lower panel is the  $v$ – $w$  phase plane trajectory of this same solution, with the  $v$  and  $w$  nullclines shown as dashed curves. Arrows in the phase plane indicate the direction of increasing  $x$ .

connecting the smallest zero of  $I_{ion}$ , say  $v_-$ , with its largest zero  $v_+$ , with  $w$  fixed. It is easy to show that the sign of  $c$  is the opposite of the sign of

$$\int_{v_-}^{v_+} I_{ion} dv. \quad (16)$$

As a result, there is a value of  $w$ , called the zero speed level, at which  $c = 0$ . For the cubic FHN dynamics, the zero speed level is  $w_0 = \frac{2}{27}(\alpha^2 - \alpha + 1)^{3/2}$ . A transition with  $w$  below  $w_0$  has positive speed and is therefore a front, while a transition above this level has negative speed and is a back.

### 3. Defibrillation of a ring

Since fibrillation is a state in which there are one or many reentrant waves, the goal of defibrillation is to eliminate all of these reentrant waves, regardless of their structure or location, allowing the tissue to return to rest, awaiting the next normal action potential. From a dynamical systems point of view, the goal of a defibrillatory shock is to change the state of the system by moving it to the attracting basin of the rest state. We need, therefore, to understand something about basins of attraction for the variety of possible behaviors, and how to move the system from one basin to another.

Since most of the fundamental ideas can be understood for a one-dimensional ring, in this section we focus on this simplified geometry. For a one-dimensional ring, reentrant activity corresponds to a wave (or waves) rotating around the ring. For a ring, there is a limit on the number of stable attractors, and the number of such attractors is always odd. This includes waves with one or more action potentials moving in the

clockwise direction, an equal number moving in the counterclockwise, as well as the uniform rest state.

For two-variable models, these different states can be distinguished by a topological criterion, as follows. At any point in time, plot the solution as a curve in phase space, parameterized by space (as in Fig. 1). Because the spatial domain is a ring, hence periodic, the solution curve in phase space is always a closed curve.

The zero speed level, described above, can be used to define a winding number for trajectories. Consider a thin ellipse, with major axis along the zero speed level and centered at the middle zero of  $f(v)$  (see Fig. 2). For this discussion, the precise size of the ellipse is not significant, so long as typical periodic traveling waves surround it. Points that lie inside this ellipse correspond to the core of spirals in two spatial dimensions, and are regarded as “phaseless points” (Winfree, 1983, 1997). For curves that do not intersect this ellipse, the winding number is defined as the integer number of times the curve wraps around this ellipse, positive if moving in space from left to right gives clockwise rotation about the ellipse, and negative if counterclockwise. If the number of windings around the ellipse is zero, the winding number is zero. For example, in Fig. 1 where there is a single periodic wave moving from right to left, the rotation in the phase portrait is counterclockwise, hence the winding number is  $-1$ .

In the limit that  $\epsilon$  is small, the winding number is an invariant of the flow for dynamics (14) for all trajectories with an appropriate restriction on the size of  $|\partial w/\partial x|$ . A restriction on  $|\partial w/\partial x|$  is necessary in order to make sure that each wrap around the ellipse takes up a sufficient amount of space  $x$  (Cytrynbaum, 2001).

It would be nice if this winding number were an invariant of the flow for all dynamics of FHN type, but it is not. It is easy to find dynamics where a wave on a periodic ring (with winding number  $\pm 1$ ) is unstable but persists for quite some time before collapsing into the rest state. However, the collapse can be identified as a transition in which the winding number changes from  $\pm 1$  to zero.

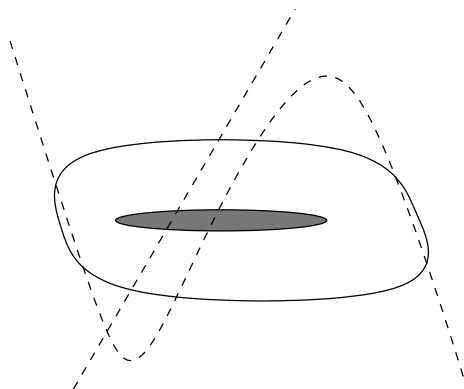


Fig. 2. A trajectory in the phase portrait with winding number  $\pm 1$ .

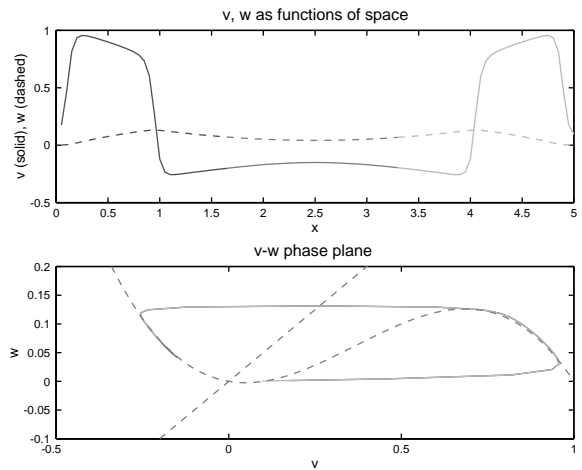


Fig. 3. The profile (two waves traveling outward from the center), having winding number 0, created following application of an S1 stimulus at the center of a ring.

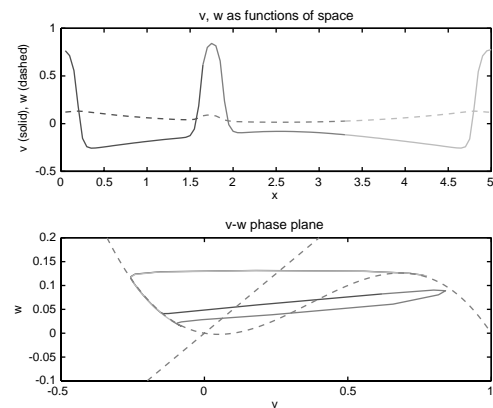


Fig. 4. The profile created after application of an S2 stimulus applied at a point to the left of center, having winding number 1. This profile evolves into a periodic traveling wave moving from left to right.

Even though the winding number is not an invariant, it gives a useful characterization of the behavior of waves because it gives criteria for transitions between different states. For example, it is easily understood that it is impossible to create a single rotating wave using an S1–S2 stimulus protocol with the two stimuli applied at the same point on the ring, since the winding number can never thereby (because of symmetry) be anything other than zero.

Similarly, to create a single traveling wave with an S1–S2 stimulus with the stimuli applied at different places requires the correct timing to turn a double cover of a single curve (the result of the S1) into a single loop with winding number  $\pm 1$ . In Figs. 3 and 4 are shown snapshots of this sequence of events. In Fig. 3 are shown two action potentials propagating outward that were initiated following a stimulus that was applied at the center of the spatial domain at time  $t = 0$ . The phase portrait for this trajectory is a double cover of a single curve. In Fig. 4 is shown the result at the end of the S2

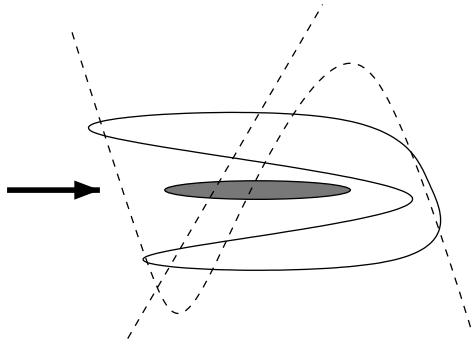


Fig. 5. A depolarizing stimulus applied in the partially refractory tail may produce winding number zero by creating a new wavefront and new waveback.

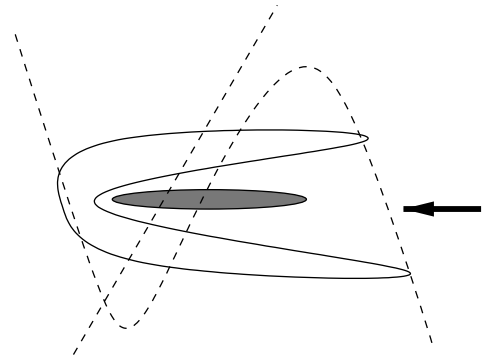


Fig. 6. A hyperpolarizing stimulus applied in the excited region may produce winding number zero by creating a new wavefront and new waveback.

stimulus that was applied slightly left of center. The phase portrait for this trajectory shows that what was before a double cover of a single curve has now been split into a loop with winding number 1. After the S2 stimulus is ended this profile quickly evolves into a periodic traveling wave moving from left to right.

(A similar winding number can be defined for general ionic models by plotting  $v(x, t)$  against  $v(x, t - \tau)$  for some fixed delay  $\tau$  along some closed curve in space.) We use the winding number to assess the long-term behavior of trajectories following a brief shock. That is, trajectories with winding numbers  $\pm 1$  converge to a traveling wave profile, while those with winding number zero converge to the uniform rest state.

Suppose the ring size is such that it allows only winding numbers of  $-1, 0$  and  $1$ . That is, the only stable attractors are the rest state and left or right moving periodic traveling waves. A time-dependent perturbation to one of these can have one of three outcomes. It could change the winding number or keep it the same. Specifically, if we start with a left moving travelling wave, a time dependent perturbation could change it so that it returns to rest, or that the wave reverses direction, or the wave could remain the same, with only a phase shift. These are the only possibilities.

It is clear how each of these transitions can be effected. To turn a winding number  $\pm 1$  trajectory into a winding number  $0$  trajectory it is sufficient to apply a depolarizing stimulus at a place where the dynamics are partially recovered in such a way that a portion of the phase plane curve is moved from left to right so that it no longer surrounds the defining ellipse (see Fig. 5). Similarly, it is sufficient to apply a hyperpolarizing stimulus at a place where the dynamics are excited in such a way that a portion of the phase plane curve is moved from right to left so that it also fails to surround the ellipse (see Fig. 6). Of course, if both of these events occur at the same time then the winding number changes sign, leading to a reversal of the direction of travel of the wave.

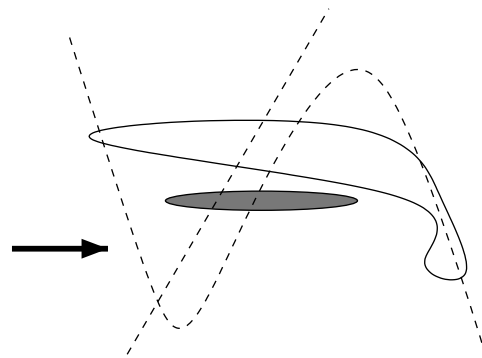


Fig. 7. A depolarizing stimulus applied in the recovered region may produce winding number zero by converting a wavefront into a waveback.

Each of these has three subcases which have slightly different physical descriptions. For example, with a depolarizing stimulus applied to the partially refractory tail, the stimulus might create a new front and a new back (Fig. 5), it might convert a front into a back (Fig. 7), or it might convert a back into a front (Fig. 8).

To demonstrate that each of these transitions is possible, and to explore the effect of inhomogeneities of resistance of different spatial scale, we simulated Eq. (8) on a one-dimensional periodic ring, using FitzHugh–Nagumo kinetics (14). The conductivities were taken to be

$$\sigma_i = 0.01 \left( 1 + \frac{1}{2} \sin \frac{2\pi(\Omega x + \phi)}{L} \right), \quad \sigma_e = 0.01. \quad (17)$$

The length of the ring was  $L = 10$ , which was large enough to support a single stable periodic traveling wave. The spatial grid size was  $\Delta x = 0.05$ , and the time step was  $\Delta t = 3$ , using the Crank–Nicholson method. Initially, at time  $t = 0$ , there was a periodic traveling wave (winding number  $-1$ ), as shown in Fig. 1. At some time during the evolution of this wave, a brief stimulus was applied, simulated by setting  $I(t)$  to be some nonzero constant for the short duration of 10 time steps (30 time units). (This is short compared to the time

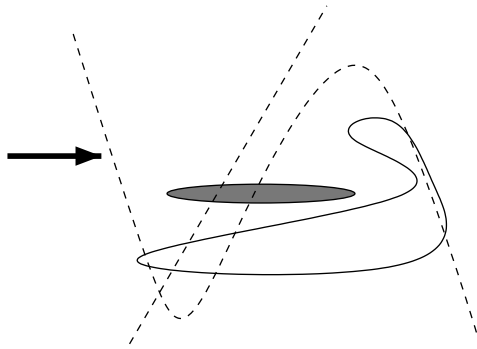


Fig. 8. A depolarizing stimulus applied in the refractory region may produce winding number zero by converting a waveback into a wavefront.

constant of the variable  $w$ ,  $\tau_w = (\gamma\epsilon)^{-1} = 300$  time units.) Notice that the effect of  $I(t)$  at various points on the ring is determined by the inhomogeneity of  $\sigma_i$  and  $\sigma_e$  in the source term (11).

By large scale we mean an inhomogeneity whose characteristic length scale is larger than the size of the critical domain, which for this model is about 1 space unit. For this simulation, we chose  $\Omega = 1$  to represent a large scale inhomogeneity. With large spatial scale inhomogeneity, the results are essentially as described above, with the important restriction that depolarizing current and hyperpolarizing current must always be in balance.

Beginning in Fig. 9 is shown the sequence of events leading to successful defibrillation. At time  $t = 15$ , the effects of the stimulus are depolarizing ahead of the wavefront, and hyperpolarizing at the waveback.

The trajectory is on its way to becoming like Fig. 7 with winding number zero. After the stimulus is removed, when the wave is reestablished, there are two wavebacks moving toward each other, which eventually collide causing the wave to collapse, leaving the medium at rest. This is shown in Fig. 10 for time  $t = 105$ .

This stimulus was successful because the depolarizing stimulus was of sufficient amplitude and was properly located so that the wavefront was converted to a waveback. The opposite also occurs (but is not shown), namely a properly located hyperpolarizing stimulus of sufficient amplitude can convert a waveback into a wavefront. Thus, successful defibrillation with  $\Omega = 1$  occurs if the depolarizing stimulus is properly timed and of sufficient amplitude or if the hyperpolarizing stimulus is properly timed and of sufficient amplitude.

If neither of these occur, the traveling wave is reestablished with a phase shift. However, if both of these occur, the wave is not only reestablished, but its direction of propagation is reversed, traveling in the opposite direction from the original.

That this last possibility occurs for FHN dynamics is demonstrated by the following figures. In Fig. 11 is shown the solution during the stimulus (at  $t = 6$ ) that

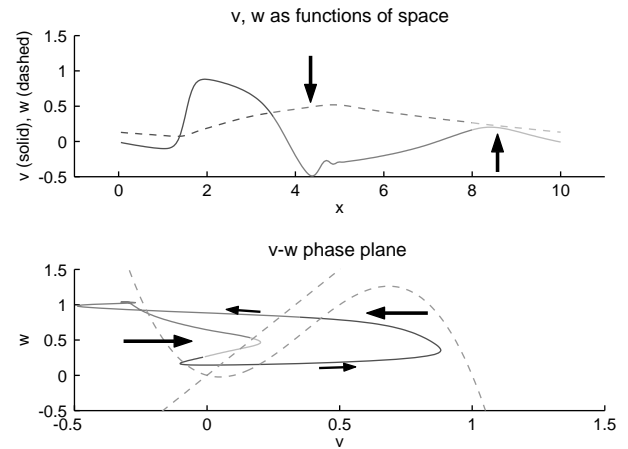


Fig. 9. Solution (with  $\Omega = 1$ ) midway through the stimulus of amplitude  $I = 0.17$  at  $t = 15$ . Large arrows indicate the location of maximal depolarizing and hyperpolarizing input.

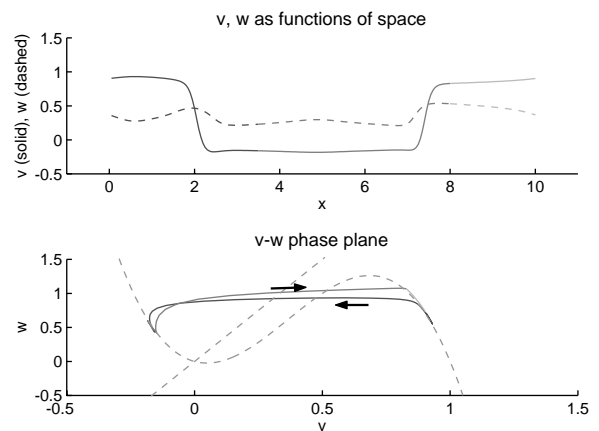


Fig. 10. Solution (with  $\Omega = 1$ ) at  $t = 105$  following a superthreshold stimulus.

will eventually rotate in the opposite direction. Notice that the depolarization is ahead of the wavefront and the hyperpolarization is slightly ahead of the waveback. It is apparent from the phase portrait that this stimulus has the potential of changing the sign of the winding number (from  $-1$  to  $+1$ ), and indeed this is the case.

After the stimulus is removed, the wave that is reestablished, shown in Fig. 12 is rotating in the opposite direction.

These simulations demonstrate that with a low spatial frequency inhomogeneity, defibrillation in one spatial dimension is not a true threshold phenomenon, because it depends crucially on the timing of the stimulus. Increasing the amplitude of a poorly timed stimulus cannot increase the likelihood of defibrillation. Only if it is properly timed is the amplitude of the stimulus significant.

This phase–amplitude dependence for large scale inhomogeneities is depicted in Fig. 13. Shown here are

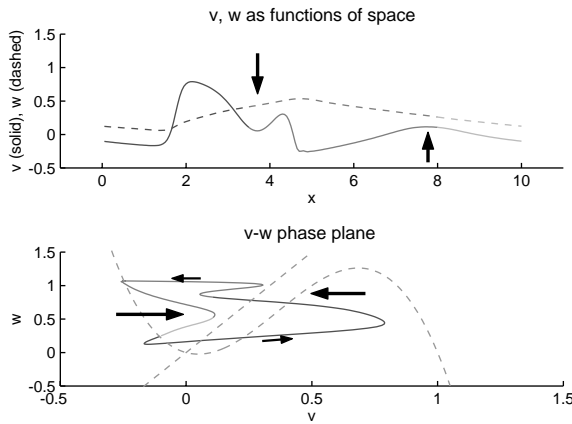


Fig. 11. Solution (with  $\Omega = 1$ ) during a stimulus of amplitude  $I = 0.17$  (at  $t = 18$ ) that eventually causes the wave to rotate in the opposite direction. Large arrows indicate the location of largest depolarizing and hyperpolarizing inputs.

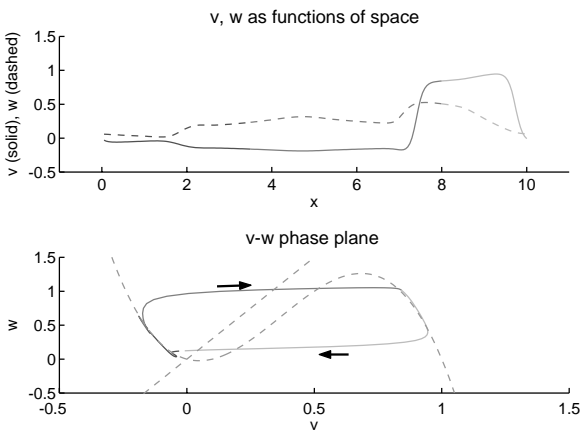


Fig. 12. Solution (with  $\Omega = 1$ ) at  $t = 285$  after the stimulus is removed that is rotating in the opposite direction (from left to right).

the three regions in which there is defibrillation success, phase resetting and propagation reversal, in the case that  $\Omega = 1$ . The region with defibrillation success has two components, one in which fronts are converted to backs via hyperpolarization and one in which backs are converted to fronts via depolarization. The region with defibrillation success is quite small, with the probability of success less than 20%.

Similar results to these were found using the Beeler–Reuter ionic model. With this full ionic model it is possible to eliminate a rotating wave by application of a depolarizing stimulus behind the tail of the action potential or by applying a hyperpolarizing stimulus near the front of the action potential. The effect of the depolarizing stimulus is to effectively turn a back into a front, as depicted by Fig. 8, even though these are not two-variable dynamics to which this concept of winding number can be applied. The effect of the hyperpolarizing stimulus is to turn a front into a back. In our simulations, however, the hyperpolarizing current required to reverse a front was significantly larger than the

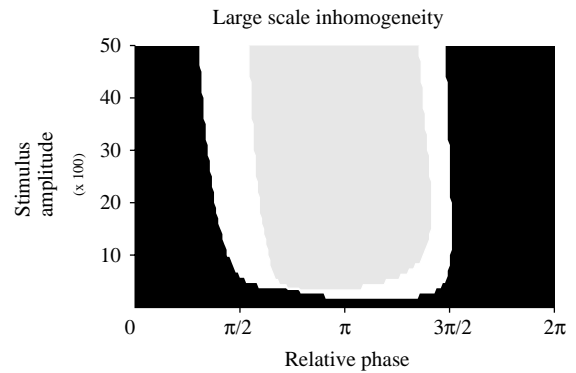


Fig. 13. Amplitude–phase diagram in which three possible outcomes occur, for  $\Omega = 1$ . Phase resetting occurs in the solid black region; direction reversal occurs in the gray region, and defibrillation occurs in the white region.

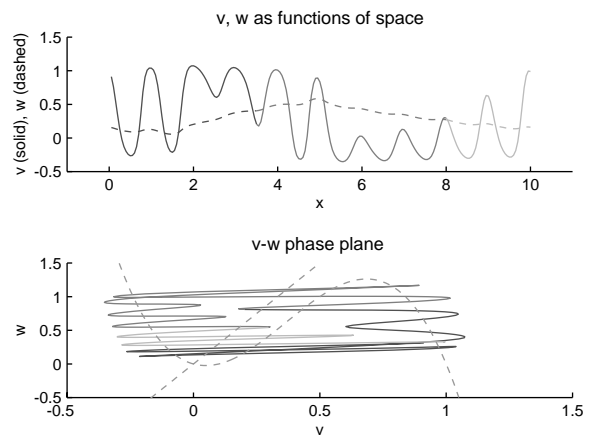


Fig. 14. Solution (with  $\Omega = 10$ , a small scale inhomogeneity) midway through the stimulus with amplitude  $I = 0.0031$ , at  $t = 15$ .

depolarizing current required to activate a back. This is readily explained by the fact that to hyperpolarize a cell during its upstroke requires sufficient hyperpolarizing current to overwhelm the inward sodium current. This current requirement is far larger than the amount of depolarizing current required to excite a partially recovered cell.

With small scale spatial inhomogeneities the results are substantially different. With small scale inhomogeneity, the locations of hyperpolarization and depolarization are closely spaced, and it is not as easy to predict the result of the stimulus, as it is with large scale inhomogeneities.

To illustrate the differences, numerical simulations were performed, this time with  $\Omega = 10$ , so that the period of the inhomogeneity was  $\frac{1}{10}$ th that of the ring.

In Fig. 14, the solution is shown midway through the stimulus. The effects of the inhomogeneity are clearly seen as sites of depolarization and hyperpolarization, however, from this figure it is not possible to predict the outcome of the stimulus. One noticeable feature is that



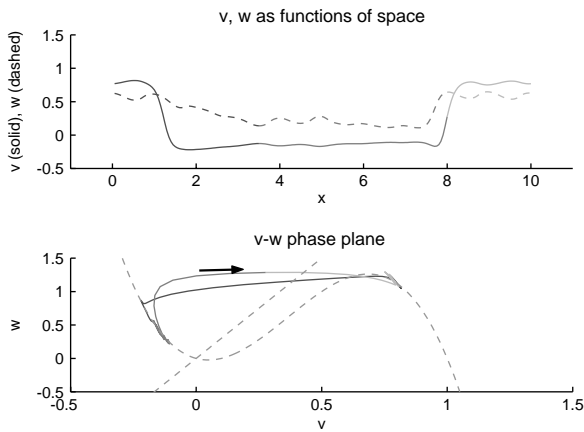


Fig. 15. Solution (with  $\Omega = 10$ ) at  $t = 165$  after the suprathreshold stimulus ( $I = 0.0031$ ) was removed and slightly before the wave collapses and the medium returns to rest.

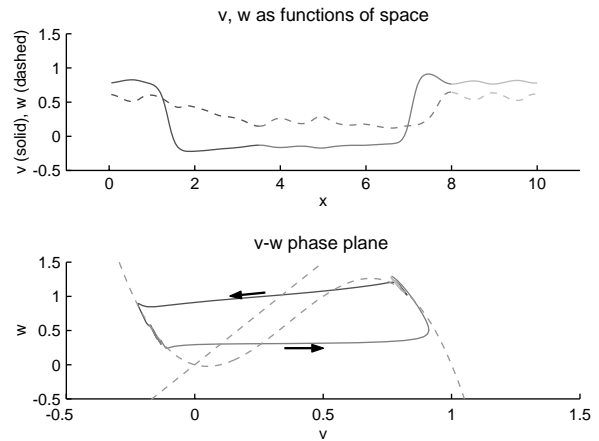


Fig. 16. Solution (with  $\Omega = 10$ ) at  $t = 165$  after the subthreshold stimulus ( $I = 0.0030$ ) was removed and the rotating wave has been reestablished.

even though the amplitude of the depolarizing and hyperpolarizing currents is everywhere the same, the effect of these on the potential depends strongly upon where in the phase of the action potential the stimulus is applied.

After the stimulus is removed, the wavelike behavior is reestablished, although it is some time before the final outcome is evident. In Fig. 15, the solution is shown some time after the stimulus has terminated (at  $t = 165$ ). By this time it is evident from the phase portrait that the winding number is zero, and the wave consists of two wavebacks which are traveling in opposite directions and, because the ring is periodic, will shortly coalesce and collapse, allowing the medium to return to rest. In other words, the shock was successful at terminating the rotating wave.

The most significant observation is that this result is practically independent of timing.

If the stimulus amplitude is reduced slightly ( $I = 0.0030$ ) the outcome is dramatically different. In this case, the early evolution is indistinguishable from that shown in Fig. 14, however, the later evolution diverges from the previous case. In Fig. 16 is shown the wave at  $t = 165$ . Here it is evident from the phase portrait that there is a wave front and a wave back with winding number  $-1$ , so that a rotating wave is reestablished. The phase of the wave has merely been reset.

From numerical simulations we conclude that while the mechanism of defibrillation is the same, namely the stimulus had the effect of turning a winding number  $\pm 1$  wave into a winding number zero wave, there is a significant difference in that the result is independent of timing, and some of the conversions seen with large scale inhomogeneity are not possible with small scale inhomogeneity.

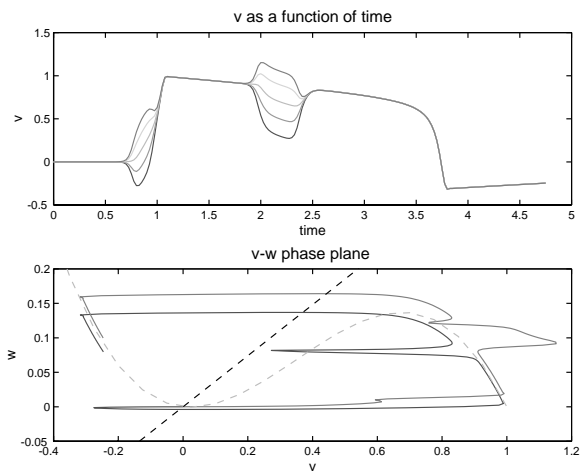


Fig. 17. Upper panel: Plots of the potential as a function of time for five different locations along a single cell during an S1–S2 stimulus protocol applied to an FHN cell. Lower panel: Phase portrait plots of  $v$  and  $w$  at the opposite ends of a cell during an S1–S2 stimulus protocol.

The natural question is to ask why these are different.

A clue to understanding this difference is provided by recent experimental studies of single cells stimulated by an applied electric field (Sharma and Tung, 2002). In this study the transmembrane potential was measured (using voltage sensitive fluorescent dyes) at five sites along the length of a single cell before, during, and after the application of two successive stimuli applied to the extracellular bath. The first stimulus was adequate to excite the resting cell, and the second stimulus was applied 20 ms after the first, during the excited phase of the action potential.

The results of these experiments are simulated in Fig. 17. The upper panel in this figure is remarkably similar to those in Sharma and Tung (2002) even though this figure was produced from a numerical simulation with the FHN model. For this simulation, a current was applied to the extracellular medium, and the length of

the cell was taken to be  $\frac{1}{6}$ th the space constant of the intact cellular medium. We used the FHN dynamics (14) with parameter values  $\varepsilon = 0.1$ ,  $a = 0.05$ ,  $\gamma = 2.7$ , and  $A = 50$ , and numerical integration used the Crank–Nicholson method. The S1 stimulus was 0.3 time units duration and the S2 had 0.5 time units duration, and were both of amplitude  $I = 4$ . The upper panel in Fig. 17 shows five traces of voltage potential from five evenly spaced locations along the cell (at  $x = 0, 0.25, 0.5, 0.75, 1.0$ ). When there is an applied current, one end of the cell is depolarized relative to the opposite end which is hyperpolarized, but when the current is removed, the membrane potential quickly becomes homogeneous throughout the cell. The lower panel shows the phase plane trajectory for the two ends of the cell.

The noticeable feature of this figure is that the effect of the stimulus is different depending on when during the action potential it is applied. If applied when the cell is recovering or at rest, the net effect is strong depolarization, while if the cell is excited, the net effect is slight hyperpolarization. For these dynamics, the net depolarization increased as the cell is more recovered, and the net hyperpolarization increased as the stimulus was applied later during the excited phase.

The analysis to understand what determines the net effect has been given previously in several places, for example, (Krassowska and Neu, 1994; Pumir and Krinsky, 1997). The conclusion of that analysis is that if the applied stimulus is short compared to the time constants of all slow gating variables, then the potential is well approximated by

$$v(y, t) = \frac{I(t)}{\sigma_e} \left( y - \frac{L}{2} \right) + \phi(t) + O(\varepsilon), \quad (18)$$

on the interval  $0 < y < L$ , where  $\phi(t)$  is the cell-averaged membrane potential, whose dynamics satisfy

$$C_m \frac{d\phi}{dt} = - \int_0^L I_{ion} \left( \frac{LI(t)}{\sigma_e} \left( x - \frac{1}{2} \right) + \phi \right) dx, \quad (19)$$

holding all slow gating variables fixed. Here,  $\varepsilon = (L/A)^2$ , where  $L$  is the length of the cell, and  $A$  is the length constant.

It follows from Eq. (19) that the effect of the stimulus is strongest for those potentials where  $I_{ion}$  is most strongly nonlinear, and if  $I_{ion}$  is linear in  $v$ , then the stimulus has no effect. This explains why the effect of the stimulus is so strong and depolarizing on a resting cell, but is much smaller on a cell that is excited. For a resting cell, the sodium current is available and the result of depolarizing one end of the cell far outweighs the effect of hyperpolarizing the opposite end, so that the net effect is overall depolarization. In contrast, during the excited phase the primary current is an outward potassium current, which on this time scale is linear in  $v$ ,

so that depolarization at one end of the cell and hyperpolarization at the opposite have nearly cancelling effects. Observations of this nature form the basis of the paper Pumir et al. (1998).

Of course, Eq. (19) is only a first-order approximation, ignoring slower time scale events. These effects can be seen in the lower panel of Fig. 17, where the recovery variable changes more rapidly at the depolarized end of the cell than at the hyperpolarized end of the cell, so that when the stimulus is removed, the recovery variable is not uniform throughout the cell. The primary observation, however, is that the response to small spatial scale inhomogeneity is a nonlinear average of depolarizing and hyperpolarizing currents with the consequence that for cardiac tissue the response is strongly depolarizing if the tissue is partially or fully recovered, and the response is weak otherwise. This difference in response is because the inward sodium current is strongly nonlinear in  $v$  while other currents are far less so. This argues that the most significant current responsible for defibrillation is the sodium current.

This analysis applies regardless of the source of resistive inhomogeneity, as long as the spatial scale of that inhomogeneity is small compared to the length constant of the medium. Thus, even though this analysis was initially done with gap junctions and sawtooth potentials in mind, it applies equally well if the primary resistive inhomogeneity is in the extracellular space or results from other small-scale structures.

#### 4. Defibrillation in higher dimensions

The mechanism for defibrillation that we proposed in the previous section is that a stimulus, however it is provided, must result in a winding number zero trajectory after it has ended. In two or three spatial dimensions, an analogous winding number cannot be defined, so we retain the definition of winding number used on a one-dimensional ring. That is, for any specific non-intersecting closed curve in physical space whose image in the phase plane does not intersect the prescribed ellipse, the winding number is non-zero if the total number of wraps around the ellipse is non-zero as one traverses the simple closed curve, and zero otherwise. (There is additional subtlety as to how to correctly orient the curve in three dimensions, so for the sake of simplicity, for the remainder of this discussion we consider only two-dimensional regions of space.) In two- or three-dimensional space, this winding number may change as the defining curve is moved around. However, we can determine when there is a reentrant pattern as follows. Let  $W_c$  be the winding number associated with a particular closed curve  $c$ , and let

$$W = \max_c |W_c| \quad (20)$$

over all closed curves for which  $W_c$  is defined. If  $W \neq 0$  there is a reentrant pattern of activity. Thus, defibrillation is successful if a short time after the stimulus is terminated,  $W = 0$ , while defibrillation is unsuccessful if not.

Notice that this definition of  $W$  is non-local. If the spatial domain contains no holes, we could use a local property determined by the number and nature of phaseless regions. However, defibrillation is *not* equivalent to the elimination of all phaseless regions if the domain contains one or more holes (e.g., a ring or annulus).

The main lesson we learned in the previous section is that to use depolarization to obtain winding number zero, a depolarizing current of sufficient amplitude must be applied to a particular region on the trajectory in the recovery phase, illustrated by Fig. 5. Depolarization of other parts of the action potential are of no consequence to changing the winding number. The domain on the phase portrait that requires depolarization to produce winding number zero we identify as the critical domain. (The critical domain is closely related to, but not exactly the same as, the excitable gap.) There are analogous hyperpolarization critical domains whose significance will be described below.

If  $W$  is not zero, there are a number of regions in space which map to the ellipse in phase space. These are analogous to Winfree's phaseless points (Winfree, 1983, 1997), and are identified as the core region of reentrant activity. Any simple curve surrounding one of these regions must necessarily intersect the critical domain. In fact, critical domains are continuous in space and can begin or end only at the preimage of phaseless ellipses or at the domain boundary. If there is but a single spiral, then there would be a single phaseless region with a spiral shaped critical domain emanating outward from it. A typical arrangement of critical domains is shown in Fig. 18. In this figure the contours are those of the potential  $v$  and the darkened regions are the critical domains, numerically computed from a simulation of the FHN model in a fibrillatory state. (The axes are in non-dimensional space units and the length constant is one space unit.)

A criterion for defibrillation is that all of the critical domains be depolarized with sufficient amplitude, and if defibrillation fails, it is because some portion of a critical domain was not adequately depolarized. If a new reentrant pattern is created, it is because some preexisting critical domain was differentially depolarized. In fact, if some portion of a critical regions fails to be depolarized it is the ends of that region that become the new centers of reentrant activity after the stimulus has terminated (Efimov et al., 2000a, b; Winfree, 1983).

The next issue to face is how the stimulus is generated. If the stimulus is generated by a large scale spatial

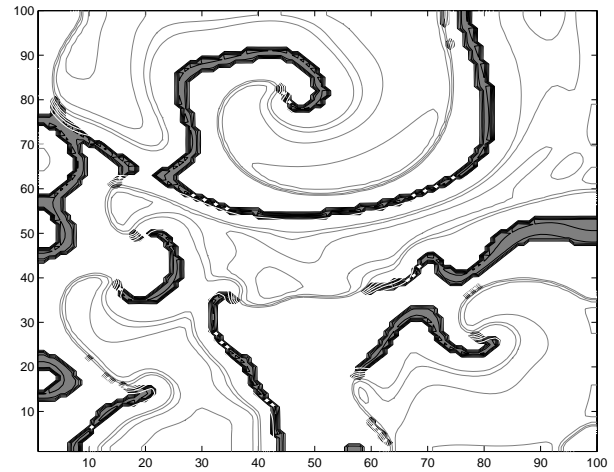


Fig. 18. An example of the arrangement of critical domains for FHN dynamics in a fibrillatory state. Critical domains end at phaseless regions or at a boundary.

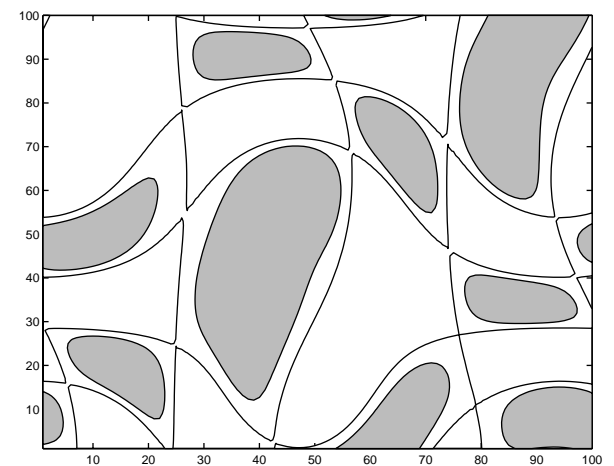


Fig. 19. An example of the depolarization tiling due to virtual electrodes. The shaded regions denote superthreshold depolarization.

inhomogeneity, then because charge is conserved, the net transmembrane current from the stimulus is zero; total depolarizing and hyperpolarizing currents must be in exact balance. This means that the size of the region that receives depolarizing current is roughly the same as the size of the region that receives hyperpolarizing current. Of course, the regions with sufficiently large depolarizing current are smaller. Thus, the distribution of virtual anodes and cathodes creates a tiling pattern of depolarization and hyperpolarization. An example of such a tiling is shown in Fig. 19.

Now, according to the criterion developed above, defibrillation will be successful if the superthreshold depolarization tiling completely covers the critical regions. An example of this overlay is shown in Fig. 20.

For Fig. 20 defibrillation would not be successful. In fact, the probability that the distribution of critical regions fits entirely within the superthreshold depolarization tiling is zero. Thus, with a typical depolarization

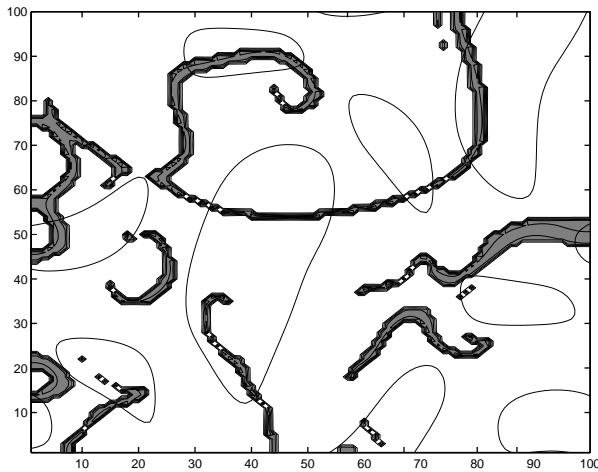


Fig. 20. An example of the overlay of the superthreshold depolarization tiling with the critical domains. Defibrillation is successful if the critical domains lie within the superthreshold depolarization tiling.

tiling and a typical distribution of critical domains (such as shown in Fig. 18), the probability of successful defibrillation is zero, and it is not improved with an increase in stimulus amplitude.

However, the complete picture requires consideration of the complementary hyperpolarizing critical regions. The hyperpolarization critical regions are connected to depolarization critical at the core of a spiral and their partial hyperpolarization leads to the creation of new spiral cores. If defibrillation required the immediate elimination of all spirals, defibrillation by this mechanism would be impossible. However, spirals formed by depolarization are typically paired with spirals formed by hyperpolarization. If the total parity of the spirals is zero, and if all the spiral pairs are sufficiently close together, they will subsequently collapse (since spiral pairs require a sufficient domain size to maintain themselves). If all such spiral pairs collapse, defibrillation is achieved. We believe this is the mechanism by which defibrillation is achieved by Eason and Trayanova (2002).

Since all the new spirals are formed on the boundaries of the superthreshold virtual electrodes, this mechanism requires that all of the space be closely packed with superthreshold virtual electrodes. In this way, newly formed spiral pairs are not too far apart and will collapse, and all preexisting cores receive superthreshold stimulus and are eliminated. It is also for this reason that reentry with a “virtual core” (rotation around an inexcitable hole) cannot be readily eliminated, unless a superthreshold virtual electrode completely surrounds the hole. If the virtual core is not surrounded by a virtual electrode, then as with a one dimensional ring, elimination of the reentrant wave is dependent on timing.

The story is much different with small scale inhomogeneity. The physics of anodes and cathodes is exactly the same, namely there is no net transmembrane current. Thus, the area where there is depolarization is roughly the same as the area where there is hyperpolarization. The difference, however, is that because the spatial scale of this pattern is small compared to the length scale of the tissue, the effect of the stimulus is quickly smeared out, or averaged, spatially. If there is no amplification of these currents, (if the tissue is effectively linear i.e., passive) then this average effect is zero; the depolarizing and hyperpolarizing currents cancel each other out for no net effect. On the other hand, if the tissue is locally active and excitable, so that it amplifies depolarizing currents, and deamplifies hyperpolarizing currents, then the net effect is always depolarization. Thus, with small scale inhomogeneity the critical domain, because it is partially recovered, always receives depolarizing stimulus; there is no need to align the depolarizing virtual electrode with the critical domain. The only concern is whether the amplitude of the stimulus is sufficient to excite the critical domain. Thus, spatial averaging has the effect of eliminating the need for critical domains to be precisely aligned with the regions of depolarization. All critical domains are depolarized. The only variable is the amplitude of the effective depolarization.

When a stimulus is applied to two- or three-dimensional tissue, the distribution of total current is not uniform, as it is in one dimension. The consequence of this is that with a fixed stimulus amplitude, in some regions of space the critical domain may receive adequate stimulus, while in other regions of space the critical domain may not be sufficiently excited. Thus, for a fixed stimulus protocol and electrode placement, there are regions of space that are superthreshold and regions that are subthreshold. The probability of defibrillation is the probability that no part of the critical domain lies in the subthreshold region at the time the stimulus is applied.

Notice one very important difference between small-scale and large-scale inhomogeneity. With large-scale inhomogeneity, the size of the superthreshold old depolarizing region is highly unlikely to exceed 50% and will generally be much smaller than that, whereas, with small-scale inhomogeneity, the size of the superthreshold region increases with the stimulus amplitude and could cover the entire tissue region. In Keener and Panfilov (1996), homogenization theory was used to derive an estimate for the effective amplitude of the stimulus from a more detailed tissue model. From their analysis, it was shown that the amplitude of the effective stimulus is related to several factors, written as

$$A = I(t) |K_i(x) \cdot \psi_i(x) + K_e(x) \cdot \psi_e(x)|. \quad (21)$$

Here  $I(t)$  is the stimulus amplitude,  $K_i(x)$  and  $K_e(x)$  describe the spatial distribution of intracellular and

extracellular current as determined by the large scale tissue structure (the mean field) and electrode placement, and  $\psi_i(x)$  and  $\psi_e(x)$  are vectors determined from local features (the small scale structure) of the cellular medium. We define  $\theta(x)$  to be a local tissue property such that if  $\int A dt > \theta$ , then the refractory tissue within a critical domain at position  $x$  will receive adequate stimulus to effect defibrillation locally. Thus, a stimulus is globally superthreshold if

$$\int A(t, x) dt > \chi(x)\theta, \tag{22}$$

for all  $x$ , where  $\chi(x)$  is the characteristic function of the critical domain, zero outside the critical domain and one inside. Since  $\chi(x)$  is continually changing during fibrillation, the probability of defibrillation is the probability that Eq. (22) holds on the entire tissue domain at the time the stimulus is applied (i.e., for a randomly chosen spatial distribution of critical domains).

This criterion is actually too demanding for two reasons. First, it is not necessary that all critical domains be excited, but only that the size of any critical domains that receive insufficient stimulus be small enough so that new reentrant patterns cannot emerge from them. Second, critical domains in the vicinity of tissue boundaries need not be completely stimulated for the same reason, since if they are close to a boundary they are not capable of sustaining a reentrant pattern but will spontaneously disappear.

With these caveats, we see that there are five features that determine whether or not a stimulus will successfully defibrillate a piece of tissue. These are:

- the amplitude (and/or duration) of the stimulus  $I(t)$ ;
- the distribution of current within the tissue specified by  $K_i(x)$  and  $K_e(x)$ ;
- the microstructure of the tissue specified by  $\psi_i(x)$  and  $\psi_e(x)$ ;
- the spatial distribution of the critical domain  $\chi(x)$ , and
- the threshold of the critical domain  $\theta(x)$ .

It is presumably easiest to change the size of the subthreshold region. In fact, increasing the stimulus amplitude  $I$  has exactly this effect. Since the function  $|K_i(x) \cdot \psi_i(x) + K_e(x) \cdot \psi_e(x)|$  is positive, increasing  $I$  has the effect of decreasing the size of the region that is subthreshold, thereby increasing the probability of defibrillation success.

The functions  $K_e(x)$  and  $K_i(x)$  are modified by electrode design and electrode placement. Clearly, satisfying Eq. (22) is more difficult if these have large spatial variations. Thus, one criterion for the design and placement of electrodes is that the field produced be as uniform as possible.

The functions  $\psi_i(x)$  and  $\psi_e(x)$  depend on the amplitude of local resistive inhomogeneity. Indeed, this theory predicts that if the amplitude of local resistive inhomogeneity is increased, then the DFT will be decreased. One way to increase local intracellular resistive inhomogeneity is with the application of heptanol, a gap junction decoupler, and this theory therefore predicts that addition of heptanol should decrease the DFT. Indeed, this is known to be true experimentally (Qi et al., 2001). It is also predicted by this theory that increased extracellular resistive inhomogeneity, such as might be caused by cell swelling, should decrease the defibrillation threshold. According to this theory, the DFT is also affected by the orientation of the mean field (a vector) relative to the resistive inhomogeneity (also a vector). Since the vector dot product  $K \cdot \psi$  is maximized when the vectors are parallel, this theory predicts that the DFT is lowest if the mean field is oriented to be parallel to the direction of greatest resistive inhomogeneity.

The most significant determinant of defibrillation success is the distribution of the critical domain  $\chi(x)$  compared to the location of the subthreshold regions. For example, in Fig. 20, the critical domain consists of 8 narrow strips of varying lengths. To get some understanding of the importance of this distribution, we suppose that the critical domain consists of several disjoint striplike components. We can model this in a simple way with a one dimensional example. Suppose the unit interval contains several ( $k$ ) subintervals of total length  $B$  that receive subthreshold stimulus. Suppose further that there are  $n$  critical regions with total length  $A$ . For simplicity, we suppose that the  $n$  critical subregions are all intervals of identical length  $A/n$ . The probability of defibrillation success is the same as the probability that when the critical intervals are randomly distributed on the unit interval, there is no overlap with the subthreshold regions of total length  $B$ .

We can calculate this probability for large enough  $n$ . Since there are  $k$  distinct subthreshold regions, the area of the region in which the midpoint of a critical interval may fall without jeopardizing successful defibrillation is  $1 - B - kA/n$ , at least if  $n$  is large enough. This means that the probability of defibrillation success is

$$P(\text{success}) = \left(1 - B - \frac{kA}{n}\right)^n, \tag{23}$$

a decreasing function of  $n$ . In other words, as the critical domains become smaller and less organized but more numerous, the probability of defibrillation success decreases.

This argument is readily generalized to two dimensional regions. Suppose that the subthreshold region has total area fraction  $B$ , and suppose that there are  $n$  critical domains that require superthreshold stimulus. We suppose that the critical domains are chosen

randomly and are characterized by a single size parameter, say length  $l$ . The distribution of sizes is given by some probability distribution function, say  $p_n(l)$ , with mean value  $A/n$ ,

$$\int_0^{\infty} l p_n(l) dl = \frac{A}{n} \quad (24)$$

so that the total length of critical domains is  $A$ , on average. The probability that a critical domain of size  $l$  can be placed entirely within the superthreshold region is some function,

$$P(\text{success}|l) = F(B, l) \quad (25)$$

a monotone decreasing function of  $l$  with  $F(B, 0) = 1 - B < 1$ . The probability that defibrillation of all  $n$  excitable components will be successful is

$$P(\text{success}) = \left( \int_0^{\infty} F(B, l) p_n(l) dl \right)^n. \quad (26)$$

Clearly,

$$\begin{aligned} \int_0^{\infty} F(B, l) p_n(l) dl &\leq \int_0^{\infty} F(B, 0) p_n(l) dl \\ &= 1 - B < 1, \end{aligned} \quad (27)$$

so that

$$\lim_{n \rightarrow \infty} P(\text{success}) = 0. \quad (28)$$

In other words, as the critical regions become smaller and more disorganized, the probability of defibrillation success decreases.

It is known that the defibrillation threshold is lower within the first few cycles of ventricular fibrillation than after 10 seconds of fibrillation (Gradaus et al., 2002), and that the defibrillation threshold for monomorphic tachycardia is lower than for fully developed fibrillation (AMA Standards, 1986).

If fibrillation corresponds to multiple unstable rotating wavelets, and if fibrillation is initiated by the degradation of one or two larger spirals, then this experimental observation is consistent with the present theory. This is because for spirals, the critical regions are highly organized, however, as breakup occurs and the spirals degrade into fully developed fibrillation, so also the organization of the critical regions degrades, although the total length of the critical domain remains more or less the same. According to this theory, the defibrillation threshold should be lower for monomorphic tachycardia or early fibrillation.

The final way that the success of defibrillation is determined is by the threshold  $\theta$ . Recall that  $\theta$  is a measure of how large a stimulus is required to depolarize marginally recovered tissue. One would expect that  $\theta$  could be modified by the presence of drugs.

Sotolol is a potassium channel blocker, specifically blocking the delayed rectifier potassium channel, and as a consequence is known to lengthen action potential

duration. According to this theory, a potassium channel blocker should have little effect on the DFT since potassium currents are effectively linear and not responsive to small spatial scale hyperpolarization/depolarization pairs. Indeed, it is known that sotolol has no significant effect on the DFT (Ujhelyi et al., 1999).

The story for sodium channel blockers is more complicated. A sodium channel blocker has two effects on the dynamics that are related to this theory. First, with fewer sodium channels available, the action potential upstroke and the speed of propagation are slowed. This means that the critical domain, defined by the zero wave speed, is shifted. In addition, with fewer available sodium channels it would seem likely that the threshold for depolarization of the critical region is increased. This loose argument suggests that sodium channel blockers should increase the DFT. Unfortunately, we do not yet have solid evidence from a detailed ionic model that this line of reasoning is correct. Nonetheless, the experimental evidence is that sodium channel blockers, such as lidocaine and mexiletine, increase the DFT (Crystal et al., 2002; Ujhelyi et al., 1999).

## 5. Discussion

These results can be described in terms of dynamical systems theory. The system we examined has several stable attractors, including the rest state and the fibrillatory state in which there are several or many reentrant patterns. Defibrillation can be viewed as the attempt to move the system from one attractor, a fibrillatory state, to another, the uniform rest state, by application of a time dependent perturbation. However, the form of the perturbation is not arbitrary, but is constrained by the physics of cardiac tissue. Specifically, the stimulus cannot be applied directly to interior points of the tissue, but can be applied only at the tissue boundary. The way this stimulus is translated into transmembrane stimulus in the interior of the tissue is related to its resistive inhomogeneity, among other things.

We have shown here that the effect of stimuli on reentrant patterns depends strongly on the spatial scale of the inhomogeneity by which the stimulus is mediated. When the inhomogeneity is of small spatial scale, and the applied field is uniform, the effect of a stimulus is uniform. In the limit that the spatial scale is zero, this conclusion can be verified using homogenization theory (Keener and Panfilov, 1996), and in this situation, defibrillation is a true threshold phenomenon; defibrillation succeeds or fails depending solely on the amplitude of the stimulus. If the applied field and tissue properties are nonuniform (which is the physically realistic case),

then the probability of defibrillation success is an increasing function of stimulus amplitude. Failure to defibrillate can be explained as a failure of some critical domain to be adequately depolarized, leading to the reestablishment of reentrant waves.

The small scale hypothesis produces a theory that agrees with the experimental data in several ways. That the defibrillation threshold should decrease with the addition of heptanol, remain unchanged with application of potassium channel blockers and increase with addition of sodium channel blockers is consistent with this theory. This theory also can be used to show that a biphasic shock is more efficient than a monophasic shock (Keener and Lewis, 1999).

With large spatial scale inhomogeneities, the mechanism of defibrillation is quite different. Here, many new spiral pairs are created and if they are close enough together, they soon collapse, leaving the tissue free of reentrant waves. Since spirals are formed on the boundaries of superthreshold virtual electrodes, spiral pairs will be close together if the superthreshold virtual electrodes nearly cover all of space. It is also required that the cores of preexisting spirals all lie inside a superthreshold virtual electrode. A large domain that is not covered by superthreshold virtual electrodes cannot be defibrillated by this mechanism. Similarly, anatomical reentry is difficult to eliminate by this mechanism, because, as with one-dimensional reentry, elimination of anatomical reentry requires proper alignment with the virtual electrodes. It is experimentally well established that injection of a depolarizing current at the right place at the right time can terminate a rotating wave on a ring, but if the timing is not correct, the rotating wave is merely reset to a different phase (Frame and Rhee, 1988; Glass and Josephson, 1995). In our numerical simulations of FHN dynamics, we also found that reversal of direction was possible, with a properly timed stimulus.

As a consequence of this need for proper alignment, it is quite difficult to eliminate anatomical reentry. Several numerical studies have shown success at eliminating functional reentry with large scale virtual electrodes for a tissue size that was small enough to be covered by a few virtual electrodes (Anderson et al., 2000; Efimov et al., 2000a, b; Trayanova et al., 1998; Eason and Trayanova, 2002). It is the prediction of the large scale theory that if a spiral were to drift and become pinned by an inexcitable anatomical obstacle, it would become much more difficult to eliminate, unless the anatomical obstacle were completely surrounded by a superthreshold virtual electrode.

The main uncertainty for the small scale theory remains the physical source of small spatial scale inhomogeneity. The effect of small scale resistive inhomogeneity due to gap junctions is clearly seen in single isolated cells, but less so in coupled pairs of cells (Sharma and Tung, 2001). In intact tissue, sawtooth

potentials of the amplitude required of this theory have not been seen (Zhou et al., 1998).

However, other sources of small scale resistive inhomogeneity exist and may be more important than gap junctions. For example, recent simulations (Hooks et al., 2002) suggest that the interlaminal clefts between tissue layers may be much more significant than previously thought. The spatial scale of these clefts is of the right order of magnitude for this theory to apply, about six cell widths. Furthermore, most gap junction connections occur as end-to-end couplers, so that lateral resistive inhomogeneity may be more significant than transverse. Other inhomogeneities such as fiber splitting or tapering may play a role as well.

So we are left with an unexplained conundrum. The small-scale hypothesis produces a theory which agrees qualitatively with many of the experimental observations, but the fundamental hypothesis of the nature of the tissue structure has not been verified. On the other hand, the large-scale sources of resistive inhomogeneity that are readily observed seems to work only for a restricted class of reentrant patterns and domains.

## Acknowledgements

This research was supported in part by NSF Grant DMS-99700876 and DMS-0211366.

We acknowledge N. Trayanova for clarifying several misconceptions we had concerning her work, and S. Folias for valuable discussions and numerical simulations.

## References

- AMA Standards, 1986. American heart association standards and guidelines for cardiopulmonary resuscitation and emergency cardiac care. *JAMA* 255, 2841–3044.
- Anderson, C., Trayanova, N., Skouibine, K., 2000. Termination of spiral waves with biphasic shocks: Role of virtual electrode polarization. *J. Cardiovasc. Electrophys.* 11, 1386–1396.
- Beeler, G.W., Reuter, H.J., 1977. Reconstruction of the action potential of ventricular myocardial fibers. *J. Physiol.* 268, 177–210.
- Caulfield, J.B., Borg, T.K., 1979. The collagen network of the heart. *Lab. Invest.* 40, 364–372.
- Choi, B.-R., Nho, W., Liu, T., Salama, G., 2002. Life span of ventricular fibrillation frequencies. *Circ. Res.* 91, 339.
- Crystal, E., Ovsyshcher, I.E., Wagshal, A.B., Katz, A., Ilia, R., 2002. Mexiletine related chronic defibrillation threshold elevation: case report and review of the literature. *PACE* 25, 507–508.
- Cytrynbaum, E., 2001. Using low dimensional models to understand cardiac arrhythmias. Ph.D. Thesis, University of Utah, Salt Lake City, UT.
- Eason, J., Trayanova, N., 2002. Phase singularities and termination of spiral wave reentry. *J. Cardiovasc. Electrophys.* 13, 672–679.
- Efimov, I.R., Aguel, F., Cheng, Y., Wollenzier, B., Trayanova, N., 2000a. Virtual electrode polarization in the far field: implications for external defibrillation. *Am. J. Physiol.* 279, H1055–H1070.

- Efimov, I.R., Gray, R.A., Roth, B.J., 2000b. Virtual electrodes and deexcitation: New insights into fibrillation induction and defibrillation. *J. Cardiovasc. Electrophysiol.* 11, 339–353.
- Fast, V.G., Rohr, S., Gillis, A.M., Kleber, A.G., 1998. Activation of cardiac tissue by extracellular electrical shocks. *Circ. Res.* 82, 375–385.
- Fishler, M.G., 1998. Syncytial heterogeneity as a mechanism underlying cardiac far-field stimulation during defibrillation-level shocks. *J. Cardiovascular Electrophysiology* 9 (4), 384–394.
- Fishler, M.G., Vepa, K., 1998. Spatiotemporal effects of syncytial heterogeneities on cardiac far-field excitations during monophasic and biphasic shocks. *J. Cardiovasc. Electrophysiol.* 9 (12), 1310–1324.
- Frame, L.H., Rhee, E.K., 1988. Reversal of reentry by pacing: relation to termination. *Circulation* 80, 11–96.
- Gillis, A.M., Fast, V.G., Rohr, S., Kleber, A.G., 1996. Spatial changes in transmembrane potential during extracellular electric shocks in cultured monolayers of neonatal rat ventricular myocytes. *Circ. Res.* 79, 676–690.
- Glass, L., Josephson, M.E., 1995. Resetting and annihilation of reentrant abnormally rapid heartbeat. *Phys. Rev. Lett.* 75, 2059–2062.
- Gold, M.R., Higgins, S., Gilliam, F.R., Koppelman, H., Hessen, S., Payne, J., Strickberger, S.A., Breiter, D., Hahn, S., 2002. Efficacy and temporal stability of reduced safety margins for ventricular defibrillation: primary results from the Low Energy Safety Study (LESS). *Circulation* 105, 2043–2048.
- Gradaus, R., Bode-Schnurbus, L., Weber, M., Rotker, J., Hammel, D., Breithardt, G., Bocker, D., 2002. Effect of ventricular fibrillation duration on the defibrillation threshold in humans. *Pacing Clin. Electrophysiol.* 25, 14–19.
- Gray, R.A., Jalife, J., Panfilov, A., Baxter, W.T., Cabo, C., Davidenko, J.M., Pertsov, A.M., 1995. Mechanisms of cardiac fibrillation. *Science* 270, 1222–1223.
- Henriquez, C.S., 1993. Simulating the electrical behavior of cardiac tissue using the bidomain model. *Crit. Revs. Biomed. Eng.* 21, 1–77.
- Hooks, D.A., Tomlinson, K.A., Marsden, S.G., LeGrice, I.J., Smaill, B.H., Pullan, A.J., Hunter, P.J., 2002. Cardiac microstructure: Implications for electrical propagation and defibrillation in the heart. *Circ. Res.* 91, 331–338.
- Ideker, R.E., Tang, A.S.L., Frazier, D.W., Shibata, N., Chen, P.-S., Wharton, J.M., 1991. Basic mechanisms of ventricular defibrillation. In: Glass, L., Hunter, P., McCulloch, A. (Eds.), *Theory of Heart*. Springer, New York, pp. 533–560 (Chapter 20).
- Keener, J., Sneyd, J., 1998. *Mathematical Physiology*. Springer, New York.
- Keener, J.P., 1996. Direct activation and defibrillation of cardiac tissue. *J. Theor. Biol.* 178, 313–324.
- Keener, J.P., 1998. The effect of gap junctional distribution on defibrillation. *Chaos* 8, 175–187.
- Keener, J.P., Lewis, T.J., 1999. The biphasic mystery: why a biphasic shock is more effective than a monophasic shock for defibrillation. *J. Theor. Biol.* 200, 1–17.
- Keener, J.P., Panfilov, A.V., 1996. A biophysical model for defibrillation of cardiac tissue. *Biophys. J.* 71, 1335–1345.
- Knisely, S.B., Hill, B.C., Ideker, R.E., 1994. Virtual electrode effects in myocardial fibers. *Biophys. J.* 66, 719–728.
- Krassowska, W., Neu, J.C., 1994. Response of a single cell to an external electric field. *Biophys. J.* 66, 1768–1776.
- Krassowska, W., Pilkington, T.C., Ideker, R.E., 1987. Periodic conductivity as a mechanism for cardiac stimulation and defibrillation. *IEEE Trans. Biomed. Eng.* 34, 555–559.
- Krassowska, W., Pilkington, T.C., Ideker, R.E., 1990. Potential distribution in three-dimensional periodic myocardium—Part I: Solution with two-scale asymptotic analysis. *IEEE Trans. Biomed. Eng.* 37, 252–266.
- Krinsky, V.I., Pumir, A., 1998. Models of defibrillation of cardiac tissue. *Chaos* 8, 188–203.
- Neu, J.C., Krassowska, W., 1993. Homogenization of syncytial tissues. *Crit. Rev. Biomed. Eng.* 21, 137–199.
- Panfilov, A.V., 1998. Spiral breakup as a model of ventricular fibrillation. *Chaos* 8, 57–64.
- Panfilov, A.V., 1999. Three-dimensional organization of electrical turbulence in the heart. *Phys. Rev. Lett.* 59, R6251–R6254.
- Pumir, A., Krinsky, V.I., 1997. Two biophysical mechanisms of defibrillation of cardiac tissue. *J. Theor. Biol.* 185, 189–199.
- Pumir, A., Romey, G., Krinsky, V.I., 1998. De-excitation of cardiac cells. *Biophys. J.* 74, 2850–2861.
- Qi, X., Varma, P., Newman, D., Dorian, P., 2001. Gap junction blockers decrease defibrillation thresholds without changes in ventricular refractoriness in isolate rabbit hearts. *Circulation* 104, 1544–1549.
- Robinson, T.F., Gould-Cohen, L., Factor, S.M., 1983. Skeletal framework of mammalian heart muscle. *Lab. Invest.* 49, 482–498.
- Sepulveda, N.G., Wikswo, J.P., 1987. Electric and magnetic fields from two-dimensional anisotropic bisyncytia. *Biophys. J.* 51, 557–568.
- Sharma, V., Tung, L., 2001. Theoretical and experimental study of sawtooth effect in isolated cardiac cell-pairs. *J. Cardiovasc. Electrophys.* 12, 1164–1173.
- Sharma, V., Tung, L., 2002. Spatial heterogeneity of transmembrane potential responses of single guinea-pig cells during electric field stimulation. *J. Phys.* 542, 477–492.
- Sobie, E.A., Susil, R.C., Tung, L., 1997. Generalized activating function for predicting virtual electrodes in cardiac tissue. *Biophys. J.* 73, 1410–1423.
- Trayanova, N., 1999. Far-field stimulation of cardiac tissue. *Herz Elektrophysiol.* 43, 1141–1150.
- Trayanova, N., Skouibine, L., 1998. Modelling defibrillation: effects of fiber curvature. *J. Electrophysiol.* 31, 29–32.
- Trayanova, N., Skouibine, K., Aguel, F., 1998. The role of cardiac tissue structure in defibrillation. *Chaos* 8, 221–233.
- Ujhelyi, M.R., Sims, J.J., Miller, A.W., 1999. Induction of electrical heterogeneity impairs ventricular defibrillation: an effect specific to regional conduction velocity slowing. *Circulation* 100, 2534–2540.
- White, J.B., Walcott, G.P., Pollard, A.E., Ideker, R.E., 1998. Myocardial discontinuities: a substrate for producing virtual electrodes that directly excite the myocardium by shocks. *Circulation* 97, 1738–1745.
- Wikswo Jr., J.P., Lin, S.F., Abbas, R.A., 1994. The complexities of cardiac cables: virtual electrode effects. *Biophys. J.* 66, 551b–553b.
- Wikswo Jr., J.P., Lin, S.F., Abbas, R.A., 1995. Virtual electrodes in cardiac tissue: A common mechanism for anodal and cathodal stimulation. *Biophys. J.* 69, 2195–2210.
- Winfree, A.T., 1983. Sudden cardiac death, a problem in topology. *Sci. Am.* 248, 114–161.
- Winfree, A.T., 1997. Rotors, fibrillation and dimensionality. In: Panfilov, A.V., Holden, A.V. (Eds.), *Computational Biology of the Heart*. Wiley, Chichester, pp. 101–135.
- Zhou, X., Knisely, S.B., Smith, W.M., Rollins, D., Pollard, A.E., Ideker, R.E., 1998. Spatial changes in the transmembrane potential during extracellular electric stimulation. *Circ. Res.* 83, 1003–1014.