2007

Thermodynamically Equivalent Silicon Models of Voltage-Dependent Ion Channels

Kai M. Hynna
University of Pennsylvania

Kwabena A. Boahen
University of Pennsylvania, boahen@seas.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/be_papers
Part of the Biomedical Engineering and Bioengineering Commons

Recommended Citation

Suggested Citation:

© 2007 Massachusetts Institute of Technology
http://www.mitpressjournals.org/loi/neco

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/be_papers/177
For more information, please contact repository@pobox.upenn.edu.
Thermodynamically Equivalent Silicon Models of Voltage-Dependent Ion Channels

Abstract
We model ion channels in silicon by exploiting similarities between the thermodynamic principles that govern ion channels and those that govern transistors. Using just eight transistors, we replicate—for the first time in silicon—the sigmoidal voltage dependence of activation (or inactivation) and the bell-shaped voltage-dependence of its time constant. We derive equations describing the dynamics of our silicon analog and explore its flexibility by varying various parameters. In addition, we validate the design by implementing a channel with a single activation variable. The design's compactness allows tens of thousands of copies to be built on a single chip, facilitating the study of biologically realistic models of neural computation at the network level in silicon.

Disciplines
Biomedical Engineering and Bioengineering | Engineering

Comments
Suggested Citation:

© 2007 Massachusetts Institute of Technology
http://www.mitpressjournals.org/loi/neco

This journal article is available at ScholarlyCommons: http://repository.upenn.edu/be_papers/177
We model ion channels in silicon by exploiting similarities between the thermodynamic principles that govern ion channels and those that govern transistors. Using just eight transistors, we replicate—for the first time in silicon—the sigmoidal voltage dependence of activation (or inactivation) and the bell-shaped voltage-dependence of its time constant. We derive equations describing the dynamics of our silicon analog and explore its flexibility by varying various parameters. In addition, we validate the design by implementing a channel with a single activation variable. The design’s compactness allows tens of thousands of copies to be built on a single chip, facilitating the study of biologically realistic models of neural computation at the network level in silicon.
Neurobiologists perform experiments on real cells, both in vivo and in vitro; however, while working with real cells eliminates the need for justifying assumptions within a model, limitations in technology restrict recordings to tens of neurons. Computational neuroscientists create computer models of real cells, using simulations to test the function of a channel within a single cell or a network of cells. While providing a great deal of flexibility, computational neuroscientists are often limited by the processing power of the computer and must constantly balance the complexity of the model with practical simulation times. For instance, a Sun Fire 480R takes 20 minutes to simulate 1 second of the 4000-neuron network (M. Shelley, personal communication, 2004) of Tao, Shelley, McLaughlin, and Shapley (2004), just 1 mm² (a single hypercolumn) of a sublayer (4Cα) of the primary visual cortex (V1).

An emerging medium for modeling neural circuits is the silicon chip, a technique at the heart of neuromorphic engineering. To model the brain, neuromorphic engineers use the transistor’s physical properties to create silicon analogs of neural circuits. Rather than build abstractions of neural processing, which make gross simplifications of brain function, the neuromorphic engineer designs circuit components, such as ion channels, from which silicon neurons are built. Silicon is an attractive medium because a single chip can have thousands of heterogeneous silicon neurons that operate in real time. Thus, network phenomena can be studied without waiting hours, or days, for a simulation to run.

To date, however, neuromorphic models have not captured the voltage dependence of the ion channel’s temporal dynamics, a problem outstanding since 1991, the year Mahowald and Douglas (1991) published their seminal work on silicon neurons. Neuromorphic circuits are constrained by surface area on the silicon die, limiting their complexity, as more complex circuits translate to fewer silicon neurons on a chip. In the face of this trade-off, previous attempts at designing neuromorphic models of voltage-dependent ion channels (Mahowald and Douglas, 1991; Simoni, Cymbalyuk, Sorensen, Calabrese, & DeWeerth, 2004) sacrificed the time constant’s voltage dependence, keeping the time constant fixed. In some cases, however, this nonlinear property is critical. An example is the low-threshold calcium channel in the thalamus’s relay neurons. The time constant for inactivation can vary over an order of magnitude depending on the membrane voltage. This variation defines the relative lengths of the interburst interval (long) and the burst duration (short) when the cell bursts rhythmically.

Sodium channels involved in spike generation also possess a voltage-dependent inactivation time constant that varies from a peak of approximately 8 ms just below spike threshold to as fast as 1 ms at the peak of the voltage spike (Hodgkin & Huxley, 1952). Ignoring this variation by fixing the time constant alters the dynamics that shape the action potential. For example, lowering the maximum (peak) time constant would reduce the
size of the voltage spike due to faster inactivation of sodium channels below threshold. Inactivation of these channels is a factor in the failure of action potential propagation in Purkinje cells (Monsivais, Clark, Roth, & Hausser, 2005). On the other hand, increasing the minimum time constant at more depolarized levels—that is, near the peak voltage of the spike—would increase the width of the action potential, as cell repolarization begins once the potassium channels overcome the inactivating sodium channel. A wider spike could influence the cell’s behavior through several mechanisms, such as those triggered by increased calcium entry through voltage-dependent calcium channels.

In this letter, we present a compact circuit that models the nonlinear dynamics of the ion channel’s gating particles. Our circuit is based on linear thermodynamic models of ion channels (Destexhe & Huguenard, 2000), which apply thermodynamic considerations to the gating particle’s movement, due to conformation of the ion channel protein in an electric field. Similar considerations of the transistor make clear that both the ion channel and the transistor operate under similar principles. This observation, originally recognized by Carver Mead (1989), allows us to implement the voltage dependence of the ion channel’s temporal dynamics, while at the same time using fewer transistors than previous neuromorphic models that do not possess these nonlinear dynamics. With a more compact design, we can incorporate a larger number of silicon neurons on a chip without sacrificing biological realism.

The next section provides a brief tutorial on the similarities between the underlying physics of thermodynamic models and that of transistors. In section 3, we derive a circuit that captures the gating particle’s dynamics. In section 4, we derive the equations defining the dynamics of our variable circuit, and section 5 describes the implementation of an ion channel population with a single activation variable. Finally, we discuss the ramifications of this design in the conclusion.

2 Ion Channels and Transistors

Thermodynamic models of ion channels are founded on Hodgkin and Huxley’s (empirical) model of the ion channel. A channel model consists of a series of independent gating particles whose binary state—open or closed—determines the channel permeability. A Hodgkin-Huxley (HH) variable represents the probability of a particle being in the open state or, with respect to the channel population, the fraction of gating particles that are open. The kinetics of the variable are simply described by

\[ \alpha(V) (1 - u) \leftrightarrow u, \beta(V) \]  

(2.1)
where \( \alpha(V) \) and \( \beta(V) \) define the voltage-dependent transition rates between the states (indicated by the arrows), \( u \) is the HH variable, and \((1-u)\) represents the closed fraction.

\( \alpha(V) \) and \( \beta(V) \) define the voltage-dependent dynamics of the two-state gating particle. At steady state, the total number of gating particles that are opening—that is, the opening flux, which depends on the number of channels closed and the opening rate—are balanced by the total number of gating particles that are closing (the closing flux). Increasing one of the transition rates—through, for example, a shift in the membrane voltage—will increase the respective flow of particles changing state; the system will find a new steady state with new open and closed fractions such that the fluxes again cancel each other out.

We can describe these dynamics simply using a differential equation:

\[
\frac{du}{dt} = \alpha(V)(1-u) - \beta(V)u. \tag{2.2}
\]

The first term represents the opening flux, the product of the opening transition rate \( \alpha(V) \) and the fraction of particles closed \((1-u)\). The second term represents the closing flux. Depending on which flux is larger (opening or closing), the fraction of open channels \( u \) will increase or decrease accordingly.

Equation 2.2 is often expressed in the following form:

\[
\frac{du}{dt} = -\frac{1}{\tau_u(V)}(u - u_\infty(V)), \tag{2.3}
\]

where

\[
u_\infty(V) = \frac{\alpha(V)}{\alpha(V) + \beta(V)}, \tag{2.4}\]

\[
\tau_u(V) = \frac{1}{\alpha(V) + \beta(V)}, \tag{2.5}\]

represent the steady-state level and time constant (respectively) for \( u \). This form is much more intuitive to use, as it describes, for a given membrane voltage, where \( u \) will settle and how fast. In addition, these quantities are much easier for neuroscientists to extract from real cells through voltage-clamp experiments. We will come back to the form of equation 2.3 in section 4. For now, we will focus on the dynamics of the gating particle in terms of transition rates.

In thermodynamic models, state changes of a gating particle are related to changes in the conformation of the ion channel protein (Hill & Chen, 1972; Destexhe & Huguenard, 2000, 2001). Each state possesses a certain energy
Figure 1: Energy diagram of a reaction. The transition rates between two states are dependent on the heights of the energy barriers ($\Delta G_C$ and $\Delta G_O$), the differences in energy between the activated state ($G^*$) and the initial states ($G_C$ or $G_O$). Thus, the time constant depends on the height of the energy barriers, and the steady state depends on the difference in energy between the closed and open states ($\Delta G$).

(see Figure 1), dependent on the interactions of the protein molecule with the electric field across the membrane. For a state transition to occur, the molecule must overcome an energy barrier (see Figure 1), defined as the difference in energy between the initial state and an intermediate activated state. The size of the barrier controls the rate of transition between states (Hille, 1992):

$$\alpha(V) = \alpha_0 e^{-\Delta G_C(V)/RT}$$

$$\beta(V) = \beta_0 e^{-\Delta G_O(V)/RT},$$

where $\alpha_0$ and $\beta_0$ are constants representing base transition rates (at zero barrier height), $\Delta G_C(V)$ and $\Delta G_O(V)$ are the voltage-dependent energy barriers, $R$ is the gas constant, and $T$ is the temperature in Kelvin.

Changes in the membrane voltage of the cell, and thus the electric field across the membrane, influence the energies of the protein’s conformations differently, changing the sizes of the barriers and altering the transition rates between states. Increasing a barrier decreases the respective transition rate, slowing the dynamics, since fewer proteins will have sufficient energy. The steady state depends on the energy difference between the two states. For a difference of zero and equivalent base transition rates, particles are
equally distributed between the two states. Otherwise, the state with lower energy is the preferred one.

The voltage dependence of an energy barrier has many components, both linear and nonlinear. Linear thermodynamic models, as the name implies, assume that the linear voltage dependence dominates. This dependence may be produced by the movement of a monopole or dipole through an electric field (Hill & Chen, 1972; Stevens, 1978). In this situation, the above rate equations simplify to

\[
\alpha (V) = A e^{-b_1 (V-V_{1/2})/RT}, \quad (2.8)
\]
\[
\beta (V) = A e^{-b_2 (V-V_{1/2})/RT}, \quad (2.9)
\]

where \(V_{1/2}\) and \(A\) represent the half-activation voltage and rate, while \(b_1\) and \(b_2\) define the linear relationship between each barrier and the membrane voltage. The magnitude of the linear term depends on such factors as the net movement of charge or net change in charge due to the conformation of the channel protein. Thus, ion channels use structural differences to define different membrane voltage dependencies.

While linear thermodynamic models have simple governing equations, they possess a significant flaw: time constants can reach extremely small values at voltages where either \(\alpha (V)\) and \(\beta (V)\) become large (see equation 2.5), which is unrealistic since it does not occur in biology. Adding nonlinear terms in the energy expansion of \(\alpha (V)\) and \(\beta (V)\) can counter this effect (Destexhe & Huguenard, 2000). Other solutions involve either saturating the transition rate (Willms, Baro, Harris-Warrick, & Guckenheimer, 1999) or using a three-state model (Destexhe & Huguenard, 2001), where the forward and reverse transition rates between two of the states are fixed, effectively setting the maximum transition rate. Linear models, however, bear the closest resemblance to the MOS transistor, which operates under similar thermodynamic principles.

Short for metal oxide semiconductor, the MOS transistor is named for its structure: a metallic gate (today, a polysilicon gate) atop a thin oxide, which insulates the gate from a semiconductor channel. The channel, part of the body or substrate of the transistor, lies between two heavily doped regions called the source and the drain (see Figure 2). There are two types of MOS transistors: negative or n-type (NMOS) and positive or p-type (PMOS). NMOS transistors possess a drain and a source that are negatively doped—areas where the charge carriers are negatively charged electrons. These two areas exist within a p-type substrate, a positively doped area, where the charge carriers are positively charged holes. A PMOS transistor consists of a p-type source and drain within an n-type well. While the rest of this discussion focuses on NMOS transistor operation, the same principles apply to PMOS transistors, except that the charge carrier is of the opposite sign.
In the subthreshold regime, charge flows across the channel by diffusion from the source end of the channel, where the density is high, to the drain, where the density is low. Governed by the same laws of thermodynamics that govern protein conformations, the density of charge carriers at the source and drain ends of the channel depends exponentially on the size of the energy barriers there (see Figure 3). These energy barriers exist due to a built-in potential difference, and thus an electric field, between the channel and the source or the drain. Adjusting the voltage at the source, or the drain, changes the charge carriers’ energy level. For the NMOS transistor’s negatively charged electrons, increasing the source voltage decreases the energy level; hence, the barrier height increases. This decreases the charge density at that end of the channel, as fewer electrons have the energy required to overcome the barrier. The voltage applied to the gate, which influences the potential at the surface of the channel, has the opposite effect: increasing it (e.g., from $V_G$ to $V_G1$ in Figure 3) decreases the barrier height—at both ends of the channel.

Factoring in the exponential charge density dependence on barrier height yields the relationship between an NMOS transistor’s channel current and its terminal voltages (Mead, 1989):

$$I_{ds} = I_{ds0} \left( e^{\frac{e(V_G-V_S)}{kT}} - e^{\frac{e(V_G-V_D)}{kT}} \right), \quad (2.10)$$

where $\kappa$ describes the relationship between the gate voltage and the potential at the channel surface. $U_T$ is called the thermal voltage (25.4 mV at room temperature), and $I_{ds0}$ is the baseline diffusion current, defined by the barrier introduced when the oppositely doped regions (p-type and n-type) were fabricated. Note that, for clarity, $U_T$ will not appear in the
remaining transistor current equations, as all transistor voltages from here on are given in units of $U_T$. When $V_{DB}$ exceeds $V_{SB}$ by $4U_T$ or more, the drain term becomes negligible and is ignored; the transistor is then said to be in saturation.

The similarities in the underlying physics of ion channels and transistors allow us to use transistors as thermodynamic isomorphs of ion channels. In both, there is a linear relationship between the energy barrier and the controlling voltage. For the ion channel, either isolated charges or dipoles of the channel protein have to overcome the electric field created by the voltage across the membrane. For the transistor, electrons, or holes, have to overcome the electric field created by the voltage difference between the source, or drain, and the transistor channel. In both instances, the transport of charge across the energy barrier is governed by a Boltzman distribution, which results in an exponential voltage dependence. In the next section, we use these similarities to design an efficient transistor representation of the gating particle dynamics.

3 Variable Circuit

Based on the discussion from the previous section, it is tempting to think we may be able to use a single transistor to model the gating dynamics of a channel particle completely. However, obtaining the transition rates solves only part of the problem. We still need to multiply the rates with the number of gating particles in each state to obtain the opening and closing fluxes, and then integrate the flux difference to update the particle counts.
Figure 4: Channel variable circuit. (a) The voltage $u_V$ represents the logarithm of the channel variable $u$. $V_O$ and $V_C$ are linearly related to the membrane voltage, with slopes of opposite sign. $u_H$ and $u_L$ are adjustable bias voltages.
(b) Two transistors (N3 and N4) are added to saturate the variable’s opening and closing rates; the bias voltages $u_{\tau_H}$ and $u_{\tau_L}$ set the saturation level.

(see equation 2.2). A capacitor can perform the integration if we use charge to represent particle count and current to represent flux. The voltage on the capacitor, which is linearly proportional to its charge, yields the result.

The first sign of trouble appears when we attempt to connect a capacitor to the transistor’s source (or drain) terminal to perform the integration. As the capacitor integrates the current, the voltage changes, and hence the transistor’s barrier height changes. Thus, the barrier height depends on the particle count, which is not the case in biology; gating particles do not (directly) affect the barrier height when they switch state. Our only remaining option, the gate voltage, is unsuitable for defining the barrier height, as it influences the barrier at both ends of the channel identically. $\alpha(V)$ and $\beta(V)$, however, demonstrate opposite dependencies on the membrane voltage; that is, one increases while the other decreases.

We can resolve this conundrum by connecting two transistors to a single capacitor (see Figure 4a). Each transistor defines an energy barrier for one of the transition rates: transistor N1 uses its source and gate voltages ($u_L$ and $V_C$, respectively) to define the closing rate, and transistor N2 uses its drain and gate voltages ($u_H$ and $V_O$) to define the opening rate (where $u_H > u_L$). We integrate the difference in transistor currents on the capacitor $C_u$ to update the particle count. Notice that neither barrier
depends on the capacitor voltage \( u_V \). Thus, \( u_V \) becomes representative of the fraction of open channels; it increases as particles switch to the open state.

How do we compute the fluxes from the transition rates? If \( u_V \) directly represented the particle count, we would take the product of \( u_V \) and the transition rates. However, we can avoid multiplying altogether if \( u_V \) represents the logarithm of the open fraction rather than the open fraction itself. \( u_V \)'s dynamics are described by the differential equation

\[
C_u \frac{du_V}{dt} = I_{ds0} e^\kappa V_O \left( e^{-u_V} - e^{-u_H} \right) - I_{ds0} e^\kappa V_C e^{-u_L},
\]

\[
= I_{ds0} e^\kappa V_O - u_H \left( e^{-(u_V-u_H)} - 1 \right) - I_{ds0} e^\kappa V_C - u_L,
\]

(3.1)

where \( I_{ds0} \) and \( \kappa \) are transistor parameters (defined in equation 2.10), and \( V_O, V_C, u_H, u_L, \) and \( u_V \) are voltages (defined in Figure 4a). We assume N1 remains in saturation during the channel’s operation; that is, \( u_V > u_L + 4U_T \), making the drain voltage’s influence negligible.

The analogies between equation 3.1 and equation 2.2 become clear when we divide the latter by \( u \). Our barriers—N1’s source-gate for the closing rate and N2’s drain-gate for the opening rate—correspond to \( \alpha (V) \) and \( \beta (V) \), while \( e^{-(u_V-u_H)} \) corresponds to \( u^{-1} \). Thus, our circuit computes (and integrates) the net flux divided by \( u \), the open fraction. Fortuitously, the net flux scaled by the open fraction is exactly what we need to update the fraction’s logarithm, since \( d \log(u)/dt = (du/dt)/u \). Indeed, substituting \( u_V = \log(u + u_H) \)—our log-domain representation of \( u \)—into equation 3.1 yields

\[
\frac{Q_u}{u} \frac{du}{dt} = I_{ds0} e^\kappa V_O - u_H \left( \frac{1}{u} - 1 \right) - I_{ds0} e^\kappa V_C - u_L,
\]

\[
\frac{du}{dt} = \frac{I_{ds0}}{Q_u} e^\kappa V_O - u_H \left( 1 - u \right) - \frac{I_{ds0}}{Q_u} e^\kappa V_C - u_L u,
\]

(3.2)

where \( Q_u = C_u U_T \). If we design \( V_C \) and \( V_O \) to be functions of the membrane voltage \( V \), equation 3.2 becomes directly analogous to equation 2.2.

In linear thermodynamic models, the transition rates depend exponentially on the membrane voltage. We can realize this by designing \( V_C \) and \( V_O \) to be linear functions of \( V \), albeit with slopes of opposite sign. The opposite slopes ensure that as the membrane voltage shifts in one direction, the opening and closing rates will change in opposite directions relative to each other.

Thus far in our circuit design, we have not specified whether the variable activates or inactivates as the membrane voltage increases. Recall that for activation, the gating particle opens as the cell depolarizes, whereas for
inactivation, the gating particle opens as the cell hyperpolarizes. In our circuit, whether activation or inactivation occurs depends on how we define $V_O$ and $V_C$ with respect to $V$. Increasing $V_O$, and decreasing $V_C$, with $V$ defines an activation variable, as at depolarized levels this results in $\alpha (V) > \beta (V)$. Conversely, increasing $V_C$, and decreasing $V_O$, with $V$ defines an inactivation variable, as now at depolarized voltages, $\beta (V) > \alpha (V)$, and the variable will equilibrate in a closed state.

Naturally, our circuit has the same limitation that all two-state linear thermodynamic models have: its time constant approaches zero when either $V_O$ or $V_C$ grows large, as the transition rates $\alpha (V)$ and $\beta (V)$ become unrealistically large. This shortcoming is easily rectified by imposing an upper limit on the transition rates, as has been done for other thermodynamic models (Willms et al., 1999). We realize this saturation by placing two additional transistors in series with the original two (see Figure 4b). With these transistors, the transition rates become

$$\alpha (V) = \frac{I_{ds0}}{Q_u} \frac{e^{e^{V_O - u_H}}}{1 + e^{e^{(V_O - u_H)}}} \quad (3.3)$$

$$\beta (V) = \frac{I_{ds0}}{Q_u} \frac{e^{e^{V_C - u_L}}}{1 + e^{e^{(V_C - u_L)}}} \quad (3.4)$$

where the single exponentials in equation 3.2 are now scaled by an additional exponential term. The voltages $u_{rH}$ and $u_{rL}$ (fixed biases) set the maximum transition rate for opening and closing, respectively. That is, when $V_O < u_{rH} - 4 U_T$, $\alpha (V) \propto e^{e^{V_O}}$, a rate function exponentially dependent on the membrane voltage. But when $V_O > u_{rH} + 4 U_T$, $\alpha (V) \propto e^{e^{u_{rH}}}$, limiting the transition rate and fixing the minimum time constant for channel opening. The behavior of $\beta (V)$ is similarly defined by $V_C$’s value relative to $u_{rL}$.

In the following section, we explore how the channel variable computed by our circuit changes with the membrane voltage and how quickly it approaches steady state. To do so, we must relate the steady state and time constant to the opening and closing rates and specify how the circuit’s opening and closing voltages depend on the membrane voltage.

### 4 Circuit Operation

To help understand the operation of the channel circuit and the influence of various circuit parameters, we will derive $u_\infty (V)$ and $\tau_u (V)$ for the circuit in Figure 4b using equations 2.4 and 2.5 and the transistor rate equations (equations 3.3 and 3.4), limiting our presentation to the activation version. The derivation, and the influence of various parameters, is similar for the inactivation version.
For the activation version of the channel circuit (see Figure 4b), we define the opening and closing voltages’ dependence on the membrane voltage as:

\[ V_O = \phi_o + \gamma_o \, V \]  
\[ V_C = \phi_c - \gamma_c \, V, \]  

(4.1)\hspace{1cm}(4.2)

where \( \phi_o, \gamma_o, \phi_c, \) and \( \gamma_c \) are positive constants representing the offsets and slopes for the opening and closing voltages. Additional circuitry is required to define these constants; one example is described in the next section. In this section, however, we will leave the definition as such while we derive the equations for the circuit.

Under certain restrictions (see appendix A), \( u \)'s steady-state level has a sigmoidal voltage dependence:

\[ u_\infty (V) = \frac{1}{1 + \exp \left[ -\frac{V - V_{\text{mid}}}{V_u^*} \right]}. \]  

(4.3)

where

\[ V_{\text{mid}}^u = \frac{1}{\gamma_o + \gamma_c} \left( \phi_c - \phi_o + (u_H - u_L)/\kappa \right) \]  
\[ V_u^* = \frac{1}{\gamma_o + \gamma_c} \frac{U_T}{\kappa}. \]  

(4.4)\hspace{1.5cm}(4.5)

Figure 5a shows how the sigmoid arises from the transition rates and, through them, its relationship to the voltage biases. The midpoint of the sigmoid, where the open probability equals half, occurs when \( \alpha (V) = \beta (V) \); it will thus shift with any voltage biases (\( \phi_o, \phi_c, u_H, \) or \( u_L \)) that scale either of the transition rate currents. For example, increasing \( u_H \) reduces \( \alpha (V) \), shifting the midpoint to higher voltages. The slope of the sigmoid around the midpoint is defined by the slopes \( \gamma_o \) and \( \gamma_c \) of \( V_O (V) \) and \( V_C (V) \).

To obtain the sigmoid shape, we restricted the effect of saturation. It is assumed that the bias voltages \( u_{tH} \) and \( u_{tL} \) are set such that saturation is negligible in the linear midsegment of the sigmoid (i.e., \( V_O < u_{tH} - 4 \, U_T \) and \( V_C < u_{tL} - 4 \, U_T \) when \( V \sim V_{u}^{\text{mid}} \)). That is why \( \alpha (V) \) and \( \beta (V) \) appear to be pure exponentials in Figure 5a. This restriction is reasonable as saturation is supposed to occur only for large excursions from \( V_{u}^{\text{mid}} \), where it imposes a lower limit on the time constant. Therefore, under this assumption, the sigmoid lacks any dependence on the biases \( u_{tH} \) and \( u_{tL} \).
Figure 5: Steady state and time constants for channel circuit. (a) The variable’s steady-state value ($u_\infty$) changes sigmoidally with membrane voltage ($V$), dictated by the ratio of the opening and closing rates (dashed lines). The midpoint occurs when the rates are equal, and hence its horizontal location is affected by the bias voltages ($u_H$ and $u_L$) applied to the circuit (see Figure 4b). (b) The variable’s time constant ($\tau_u$) has a bell-shaped dependence on the membrane voltage ($V$), dictated by the reciprocal of the opening and closing rates (dashed lines). The time constant diverges from these asymptotes at intermediate voltages, where neither rate dominates; it follows the reciprocal of their sum, peaking when the sum is minimized.

Under certain further restrictions (see appendix A), $u$’s time constant has a bell-shaped voltage dependence:

$$\tau_u (V) = \tau_{\min} \left( 1 + \frac{1}{\exp \left( \frac{V - V_{1u}}{V_{1u}^{*}} \right) + \exp \left( -\frac{V - V_{2u}}{V_{2u}^{*}} \right) } \right)$$

(4.6)

where

$$V_{1u} = ( u_H - \phi_o ) / \gamma_o$$

$$V_{1u}^{*} = U_T / ( \kappa \gamma_o )$$

$$V_{2u} = ( \phi_c - u_H + (u_H - u_L) / \kappa ) / \gamma_c$$

$$V_{2u}^{*} = U_T / ( \kappa \gamma_c )$$

and $\tau_{\min} = (Q_u / I_{ds0}) \ e^{-(\kappa (u_H - u_L))}$. Figure 5b shows how the bell shape arises from the transition rates and, through them, its relationship to the voltage biases. For large excursions of the membrane voltage, one transition rate dominates, and the time constant closely follows its inverse. For small excursions, neither rate dominates, and the time constant diverges from the inverses, peaking at the membrane voltage where the sum of the transition rates is minimized.
To obtain the bell shape, we saturated the opening and closing rates at the same level by setting \( \kappa \ u_H - u_H = \kappa \ u_L - u_L \). Though not strictly necessary, this assumption simplifies the expression for \( \tau_u (V) \) by matching the minimum time constants at hyperpolarized and depolarized voltages, yielding the result given in equation 4.6. The bell shape also requires this so-called minimum time constant to be smaller than the peak time constant in the absence of saturation.

The free parameters within the circuit—\( \phi_o, \gamma_o, \phi_c, \phi_{o,H}, \phi_{o,L}, \ u_H, \ u_L, \ u_{t,H}, \) and \( u_{t,L} \)—allow for much flexibility in designing a channel. Appendix B provides an expanded discussion on the influence of the various parameters in the equations above. In the following section, we present measurements from a simple activation channel designed using this circuit, which was fabricated in a standard 0.25 \( \mu \)m CMOS process.

5 A Simple Activation Channel

Our goal here is to implement an activating channel to serve as a concrete example and examine its behavior through experiment. We start with the channel variable circuit, which computes the logarithm of the channel variable, and attach its output voltage to the gate of a transistor, which uses the subthreshold regime’s exponential current-voltage relationship to invert the logarithm. The current this transistor produces, which is directly proportional to the variable, can be injected directly into a silicon neuron (Hynna & Boahen, 2001) or can be used to define a conductance (Simoni et al., 2004). The actual choice is irrelevant for the purposes of this article, which demonstrates only the channel variable.

In addition to the output transistor, we also need circuitry to compute the opening and closing voltages from the membrane voltage. For the opening voltage \( (V_O) \), we simply use a wire to tie it to the membrane voltage \( (V) \), which yields a slope of unity \( (\gamma_o = 1) \) and an intercept of zero \( (\phi_o = 0) \). For the closing voltage \( (V_C) \), we use four transistors to invert the membrane voltage. The end result is shown in Figure 6. For the voltage inverter circuit we chose, the closing voltage’s intercept \( \phi_c = \kappa \ V_{C0} \) (set by a bias voltage \( V_{C0} \)) and its slope \( \gamma_c = \kappa^2 / (\kappa + 1) \) (set by the transistor parameter defined in equation 2.10). Since \( \kappa \approx 0.7 \), the closing voltage has a shallower slope than the opening voltage, which makes the closing rate change more gradually, skewing the bell curve in the hyperpolarizing direction as intended for the application in which this circuit was used (Hynna, 2005).

This eight-transistor design captures the ion channel’s nonlinear dynamics, which we demonstrated by performing voltage clamp experiments (see Figure 7). As the command voltage (i.e., step size) increases, the output current’s time course and final amplitude both change. The clustering and speed at low and high voltages are what we would expect from a sigmoidal steady-state dependence with a bell-shaped time constant. The
Figure 6: A simple activating channel. A voltage inverter (N5-8) produces the closing voltage ($V_C$); a channel variable circuit (N1-3) implements the variable’s dynamics in the log domain ($u_V$); and an antilog transistor (N4) produces a current ($I_T$) proportional to the variable. The opening voltage ($V_O$) is identical to the membrane voltage ($V$). The series transistor (N2) sets the minimum time constant at depolarized levels. The same circuit can be used to implement an inactivating channel simply by swapping $V_O$ and $V_C$.

Figure 7: Channel circuit’s measured voltage-dependent activation. When the membrane voltage is stepped to increasing levels, from the same starting level, the output current becomes increasingly larger, approaching its steady-state amplitudes at varying speeds.

The relationship between this output current ($I_T$) and the activation variable, defined as $u = e^{u_V - u_H}$, has the form

$$I_T = u^\kappa I_T.$$  \hspace{1cm} (5.1)
Its maximum value $\bar{I}_T = e^{\kappa u_H - u_G}$ and its exponent $\kappa \approx 0.7$ (the same transistor parameter). It is possible to achieve an exponent of unity, or even a square or a cube, but this requires a lot more transistors.

We measured the sigmoidal change in activation directly, by sweeping the membrane voltage slowly, and its bell-shaped time constant indirectly, by fitting the voltage clamp data in Figure 7 with exponentials. The results are shown in Figure 8; the relationship above was used to obtain $u$ from $I_T$. The solid lines are the fits of equations 4.3 and 4.6, which reasonably capture the behavior in both sets of data. The range in the time constant data is limited due to the experimental protocol used. Since we modulated only the step size from a fixed hyperpolarized position, we need a measurable change in the steady-state output current to be able to measure the temporal dynamics for opening. However, this worked to our benefit as, given the range of the fit, there was no need to modify equation 4.6 to allow the time constant to go to zero at hyperpolarized levels (this circuit omits the second saturation transistor in Figure 4b).

All of the extracted parameters from the fit are reasonably close to our expectations—based on equations 4.3 and 4.6 and our applied voltage biases—except for $V_{2u}^*$. For $\kappa \approx 0.7$ and $U_T = 25.4$ mV, we expected $V_{2u}^* \approx 125$ mV, but our fit yielded $V_{2u}^* \approx 71.8$ mV. There are two possible explanations, not mutually exclusive. First, the fact that equation 4.6 assumes the presence of a saturation transistor, in addition to the limited data along the left side of the bell curve, may have contributed to the underfitting of that value. Second, $\kappa$ is not constant within the chip but possesses voltage dependence. Overall, however, the analysis matches reasonably well the performance of the circuit.
6 Conclusion

We showed that the transistor is a thermodynamic isomorph of a channel gating particle. The analogy is accomplished by considering the operation of both within the framework of energy models. Both involve the movement of charge within an electric field: for the channel, due to conformations of the ion channel protein; for the transistor, due to charge carriers entering the transistor channel. Using this analogy, we generated a compact channel variable circuit.

We demonstrated our variable circuit’s operation by implementing a simple channel with a single activation variable, showing that the steady state is sigmoid and the time constant bell shaped. Our measured results, obtained through voltage clamp experiments, matched our analytical results, derived from knowledge of transistors. Bias voltages applied to the circuit allow us to shift the sigmoid and the bell curve and set the bell curve’s height independently. However, the sigmoid’s slope, and its location relative to the bell curve, which is determined by the slope, cannot be changed (it is set by the MOS transistor’s \( \kappa \) parameter).

Our variable circuit is not limited to activation variables: reversing the opening and closing voltages’ linear dependence on the membrane voltage will change the circuit into an inactivation variable. In addition, channels that activate and inactivate are easily modeled by including additional circuitry to multiply the variable circuits’ output currents (Simoni et al., 2004; Delbruck, 1991).

The change in temporal dynamics of gating particles plays a critical role in some voltage-gated ion channels. As discussed in section 1, the inactivation time constant of T channels in thalamic relays changes dramatically, defining properties of the relay cell burst response, such as the interburst interval and the length of the burst itself. Activation time constants are also influential: they can modify the delay with which the channel responds (Zhan, Cox, Rinzel, & Sherman, 1999), an important determinant of a neuron’s temporal precision. Incorporating these nonlinear temporal dynamics into silicon models will yield further insights into neural computation.

Equally important in our design, not only were we able to capture the nonlinear dynamics of gating particles, we were able to do so using fewer transistors than previous silicon models. Rather than exploit the parallels between transistors and ion channels, as we did, previous silicon modelers attempted to “linearize” the transistor, to make it approximate a resistor. After designing a circuit to accomplish this, the resistor’s value had to be adjusted dynamically, so more circuitry was added to filter the membrane voltage. The time constant of this filter was kept constant, sacrificing the ion channel’s voltage-dependent nonlinear dynamics for simplicity. We avoided all these complications by recognizing that the transistor is a
thermodynamic isomorph of the ion channel. Thus, we were able to come up with a compact replica.

The size of the circuit is an important consideration within silicon models, as smaller circuits translate to more neurons on a silicon die. To illustrate, a simple silicon neuron (Zaghloul & Boahen, 2004), with a single input synapse, possessing the activation channel from section 5, requires about 330 µm² of area. This corresponds to around 30,000 neurons on a silicon die, 10 mm² in area.

Adding additional circuits, such as inactivation to the channel, increases the area of the cell design, reducing the size of the population on the chip (assuming, of course, that the total area of the die remains constant). To compensate for larger cell footprints, we can either increase the size of the whole silicon die (which costs money), or we can simply incorporate multiple chips into the system, easily doubling or tripling the network size. And unlike computer simulations, the increase in network sizes comes with minimal cost in performance or “simulation” time.

Of course, like all other modeling media, silicon has its own drawbacks. For one, silicon is not forgiving with respect to design flaws. Once the chip has been fabricated, we are limited to manipulating our models using only external voltage biases within our design. This places a great deal of importance on verification of the final design before submitting it for fabrication; the total time from starting the design to receiving the fabricated chip can be on the order of 6 to 12 months.

An additional characteristic of the silicon fabrication process is mismatch, a term referring to the variability among fabricated copies of the same design within a silicon chip (Pavasovic, Andreou, & Westgate, 1994). Within an array of silicon neurons, this translates into heterogeneity within the population. While we can take steps to reduce the variability within an array, generally at the expense of area, this mismatch can be considered a feature, since biology also needs to deal with variability. When we build silicon models that reproduce biological phenomena, being able to do so lends credence to our models, given their robustness to parameter variability. And when we discover network phenomena within our chips, these are likely to be found in biology as well, as they will be robust to biological heterogeneity.

With the ion channel design described in this article and our ability to expand our networks without much cost, we have a great deal of potential in building many of the neural systems within the brain, which consists of numerous layers of cells, each possessing its own distinct characteristics. Not only do we have the opportunity to study the role of an ion channel within an individual cell, we have the potential to study its influence within the dynamics of a population of cells, and hence its role in neural computation.
Appendix A: Derivations

We can use the transition rates for our channel circuit (see equations 3.3 and 3.4) to calculate the voltage dependence of its steady state and time constant. Starting with the steady-state equation, equation 2.4,

\[ u_\infty (V) = \frac{\alpha (V)}{\alpha (V) + \beta (V)} \]

\[ = \frac{I_{ds0}}{Q_u} \frac{e^{-u_H}}{e^{-u_H} + e^{-u_{H1}}} + \frac{I_{ds0}}{Q_u} \frac{e^{-u_L}}{e^{-u_L} + e^{-u_{L1}}} \]

\[ = \frac{1}{1 + \frac{e^{-u_H}}{e^{-u_L} + e^{-u_{L1}}} e^{u_H - u_L}}. \]

Throughout the linear segment of the sigmoid, we set the voltage biases such that \( u_{H1} > V_O + 4U_T \) and \( u_{L1} > V_C + 4U_T \). These restrictions essentially marginalize \( u_{H1} \) and \( u_{L1} \). By the time either of the terms with these two biases becomes significant, the steady state will be sufficiently close to either unity or zero, so that their influence is negligible. Therefore, we drop the exponential terms with \( u_{H1} \) and \( u_{L1} \); substituting equations 4.1 and 4.2 yields the desired result of equation 4.3.

For the time constant, we substitute equations 3.3 and 3.4 into equation 4.1:

\[ \tau_u (V) = \frac{1}{\alpha (V) + \beta (V)} \]

\[ = \frac{Q_u}{I_{ds0}} \frac{1}{e^{-V_O - u_H} + e^{-V_C - u_{L1}}} \]

\[ = \frac{1}{e^{-V_O - u_H} + e^{-V_C - u_{L1}}} + \frac{1}{e^{-V_C - u_L} + e^{-V_H - u_{H1}}}. \]

To equalize the minimum time constant at hyperpolarized and depolarized levels, we establish the following relationship: \( \kappa u_{H1} - u_H = \kappa u_{L1} - u_L \). After additional algebraic manipulation, the time constant becomes

\[ \tau_u (V) = \tau_{\text{min}} \left( 1 + \frac{e^{-\kappa (V_O - u_H)} e^{-\kappa (V_C - u_{L1})}}{e^{-\kappa (V_O - u_{H1})} + e^{-\kappa (V_C - u_{L1})} + 2} \right) \]

\[ = \frac{1}{e^{-\kappa (V_O - u_{H1})} + e^{-\kappa (V_C - u_{L1})} + 2}. \]
where $\tau_{\text{min}} = (Q_u/I_{ds0}) e^{-(\kappa u_H - u_H)}$ is the minimum time constant. To reduce the expression further, we need to understand the relative magnitudes of the various terms.

We can drop the constant in both denominators, as one of the exponentials there will always be significantly larger. Since $V_O$ and $V_C$ have opposite signs for their slopes with respect to $V$, the sum of the exponentials in the denominator peaks at a membrane voltage somewhere within the middle section of its operational voltage range. As it happens, this peak is close to the midpoint of the sigmoid (see below), where we have defined $u_{rH} > V_O + 4U_T$ and $u_{rL} > V_C + 4U_T$. Thus, at the peak, we know the constant is negligible. As the membrane voltage moves away from the peak, in either direction, one of the exponentials will continue to increase while the other decreases. Thus, with the restriction on the bias voltage $u_{rH}$ and $u_{rL}$, the sum of the exponentials will always be much larger than the constant in the denominator.

By the same logic, we can disregard the final term, since the sum of the exponentials will always be substantially larger than the numerator, making the fraction negligible over the whole membrane voltage.

With these assumptions, and substituting the dependence of $V_O$ and $V_C$ on $V$ (see Equations 4.1 and 4.2), we obtain the desired result, equation 4.6.

Appendix B: Circuit Flexibility

This section is more theoretical in nature, using the steady-state and time-constant equations derived in appendix A to provide insight into how the various parameters influence the two voltage dependencies (steady state and time constant). This section is likely of interest only to those who wish to use our approach for modeling ion channels.

An important consideration in generating these models is defining the location (i.e., the membrane voltage) and magnitude of the peak time constant. Unlike the minimum time constant, which is determined simply by the difference between the bias voltages $u_{rH}$ and $u_H$ (or $u_{rL}$ and $u_L$), no bias directly controls the maximum time constant, since it is the point at which the sum of the transition rates is minimized (see Figure 5b). The voltage at which it occurs, however, is easily determined from equation 4.6:

$$V_{\tau pk} = \frac{U_T}{\kappa (\gamma_o + \gamma_c)} \log \left( \frac{\gamma_c}{\gamma_o} \right).$$  \hspace{1cm} (B.1)

Thus, where the bell curve lies relative to the sigmoid, whose midpoint lies at $V_u^{\text{mid}}$, is determined by the opening and closing voltages’ slopes ($\gamma_o$ and $\gamma_c$). Consequently, changing these slopes is the only way to displace the bell curve relative to the sigmoid. Shifting the bell curve by changing
Figure 9: Varying the closing voltage’s slope ($\gamma_c$). (a) Changing $\gamma_c$ adjusts both the slope and midpoint of the steady-state sigmoid. (b) $\gamma_c$ also affects the location and height (relative to the minimum) of the time constant’s bell curve; the change in height (plotted logarithmically) is dramatic. (c, d) Same as in a and b, except that we adjusted the bias voltage $u_L$ to compensate for the change in $\gamma_c$, so the sigmoid’s midpoint and the bell curve’s height remain the same. The sigmoid’s slope does change, as it did before, and the bell curve’s location shifts as well, though much less than before. In these plots, $\phi_o = 400 \text{ mV}$, $u_H = 400 \text{ mV}$, $u_L = 50 \text{ mV}$, $u_{H+} = 700 \text{ mV}$, and $\gamma_o = 1.0$. $\gamma_c$’s values are 0.5 (thin, solid line), 0.75 (short, dashed line), 1.0 (long, dashed line), and 1.25 (thick, solid line).

As we change the opening and closing voltages’ slopes ($\gamma_o$ and $\gamma_c$), two things happen. First, due to $V_{mid}^u$’s dependence on these parameters (see equation 4.4), the sigmoid and the bell curve shift together. Two, due to the dependence we just described (see equation B.1), the bell curve shifts relative to the sigmoid. These effects are illustrated in Figures 9a and 9b for $\gamma_c$ ($\gamma_o$ behaves similarly). To eliminate the first effect while preserving the second, we can compensate for the sigmoid’s shift by adjusting the bias voltage $u_L$, which scales the closing rate (see equation 4.2). Consequently, $u_L$ also rescales the left part of the bell curve, where the closing rate is dominant, reducing its height relative to the minimum and canceling its shift due to the first effect.

This is demonstrated in Figures 9c and 9d. The sigmoid remains fixed, while the bell curve shifts (slightly) due to the second effect. Although the bias voltage $u_L$ cancels the shift $\gamma_c$ produces in the sigmoid, it does not compensate for the change in the sigmoid’s slope. We can shift the sigmoid
Figure 10: Varying the closing voltage’s intercept ($\phi_c$). (a) Changing $\phi_c$ shifts the steady-state sigmoid’s midpoint, leaving its slope unaffected. (b) $\phi_c$ also affects the time constant bell curve’s location and height (relative to the minimum). In these plots, $\phi_o = 0 \text{ mV}$, $u_H = 400 \text{ mV}$, $u_L = 50 \text{ mV}$, $u_{rH} = 700 \text{ mV}$, $\gamma_o = 1.0$, and $\gamma_c = 0.5$. $\phi_c$’s values are 400 mV (thin, solid line), 425 mV (short, dashed line), 450 mV (long, dashed line), and 475 mV (thick, solid line).

Figure 11: Setting the bell curve’s location and height independently. (a) Changing the opening and closing voltages’ intercepts ($\phi_o$ and $\phi_c$) together shifts the bell curve’s location without affecting its height. The steady-state sigmoid (not shown) moves with the bell curve (see equation B.1). (b) Changing $\phi_o$ and $\phi_c$ in opposite ways increases the height (relative to the minimum) without affecting the location. The steady-state sigmoid (not shown) stays put as well. In these plots, $u_H = 400 \text{ mV}$, $u_L = 50 \text{ mV}$, $u_{rH} = 700 \text{ mV}$, $\gamma_o = 1.0$, and $\gamma_c = 0.5$. In both $a$ and $b$, $\phi_c = 400 \text{ mV}$ and $\phi_o = 0 \text{ mV}$ for the thin, solid line. In $a$, both $\phi_c$ and $\phi_o$ increment by 25 mV from the thin, solid line to the short, dashed line, to the long, dashed line and to the thick, solid line. In $b$, $\phi_c$ increments and $\phi_o$ decrements by 25 mV from the thin, solid line to the short, dashed line, to the long, dashed line, and to the thick, solid line.

and leave its slope unaffected by changing the opening and closing voltages’ intercepts, as shown in Figure 10 for $\phi_c$ ($\phi_o$ behaves similarly). However, the bell curve shifts by the same amount, and its height changes as well, since $\phi_c$ rescales the closing rate. Thus, it is not possible to shift the bell curve relative to the sigmoid without changing the latter’s slope; this is evident from equations 4.5 and B.1.
We can set the bell curve’s location and height independently if we change the opening and closing voltages’ intercepts by the same amount or by equal and opposite amounts, respectively. Equal changes in the intercepts ($\phi_o$ and $\phi_c$) shift the opening and closing rate curves by the same amount, thus shifting the bell curve (and the sigmoid) without affecting its height (see Figure 11a), whereas equal and opposite changes shift the opening and closing rate curves apart, leaving the point where they cross at the same location while rescaling the value of the rate there. As a result, the bell curve’s height is changed without affecting its location (see Figure 11b).

Choosing values for $\gamma_o$ and $\gamma_c$, however, presents a trade-off. These two parameters define the dependence of $V_o$ and $V_c$ on $V$ and are not external biases like the other parameters; rather, they are defined by the fabrication process through the transistor parameter $\kappa$. We can define their relationships with $\kappa$ through the use of different circuits; in our simple activation channel (see section 5), $\gamma_c = \kappa^2/(\kappa + 1)$, as defined by the four transistors that invert the membrane voltage.

There are a couple of drawbacks. First, not all values for $\gamma_o$ or $\gamma_c$ are possible using only a few transistors. Expressed another way, a trade-off needs to be made between achieving more precise values for $\gamma_o$ or $\gamma_c$ and using fewer transistors within the design. The other drawback is that after fabrication, $\gamma_o$ and $\gamma_c$ can no longer be modified, as they are defined as functions of the transistor parameter $\kappa$. These issues merit special consideration before submitting the chip for fabrication.

References


Received October 6, 2005; accepted June 20, 2006.