The Topology of Defibrillation

James P. Keener
Department of Mathematics
University of Utah
Salt Lake City, UT 84112

November 5, 2003

Abstract

We describe how Art Winfree’s ideas about phase singularities can be used to understand the response of cardiac tissue with a random preexisting pattern of reentrant waves (fibrillation) to a large brief current stimulus. This discussion is organized around spatial dimension, beginning with a discussion of reentry on a periodic ring, followed by reentry in a two dimensional planar domain (spiral waves), and ending with consideration of three dimensional reentrant patterns (scroll waves). In all cases we show how reentrant activity is changed by the application of a shock, describing conditions under which defibrillation is successful or not.

Using topological arguments we draw the general conclusion that with a generic placement of stimulating electrodes, large scale virtual electrodes do not give an adequate explanation for the mechanism of defibrillation.

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

1 Introduction

Fibrillation is generally thought to be a highly disorganized pattern of electrical activation of the heart consisting of reentrant “spirals” that are continually created and destroyed (Gray et al. 1995; Panfilov 1998; Panfilov 1999; Choi et al. 2002). An alternate hypothesis (Jalife and Berenfeld 2004) is that fibrillation is organized and sustained by a “mother rotor” with emanating waves of excitation that break apart into the complicated wave pattern typical of fibrillation. Regardless of the underlying mechanism, 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 Joules, which explains why this is called a shock. For implantable defibrillators, the required energy is on the order of 15-20 Joules, which is still considerable.

When it works, shortly after the defibrillating shock has ended, the reentrant activity ceases and the heart approaches its rest state, awaiting normal activation from the sinoatrial node. When it fails, the reentrant activity is temporarily disturbed, but spontaneously returns (Ideker et al. 1991).

An important experimental observation that requires explanation is that the probability of defibrillation success is a sigmoidal function of stimulus amplitude, and with a sufficiently large amplitude stimulus, defibrillation success is almost guaranteed (Gold et al. 2002).

It was Art Winfree’s observation (Winfree 1983; Winfree 1989) that reentrant activity is synonymous with the existence of phase singularities (defined below), and one of his insights was to describe how the application of a stimulus to tissue could create phase singularities, and thereby establish reentrant activity. Moreover, it is now generally understood that for defibrillation to be successful, the stimulus must somehow eliminate all phase singularities. Thus, it is highly relevant to understand how a preexisting pattern of reentrant waves is changed by the application of a large shock.

The purpose of this paper is to describe how shocks act on preexisting patterns of reentrant waves. For convenience, we call this process ”defibrillation”, even though not all reentrant wave activity is identified as fibrillation (e.g., atrial flutter, monomorphic tachycardia). Specifically, we refine Art’s definition of phase singularities and use this to determine the effect of large amplitude shocks on preexisting phase singularities, either creating them, destroying them, or moving them around, thereby gaining insight into the mechanism of defibrillation.

The nature and effect of the applied stimulus is extremely important for this discussion. We assume that a large amplitude current is briefly applied at some boundary of the tissue domain. Of course, since total charge cannot build up, the net current flux across the boundary must be zero. We view cardiac tissue as a bidomain, consisting of both intracellular and extracellular spaces, separated by cell membrane. Current is applied to extracellular space, yet the only currents that matter to a cardiac cell are transmembrane currents. When there is an applied extracellular current, transmembrane currents are generated near the domain boundary, and these must be both depolarizing and hyperpolarizing in regions of about the same size. Away from the boundary there can also be regions with depolarizing and hyperpolarizing transmembrane current, (also with a zero net transmembrane current) generated by local spatial inhomogeneity of resistance. These regions of depolarization and hyperpolarization are well known to be consequences of an anisotropic bidomain model, and are referred to as virtual electrodes, virtual because they may occur at large distances from the stimulating electrodes (Wikswo et al. 1994; Wikswo et al. 1995; Efimov et al. 2000; Knisely et al. 1994).

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. There may also be anatomical obstacles or regions of damaged tissue that generate additional virtual electrodes.

The size and spatial scale of the virtual electrodes is related to the spatial scale of the resistive inhomogeneity that generates them. There is still uncertainty as to which type of resistive inhomogeneity is primarily responsible for defibrillation success.

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; Keener 1998; Keener and Cytrynbaum 2003; Keener and Lewis 1999; Keener and Panfilov 1996; Krinsky and Pumir 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; Krassowska et al. 1990). However, there are many workers in the field who do not accept the small scale hypothesis, 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).

A new proposal that deserves consideration is that interlaminal clefts provide an adequate small scale resistive inhomogeneity (Hooks et al. 2002).

A second popular hypothesis is that large spatial scale inhomogeneities are primarily responsible for defibrillation success (Fast et al. 1998; White et al. 1998; Eason and Trayanova 2002). It is this hypothesis that is examined in this paper. In particular, we extend the ideas of Art Winfree to examine the effect of large spatial scale virtual electrodes on preexisting reentrant patterns, in order to assess the feasibility of successful defibrillation. We find that under generic conditions, the probability of defibrillation success by this mechanism is bounded well below one, in contrast to experimental observations.

The outline of this paper is as follows. In the next section, we extend Art’s definition of a phase singularity. Then in the following sections we show how this can be used to determine the effect of a brief applied stimulus in one, two or three dimensional regions.

2 Phase Singularities in Cardiac Tissue

The fundamental insight that Art had relating to reentrant cardiac arrhythmias was that, like most oscillatory behavior, it is useful to view the oscillation as circular motion on the face of a clock (analog, not digital). As is well-known the minute hand of a clock rotates through $2\pi$ radians every hour, and all that is needed to tell time is the angular direction, or phase, in which the minute hand points; the radius of the hand, provided it is not zero, is arbitrary. However, there is one point on the face of a clock at which time cannot be determined, namely the center, or phaseless point.

To view an oscillation in this manner requires two state variables that oscillate but are out of phase with each other. For cardiac tissue, it is not obvious which two variables to choose. Some investigators have used the transmembrane potential $\phi(t)$ and a delay thereof, $\phi(t - \tau)$ with $\tau$ fixed (Iyer and Gray 2001; Eason and Trayanova 2002). Other possibilities for the second variable include the time derivative of $\phi(t)$, or the Hilbert Transform of $\phi(t)$. 
Of these, the Hilbert transform is probably the best choice, because of its close relationship to $\phi'(t)$ without the associated numerical instabilities (The Hilbert transform of $f$, $H(f)$, is defined by $F(H(f)) = -i \text{sgn}(\mu)F(f)$, where $F(g)$ represents the Fourier transform of $g$ (Keener 1999), whereas the derivative of $\phi(t)$ satisfies $F(\phi') = i\mu F(\phi)$.) However, with none of these choices can one determine how phase singularities are created, eliminated, or moved in response to stimuli.

The clue for what variables to use comes from an examination of two-variable models of excitable media. We suppose that tissue dynamics are described by a fast-slow system of equations,

\begin{align*}
C_m \frac{d\phi}{dt} + I_{ion}(\phi, w) &= I_{in}, \\
\frac{dw}{dt} &= \epsilon g(\phi, w),
\end{align*}

where $w \in \mathbb{R}^n$, $n \geq 1$. A useful (but not physiological) specific example of these are the FitzHugh-Nagumo (FHN) dynamics (Keener and Sneyd 1998) which take the form

\begin{align*}
I_{ion} &= -f(\phi) + \alpha w, \\
\frac{\partial w}{\partial t} &= \epsilon(\phi - \gamma w),
\end{align*}

where $f(\phi) = A\phi(1 - \phi)(\phi - a)$. In this simple model, the rest potential $\phi = 0$ corresponds to the polarized membrane state. A positive stimulus, leading to an increase of $\phi$ is a depolarizing stimulus, and a negative stimulus, leading to a decrease of $\phi$ is a hyperpolarizing stimulus. If the stimulus duration is sufficiently short and not too large, it has little effect on the variable $w$.

In Fig. 1 is shown a typical periodic traveling wave solution on a ring, moving from right to left (parameter values $\alpha = a = 0.1$, $\epsilon = 0.01$, $A = 1$, and $\gamma = \frac{1}{3}$). The upper panel shows the variables $\phi$ 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 $\phi$ and $w$ nullclines, found by setting $I_{ion} = 0$ (the $\phi$ nullcline, cubic shaped) and $\frac{dw}{dt} = 0$ (the $w$ nullcline, a monotone increasing function of $\phi$).

The first observation is that space is subdivided into two regions, excited, with $\phi$ large (of order one) and refractory/recovered, with $\phi$ near zero. The relatively sharp transitions between these regions we identify as fronts or backs. In the $\phi$-$w$ phase plane, fronts and backs are seen as curve segments that connect the left and right branches of the cubic nullcline while keeping $w$ relatively unchanged. The speed of a front or back is defined as that number $c$ for which there is a monotone increasing, heteroclinic trajectory of the equation

\begin{align*}
\frac{d^2\phi}{d\xi^2} - c \frac{d\phi}{d\xi} - I_{ion}(\phi, w) &= 0
\end{align*}

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

\begin{align*}
\int_{\phi_-}^{\phi_+} I_{ion} d\phi.
\end{align*}
Figure 1: Periodic traveling wave solution of the FHN equations, traveling from right to left, with $\phi$ shown as the solid curve, and $w$ shown as dashed, in the upper panel. In the lower panel is the $\phi$-$w$ phase plane trajectory of this same solution, with the $\phi$ and $w$ nullclines shown as dashed curves. Arrows in the phase plane indicate the direction of increasing $x$.

As a result, there is a value of $w$, 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 identified as a front, while a transition above this level has negative speed and is a back (Keener and Sneyd 1998).

The variable $c$ provides a second way to divide state space, namely if $c$ is positive or negative. Together, the $\phi$-$c$ phase plane divides state space into four subregions: I) excited with $c > 0$; II) excited with $c < 0$; III) polarized with $c < 0$ (refractory); and IV) polarized with $c > 0$ (recovered). These four regions divide the face of our “clock” into four quadrants. The phaseless point is at $c = 0, \phi = \phi^*$, where $\phi^*$ is some (not uniquely specified) intermediate value of membrane potential. (See Fig. 2.)

This same division of state space can be made without reference to a two variable model. Fronts and backs divide space into regions that are depolarized (excited) or polarized (refractory or recovering/resting). The polarized region can be divided into two subregions of recovered or refractory tissue by the criterion that a sufficiently large depolarizing stimulus will either propagate away from the stimulus site (a front with $c > 0$) or regress back into the stimulus sight (a back with $c < 0$), respectively. Similarly, depolarized regions can be divided into two subregions by the effect that a large hyperpolarizing stimulus would have on the transitions that are thus created. A transition that moves away from the hyperpolarizing stimulus is a back ($c < 0$) and a transition that collapses back into the hyperpolarized region is a front ($c > 0$). Thus, level surfaces of the two variables $\phi$ and $c$ provide a division of space into quadrants of the face of a clock.

In terms of ionic models, the quantity $c$ can be related (loosely) to the sodium inactivation variable $h$, as follows. If sodium channels are inactivated, then $h$ is small, and action potential fronts cannot propagate so that $c$ is negative. (In making this identification, we are ignoring the role of calcium in action potential propagation). On the other hand if sodium inactivation
is removed, then $h$ is large and action potentials can propagate, so that $c$ is positive.

3 Reentry on a Periodic Ring

Since fibrillation is a state in which there are one or many reentrant waves, the goal of an applied shock 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. We need, therefore, to understand something about how an applied stimulus can modify the structure of a reentrant wave.

Since 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 attracting dynamical states, and the number of such dynamical 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 direction, as well as the uniform rest state. There is always a limit on the number of waves that can fit in a given region for the simple reason that there is a minimal amount of space required to sustain a full action potential.

The different states can be distinguished by a topological criterion, as follows. At any point in time, plot the solution as a curve in $\phi$-$c$ phase space, parameterized by space (as in Fig. 1). (For two variable models it is equivalent to plot in the $\phi$-$w$ phase plane since the relationship between $c$ and $w$ is monotone decreasing.) 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 $\phi^*$ (see Fig. 2). For this discussion, the precise size of the ellipse is not significant, so long as typical periodic traveling waves surround it. 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 counterclockwise rotation about the ellipse, and negative if clockwise. 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 $+1$.

It would be nice if this winding number were an invariant of the flow for all dynamics of cardiac type, but it is not. The reason for this is that under certain conditions fronts and backs can collide and collapse, or fronts can stall and become backs. Since information about the dependence of trajectories on the independent variable $x$ is lost in the phase portrait projection, it is not possible to predict from the winding number alone if such a collapse or reversal will take place. Collapse and reversal are prevented if $|w'(x)|$ is not too large (Cytrynbaum 2001).

Although the winding number is not an invariant, it is useful here because it provides a sufficient condition for successful defibrillation. That is, if the winding number of a trajectory is zero and if $|w'(x)|$ is not too large, then this trajectory is in the attractive basin of the rest state. Thus, if an applied stimulus manages to convert a trajectory to one with winding
Figure 2: A trajectory in the $\phi$-$w$ phase portrait projection with winding number $\pm 1$.

number zero (and if $|w'(x)|$ is not too large), then the reentrant wave will be eliminated by the stimulus. (Remark: The statement that trajectories with winding number zero and $|w'(x)|$ not too large are in the attractive basin of the rest state has not been rigorously established, although the numerical evidence suggests that this is a valid statement (Cytrynbaum 2001). See also (Glass and Josephson 1995; Gedeon and Glass 1998).)

There are several ways to initiate reentrant waves, and they can all be understood using the phase portrait projection.

Suppose we apply two stimuli (called S1 and S2) to a closed ring initially at rest. If the S1 stimulus is depolarizing and large enough, it will initiate two fronts traveling in the opposite direction away from the stimulus site. Because of symmetry, the phase portrait projection of this trajectory must have winding number zero. Indeed, if nothing else happens, the two waves will proceed around the loop, collide, and subsequently collapse. If a second depolarizing stimulus is applied at exactly the same point in space, the subsequent winding number can only be zero, since symmetry has not been broken. However, if the S2 stimulus is applied at a different location at which the medium is refractory, the trajectory in phase space can be converted from a double cover of a single curve (the result of the S1) into a single loop with winding number $\pm 1$. In Figs. 3-4 are shown snapshots of this sequence of events for FHN dynamics. In Fig. 3 are shown two action potentials propagating outward that were initiated by 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 stimulus that was applied slightly left of center. The phase portrait projection 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 self-sustained (reentrant) periodic traveling wave moving from left to right.

A convenient description of this process is that, because of the timing of the S2 stimulus, the two transitions that were created by the depolarizing stimulus were of different types, a front and a back (i.e., the front had $c > 0$ and the back had $c < 0$). This occurs when the depolarized region covers the point $C_-$ on the phase portrait projection. The significance of
Figure 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.

Figure 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.
the S1 stimulus is that it set up a gradient of refractoriness in which the S2 stimulus was applied.

The second way to create transitions is with hyperpolarization. If the S2 stimulus is hyperpolarizing, and if it is applied in a region where there is a traveling front, it may convert the front to a back, thereby changing the winding number to be nonzero. A reentrant wave will have been initiated. This occurs when the hyperpolarized region covers the point \( C_+ \) on the phase portrait projection.

Both of these mechanisms are possible in two variable models and in ionic models.

There is yet a third way that a reentrant wave can be initiated. The above discussion treated \( w \) (or \( c \)) as if it is static during the time the stimulus is applied. This is approximately correct if the time constants governing recovery are large compared to the time interval of the stimulus. However, that recovery variables are affected by both the amplitude and duration of the stimulus can clearly be seen in Fig. 4 where the recovery variable \( w \) increases more rapidly in regions where there is a large depolarizing stimulus (hence the tilt of the loop in the \( \phi-w \) trajectory in Fig. 4. Similarly, the recovery variable \( w \) decreases more rapidly in regions where there is a large hyperpolarizing stimulus. Hence hyperpolarization promotes recovery.

This is also true for ionic models, particularly for sodium inactivation, which has a relatively small time constant. In some models (for example, Hodgkin-Huxley), removal of inactivation increases excitability, so that when the hyperpolarizing stimulus is terminated, the return to rest provides an activating inward current, and an action potential upstroke is created. This depolarizing rebound is called anodal break excitation. Rebound excitation almost certainly does not occur for cardiac tissue (Pertsov et al. 1977; Pertsov et al. 1977).

However, there is a type of break excitation that has been observed for cardiac tissue that relies on closely juxtaposed depolarizing and hyperpolarizing regions. The depolarization creates a region with elevated potential, which, even if it is unrecovered and so cannot propagate outward, can electrotonically excite the neighboring region in which recovery has been promoted by hyperpolarization. Thus, a front can be created in a region which is initially unrecovered if there is adequate nearby hyperpolarization (Roth 1995; Wikswo et al. 1995). A reentrant wave on a 1D ring is created if the depolarization/hyperpolarization is not symmetric, so that one of the transitions is converted to a front while the second transition remains a back. It is interesting to note that this mechanism for creation of reentry does not require that the medium be inhomogeneous before the S2 is applied - the necessary symmetry breaking is provided by the asymmetric arrangement of depolarizing and hyperpolarizing regions. (See (Roth 2000; Winfree 2000) to see how this works in two dimensional space.)

This third mechanism is topologically equivalent to the first (by depolarization alone) if we allow trajectories in the phase portrait projection to move horizontally as well as vertically, while the stimulus is applied. In both cases, the net effect of the stimulus is to create two new transitions, a front and a back, that evolve into a rotating wave.

All of these mechanisms to initiate a reentrant wave have the topological effect of changing the winding number from zero to \( \pm 1 \). However, the goal of defibrillation is to convert a trajectory with nonzero winding number to one with zero winding number.

Suppose the ring size is small enough that it only permits winding numbers of -1, 0 and 1. That is, the only possible states are the rest state and left or right moving 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 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. There are no other possibilities.

It is clear how each of these transitions can be effected. To turn a winding number $±1$ 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 projection 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. 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.

Each of these has three subcases having slightly different mechanistic descriptions. For example, with a depolarizing stimulus applied to a 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. 6), or it might convert a back into a front (Fig. 7). In ionic models, conversion of a back into a front is extremely unlikely, however, these three possibilities are topologically equivalent. In all cases, the depolarized region must contain the point $C_-$ on the phase portrait projection. Similarly, a hyperpolarizing stimulus applied to an action potential might convert a front to a back (Fig. 8), it might initiate a new front and back pair or it might convert a back to a front (see Fig. 8). Again, in ionic models, the latter two are not observed, even though they are theoretically possible and topologically equivalent. In all cases, the hyperpolarized region must contain the point $C_+$.

Now suppose that the transmembrane stimulus is generated by two virtual electrodes. Then there is a region of depolarization and a region of hyperpolarization with about equal area. If the depolarization is of sufficient amplitude and properly timed it could change the winding number by converting fronts into backs. Similarly, if the hyperpolarization is of sufficient amplitude and properly timed, it could change the winding number by converting backs into fronts. If one but not both of these occur, then the resultant winding number will be zero and defibrillation will have been successful.
Figure 6: A depolarizing stimulus applied in the recovered region may produce winding number zero by converting a wavefront into a waveback.

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

Figure 8: A hyperpolarizing stimulus applied in the excited region may produce winding number zero by converting a wavefront to a waveback.
That all of these possibilities can be realized was established by numerical simulations with FHN dynamics, the results of which are shown in Fig. 9. Shown here are the three regions in which there is defibrillation success, phase resetting and propagation reversal. 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 entire region with defibrillation success is quite small, with the probability of success less than 20% for fixed stimulus amplitude.

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 turn a back into a front, as depicted by Fig. 7, even though these are not two-variable dynamics. The effect of the hyperpolarizing stimulus is to turn a front into a back. In our simulations, the hyperpolarizing current required to reverse a front was significantly larger than the 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 counterbalance the large inward sodium current. This current requirement is larger than the amount of depolarizing current required to excite a partially recovered cell.

The probability of defibrillation can be calculated as follows. We define $V^+$ to be that region where the depolarizing stimulus is large enough to excite resting/recovering tissue, and define $V^-$ to be that region where the hyperpolarizing stimulus is large enough to deexcite excited tissue. If there is a single reentrant wave on the ring (the winding number is $\pm 1$), then there are two critical points, namely the two places at which $c = 0$. We identify these points as $C_-$ and $C_+$, with $C_-$ the place in the recovered region where $c = 0$ and $C_+$ the point in the excited region where $c = 0$.

The probability of successful defibrillation is the probability that the winding number is converted from $\pm 1$ to zero, and is the probability that this occurs via depolarization plus...
the probability that this occurs via hyperpolarization. This occurs via depolarization if $C_-$ lies inside $V^+$ and $C_+$ does not lie inside $V^-$. This occurs via hyperpolarization if $C_+$ lies inside $V^-$ and $C_-$ does not lie inside of $V^+$. Thus,

$$P(\text{defibrillation success}) = P(C_- \in V^+ \text{ and } C_+ \not\in V^-) + P(C_+ \in V^- \text{ and } C_- \not\in V^+). \quad (6)$$

These possibilities are depicted in Fig. 10.

It is clear from this that the probability of successful defibrillation is bounded well away from one, even if the amplitude of the stimulus is extremely large so that $V^+$ and $V^-$ cover the entire circle. If this is the case, then the new transitions are pinched into the boundary between $V^+$ and $V^-$, and there can be two fronts, two backs, or a front and a back. Only if the new transitions are both fronts or both backs will the reentrant pattern be eliminated. Said another way, defibrillation success occurs only if $C_+$ and $C_-$ are both contained in $V^+$ or both contained in $V^-$ after the stimulus has ended. The probability that this occurs is less than one.

4 Defibrillation in Two Dimensions

In two spatial dimensions, the projection onto the $\phi$-$c$ plane is a continuous map. As a result, while much of physical space projects to the outer trajectory of state space (corresponding to the depolarized and polarized portions of the action potential), of necessity the interior of the loop must be covered as well. Points which project to the interior of the “phaseless ellipse” in state space constitute the core of spirals in physical space. More specifically, notice that there are two critical level curves, the $\phi$ level curve with $\phi = \phi^*$, and the $c = 0$ level curve. The physical region with $\phi > \phi^*$ we identify as excited, while if $\phi < \phi^*$ the region is refractory/recovered. The $\phi = \phi^*$ level curve has two components, namely fronts, where $c > 0$, and backs where $c < 0$. Similarly, the $c = 0$ level curve divides space into two regions, with $c > 0$ where transitions must be fronts, and with $c < 0$ where transitions must be backs. The level curve $c = 0$ is also divided into two components, namely $C_-$ where $c = 0$ and the tissue is resting/refractory, and $C_+$, where $c = 0$ and the tissue is excited. Intersections of these two level curves are phaseless points which we identify as the cores of spirals. (See Fig. 11.)

Associated with each core there is a winding number of $\pm 1$, determined as follows: Traverse a circle in physical space that encloses a core in the clockwise direction and follow the
variables $\phi$ and $c$ through the phases of the clock. If the face of the clock is traversed in the clockwise direction (i.e. sequentially through phases I, II, III, and IV), the winding number is $+1$ whereas if it is traversed in the counterclockwise direction the winding number is $-1$. For example, the spiral core shown in Fig. 11 has winding number $-1$. The total parity (or total topological charge) of a region of space is the sum of the winding numbers of all spiral cores contained within the region.

The winding number can be defined for two variable models using the $\phi$-$w$ phase plane projection (as in Fig. 2) if one associates winding number $+1$ with trajectories that traverse the loop around the phaseless point in the counterclockwise direction. This difference in direction between the $\phi$-$w$ and $\phi$-$c$ planes is necessitated by the fact that $c$ is a decreasing function of $w$.

Now the question of how to initiate spirals arises, and this was answered initially by Wiener and Roseblueth (Wiener and Rosenblueth 1946) and further developed by Art Winfree (Winfree 1983; Winfree 1989). Art’s explanation was to suppose that there is a plane wave, with resting tissue ahead of the wave, an excited region, followed by a refractory/recovered region. If a sufficiently large region of the refractory/recovered space is depolarized with a stimulus, then two spiral cores will be created from which a double spiral will emerge. The explanation for this was topological, arguing that there were four regions that covered the four phases of a clock, namely recovered with a large stimulus, refractory with a large stimulus, resting with a small stimulus, and refractory with a small stimulus. Since these four regions cover the face of a clock, the behavior that ensues should continue to cover the face of a clock and a spiral core will result.

The description we use here makes Art’s argument a bit more precise. We define $V^+$ to be that region of space that receives a depolarizing stimulus that is large enough to excite any tissue for which $c > 0$. Then that part of the boundary of $V^+$ which is not already excited and for which $c > 0$ will become a new front and the boundary of $V^+$ which is not already excited for which $c < 0$ will become a back. A new spiral core will be created wherever the boundary of $V^+$ intersects the level surface $C_-$. Because $V^+$ is a closed domain, the total parity (or total topological charge) of a region of space is the sum of the winding numbers of all spiral cores contained within the region.

Figure 11: Division of space into four subregions by a spiral core.
Figure 12: The Winfree protocol; Four subdivisions of space shown with an applied depolarizing stimulus that lies in regions III and IV.

parity of the new spiral pair will be zero. This is depicted in Figs. 12 and 13.

It is apparent that a hyperpolarizing stimulus can have the symmetrically opposite effect. We define $V^-$ to be the region of space that receives a hyperpolarizing stimulus that is large enough to deexcite tissue. That part of the boundary of $V^-$ which is already excited and for which $c < 0$ will become a new back, while that part of the boundary of $V^-$ that is already excited and has $c > 0$ will become a new front. A new spiral core will emerge from each intersection of the boundary of $V^-$ with the level set $C_+$. Because $V^-$ is a closed domain, the number of new cores must be even, and the total parity of the new pattern will be zero. This scenario is depicted in Figs. 14 and 15.

As with one dimensional rings, there is yet another way to initiate spirals with break excitation (Efimov et al. 2000; Roth 2000). In this scenario a depolarizing region is closely juxtaposed with a hyperpolarizing region (Fig. 16). The depolarization creates a back transition, but near the hyperpolarization, a portion of the back transition is converted to a front transition, creating two new spiral cores. The topology of this can be understood by observing that the hyperpolarization has the effect of dynamically moving the critical curve $C_-$ outside its boundary so that after the stimulus has ended, $C_-$ intersects the depolarizing electrode region $V^+$ (Fig. 17). This is topologically, if not mechanistically, equivalent to the first method. The end result is that two new spiral cores of opposite parity are created.

Now consider what happens when a shock is applied to tissue where there are preexisting spirals. For example, in Fig. 18 is shown a pattern of spiral cores with fronts and backs, including a “virtual core” with rotation around an anatomical obstacle, immediately before application of the stimulus.

Upon application of the stimulus, the two dimensional region is divided into three sub-regions, $V^+, V^-$, and all the rest where the stimulus amplitude is too small to effect any topological changes (Fig. 19).

After the stimulus has ended, all tissue in $V^+$ will be excited and all tissue in $V^-$ will
Figure 13: Pair of spiral cores that is created following application of the Winfree protocol.

Figure 14: A hyperpolarizing stimulus that lies in regions I and II.
Figure 15: Pair of spiral cores that is created following application of a hyperpolarizing stimulus.

Figure 16: A pattern of depolarization and hyperpolarization preceding break excitation.
Figure 17: A break excitation occurs when the curve $C_-$ is moved by the hyperpolarizing stimulus so that it intersects the depolarized region, converting fronts to backs and creating a new spiral pair.

Figure 18: Pattern of cores, fronts, and backs, immediately preceding application of a stimulus. Darkened circle in lower right hand corner is to depict an anatomical obstacle around which there is circulating a reentrant wave. Total parity for this pattern is -2.
be deexcited, and new spiral cores will be created at the intersections of the boundary of $V^+$ with $C_-$ and the boundary of $V^-$ with $C_+$. Preexisting spiral cores that lie within $V^+$ will be moved to the boundary of $V^+$ if the curve $C_-$ intersects the boundary of $V^+$, while preexisting spiral cores that lie within $V^-$ will be moved to the boundary of $V^-$ if the curve $C_+$ intersects the boundary of $V^+$ (Fig. 20). A spiral pair for which the curve $C_-$ lies entirely within $V^+$ will be destroyed by depolarization, as will a spiral pair for which the curve $C_+$ lies entirely within $V^-$ be destroyed by hyperpolarization.

It follows that the total parity in any region which completely encloses regions $V^+$ and $V^-$ is unchanged by the application of the stimulus. This is illustrated by Figs. 21 and 22 for a single depolarizing region, but the same is true for hyperpolarizing regions. This conclusion is not modified by break excitations since break excitation cannot change the total parity.

It is noteworthy that the change in parity between Fig. 18 and Fig. 20 is due entirely to a hyperpolarizing region that intersects the boundary. Indeed, the only way that the overall parity can be changed is with depolarizing or hyperpolarizing regions that intersect the domain boundary. The fact that the extra core in Fig. 18 is close to a boundary and is likely to subsequently collapse into the boundary is coincidental.

Now that we see how spirals are created, destroyed or moved, it remains to determine when defibrillation will be successful. It is apparent that if this protocol is to be successful, i.e., if all spiral cores are to be eliminated, then some additional mechanisms must be involved. This procedure by itself cannot successfully defibrillate with probability close to one.

For defibrillation to be successful, spirals must collapse spontaneously after they have been moved or created by the stimulus. This can happen if spiral pairs collapse or if individual spirals move across an exterior boundary. (It is only necessary to consider exterior boundaries here since if spirals move across interior boundaries, they convert from “functional” to “anatomical” reentry, but do not change the overall parity of the domain.)
Figure 20: New pattern of fronts, backs and spiral cores following application of the stimulus in Fig. 19. Total parity for this new pattern is -3.

Figure 21: Large depolarization region applied to the pattern of fronts, backs and spiral cores shown in Fig. 18.
Figure 22: New pattern of fronts, backs and spiral cores following application of the depolarization stimulus shown in Fig. 21. Total parity for this new pattern is -2.

spirals that are not sufficiently close to a boundary, there are two requirements. First, their total parity must be zero after the stimulus has ended, and second, the resulting spiral cores must be pairwise sufficiently close together that their subsequent dynamics cause them to spontaneously collapse.

It is worth noting that the problem of converting an anatomical reentry into a functional reentry is topologically equivalent to eliminating a reentrant wave on a one dimensional ring, where the one dimensional ring is identified with the boundary of anatomical barrier. According to the bidomain tissue model, the boundary of the anatomical region will have subdomains which are depolarizing and hyperpolarizing. (i.e., it is not physically possible to depolarize or hyperpolarize the entire boundary.) It follows that the probability of converting anatomical reentry to functional reentry is bounded away from one.

In the same way, the problem of assuring that the interior region has total parity zero is topologically equivalent to a one dimensional problem. It is not necessary to know the details of the spiral structure in the interior of the domain to calculate its total parity.

Suppose we define the boundary zone to be that part of the tissue in which spirals quickly collapse into the exterior tissue boundary, and the interior zone as the complementary region in which spirals do not collapse quickly into the exterior boundary. Suppose further that we define $\partial V^+$ and $\partial V^-$ to be the intersections of $V^+$ and $V^-$, respectively, with the boundary of the interior zone. Now, the probability that the total parity of interior spirals is zero is the same as the probability that the total parity of the interior zone is zero after the stimulus is applied. If there is exactly one interior spiral before the stimulus is applied, then the probability that the total parity is zero afterward is exactly as expressed in (6), namely

$$P(\text{postshock parity } = 0) = P(C_- \in \partial V^+ \text{ and } C_+ \notin \partial V^-) + P(C_+ \in \partial V^- \text{ and } C_- \notin \partial V^+).$$

(7)

However, if the initial parity is zero with two spirals (say) of parity $\pm 1$, then to achieve total
parity zero after the shock is applied, there is a different (more complicated) expression, but the idea is the same. The only way to be assured that the interior parity is zero is to have one of $\partial V^+$ or $\partial V^-$ cover the entire boundary of the interior zone, as for example, in Figs. 21 and 22.) Otherwise, because of the constraints on the arrangement of the curves $C_+$ and $C_-$ with respect to $\partial V^+$ and $\partial V^-$, the probability that the total parity is zero after the stimulus is ended is substantially less than one, as it was in the one dimensional case.

Since it is highly unlikely that exactly one of $\partial V^+$ or $\partial V^-$ will cover the entire interior zone boundary, we conclude that the probability of defibrillation success must be bounded below one.

However, if the total parity is zero after the stimulus is ended, there remains the additional difficulty that the new spiral core pairs must be sufficiently close together that they spontaneously collapse. From Fig. 23 it is apparent that the three things that determine the separation of newly formed spiral pairs are the orthogonal distance between the boundary of $V^+$ and $V^-$ (assuming them to be parallel), the orthogonal separation between $C_+$ and $C_-$ (also assuming them to be parallel, denoted $d$ in Fig. 23), and the angle of intersection between $C_-$ and $V^+$ (denoted $\theta$ in Fig. 23). The one of these over which there is some control is the orthogonal distance between the boundary of $V^+$ and $V^-$, which is controlled by the amplitude of the stimulus, and goes to zero as the stimulus amplitude becomes large. The orthogonal separation between $C_+$ and $C_-$ is determined by the dynamics of the spiral, and the angle of intersection between $C_-$ and $V^+$ is random and depends on the timing of the stimulus in relation to the motion of the preexisting spirals. The separation between newly formed spiral pairs is minimized if the spiral arms are orthogonal to the boundary of $V^+$ and $V^-$ (i.e., $\theta = \frac{\pi}{2}$), but if $\theta$ is small, then the separation between newly formed spirals becomes large, and it is less likely that the new spiral pairs will collapse.

We conclude that with this mechanism, the probability of defibrillation success is bounded away from one, even with a large stimulus amplitude, unless special (i.e., non-generic) geometrical and physical features are present. This is because the probability that the after shock total parity is zero and the probability that all newly formed spiral pairs are close enough together to quickly collapse are both bounded below one.
5 Defibrillation in Three Dimensions

The preceding discussion was interesting for topological reasons, but is of little physical relevance. This is for the simple reason that real cardiac tissue is either three dimensional (it has significant thickness) or if it can be viewed as two dimensional (as with atrial tissue), stimuli are applied transmurally and not at the lateral edges of the domain.

For three dimensional tissue, reentrant waves must be understood to be scroll waves; spirals on the epicardial or endocardial surface are two dimensional cross-sections of three dimensional scroll wave activity. Similarly, spiral cores on the endocardial surface are the two dimensional crossections of scroll wave filaments. (This was another of Art Winfree's insights into the behavior of reentry (Winfree 1973; Winfree 1987).)

When a stimulus is applied to three-dimensional tissue, there are regions $V^+$ and $V^-$ in which the stimulus is depolarizing and hyperpolarizing, respectively. However, these are three dimensional subregions whose boundaries are two dimensional surfaces. Similarly, the level surface $c = 0$, with components $C_+$ and $C_-$, is a two dimensional surface which divides space into two subregions.

In three dimensions, the intersection of two surfaces is a curve. When a stimulus is applied to a three-dimensional tissue, scroll wave filaments are created at the intersection of the boundary of $V^+$ with $C_-$ and at the intersection of the boundary of $V^-$ with $C_+$. An example of this is depicted in Fig. 24.

It follows immediately that the effect of a stimulus is either to move preexisting filaments or to create new ones. A closed filament that is completely contained within a region $V^+$ or $V^-$ may be eliminated, depending on the timing. However, all filaments are squeezed into the region not covered by $V^+$ and $V^-$. This has interesting consequences for a filament that is originally attached to a domain.
boundary. In particular, unless the boundary is completely enclosed by a single depolarizing or hyperpolarizing region, the new filament must remain attached to the boundary. This is in contrast to one and two dimensional domains where a core can be moved away from the boundary if the stimulus timing is correct, even if the boundary is not completely covered by a single stimulating region.

According to the bidomain model, it is impossible to surround an anatomical barrier (an interior non-conducting domain) with a single depolarizing or hyperpolarizing region. Thus, it is impossible to detach a filament from an anatomical barrier with virtual electrodes.

Similarly, suppose there is a single preexisting transmural scroll filament. If this is transformed by the stimulus into another transmural filament, then there is no reason to expect it to spontaneously disappear - transmural scrolls are extremely robust. The only way to be certain that the new scroll is entirely intramural is for $V^+$ and $V^-$ to have no intersections with the endocardial and epicardial surfaces. But without special placement of electrodes, this is impossible to achieve (since the net current across the boundary must be zero), and we expect a transmural scroll wave to reappear after the stimulus has ended.

In other words, the probability of successful defibrillation of three dimensional ventricular tissue without careful placement of stimulating electrodes is substantially less than one for all stimulus amplitudes.

6 Defibrillation of Atrial Tissue

An interesting and important difference between ventricular tissue and atrial tissue is that atrial tissue is quite thin. As a result, the foregoing arguments do not apply to atrial tissue. Since the stimulating currents are applied transmurally, the arguments of Section 4 do not apply, and since the tissue is quite thin, the arguments of the last section do not apply. For atrial tissue, depolarization and hyperpolarization is achieved by a boundary effect, not by virtual electrodes.

It is beyond the scope of this paper to describe in detail an appropriate model of defibrillation for atrial tissue. Suffice it to say that the transmural profiles of transmembrane potential are very much like sawtooth potentials, and a singular perturbation argument suggests a model similar to those models found by averaging when there are small scale resistive inhomogeneities (Keener 1996; Keener 1998; Keener and Lewis 1999; Keener and Panfilov 1996; Krinsky and Pumir 1998). Therefore, the mechanism of defibrillation in atrial tissue is quite similar to that which is produced by small scale resistive inhomogeneities in ventricular tissue, without the need for any interior inhomogeneities of resistance or anisotropy.

7 Discussion

It was Art Winfree’s idea to use topological arguments to think about spiral and scroll waves. His arguments enabled us to achieve a deeper understanding of the formation and structure of these waves. Here we have modified Art’s definitions and used them to study how spirals and scrolls are created, destroyed or moved by application of an external stimulus.

Using general arguments, we conclude that the probability of defibrillation success using large scale virtual electrodes is bounded away from one, regardless of the amplitude of the
stimulus. This is because elimination of all spirals or scrolls requires a highly restrictive structure for the stimulating electrodes as well as proper timing of the stimulus. Those (numerical) situations where defibrillation success is achieved are explained by the fact that the tissue domain and placement of electrodes were carefully chosen so that they cannot be considered “generic”.

The conclusions of these topological arguments are clearly contradicted by the experimental data. It is well established experimentally that the probability of defibrillation success is an increasing function of stimulus amplitude, and approaches one for large enough stimulus amplitudes. We conclude from this that large scale virtual electrodes do not adequately explain the mechanism of defibrillation, but that some other mechanism must account for the success. The mechanism we favor remains the small scale hypothesis, however, further discussion of this hypothesis must be relegated to other publications (for example, (Keener and Cytrynbaum 2003)).

References


