

1984

How Hamiltonian Dynamical Theory In The Complex Domain Yields Asymptotic Solutions to the Non-Hermitian Integral Equations of Visual Nerve-Networks

Bruce Knight

Follow this and additional works at: http://digitalcommons.rockefeller.edu/knight_laboratory

 Part of the [Life Sciences Commons](#)

Recommended Citation

Knight, B. W., Jr. (1984) How Hamiltonian dynamical theory in the complex domain yields asymptotic solutions to the non-Hermitian integral equations of visual nerve-networks. In *Mathematical Physics VII*. Eds: Brittin, W.E.; Gustafson, K.; Wyss, W. p. 431-453.

This Article is brought to you for free and open access by the Laboratories and Research at Digital Commons @ RU. It has been accepted for inclusion in Knight Laboratory by an authorized administrator of Digital Commons @ RU. For more information, please contact mcsweej@mail.rockefeller.edu.

Reprinted with permission from:

Mathematical Physics VII: Proceedings of the VIIth International Congress on
Mathematical Physics, Boulder, Colorado, USA, August 1-10, 1983.

Edited by W. E. Brittin...et al.

Amsterdam : North-Holland Physics Publishing, 1984.

pp. 431-454.

HOW HAMILTONIAN DYNAMICAL THEORY IN THE COMPLEX DOMAIN YIELDS ASYMPTOTIC SOLUTIONS TO THE NON-HERMITIAN INTEGRAL EQUATIONS OF VISUAL NERVE-NETWORKS

Bruce W. KNIGHT

Laboratory of Biophysics, Rockefeller University, New York, NY 10021

1. INTRODUCTION

The development of useful theoretical tools is the distinctive province of mathematical physics, and in that spirit my remarks here will concern some new tools that furnish a convenient solution to a generally important practical problem which arises in my own work. My laboratory studies the dynamics of real nerve networks^{1,2,3,4,5,6,7}. That study encounters the classical eigenvalue problem

$$K\psi = \lambda\psi \tag{1.1}$$

where the linear operator K is well defined at the outset, while the eigenvector ψ and its eigenvalue λ are to be found. Our own problem in more detail^{8,9} is

$$\int dy K\{\underline{x}, \underline{y}\} \psi(\underline{y}) = \lambda\psi(\underline{x}) \tag{1.2}$$

where the integral operator $K\{\underline{x}, \underline{y}\}$ acts on functions which live on a fully infinite two-space with points \underline{x} . In our case of greatest practical interest, when \underline{x} and \underline{y} are fixed then K is a complex number, and appears in (1.2) without major traditional simplifying features (such as Hermitian or unitary structure), though it is analytic in the components of \underline{x} and \underline{y} . Our problem offers one feature for technical exploitation:^{8,9,10,11} If we re-express K in terms of difference and average arguments

$$K\{\underline{x}, \underline{y}\} = K(\underline{x}-\underline{y}, \frac{\underline{x}+\underline{y}}{2}) \tag{1.3}$$

(with no loss of generality) then the K which our problem gives us is (technically) slow in the second argument $(\underline{x}+\underline{y})/2$. This observation leads to an asymptotic method of solution which is so robust that "slowness" may be regarded almost as a catalyst rather than a limitation.^{8,9,10} In particular, valuable qualitative features of the spectrum and eigenfunctions of K emerge, and if the spectrum has more than a few discrete points then our results are asymptotic for "large enough n" with no demand for "slowness" in K itself.

Presented at the VIIth INTERNATIONAL CONGRESS ON MATHEMATICAL PHYSICS, Boulder, Colorado, 1983.

Below we will see how the kernel $K\{x,y\}$ is related by a linear transformation to another function $\tilde{K}(p,q)$ which leads to asymptotic solution¹¹ of the eigenvalue problem (1.2) through solution of the Hamiltonian equations

$$\begin{aligned}\dot{q}_m &= \partial\tilde{K}/\partial p_m \\ \dot{p}_m &= -\partial\tilde{K}/\partial q_m .\end{aligned}\tag{1.4}$$

The spectrum of K in (1.2) has qualitative features which we will be able to relate to singular points of the ordinary differential equations (1.4), and the eigenvalues may be developed asymptotically by recourse to the machinery of Hamiltonian mechanics. If K is Hermitian then \tilde{K} is real, and some interesting relations between eigenvalue problems and Hamiltonian dynamics emerge. If K is not Hermitian^{12,13} then the singular points of interest may lie at complex p and q , where the analytic Hamiltonian procedures may still be justified, and this sheds a somewhat different light on both Hamiltonian theory and the eigenvalue problem.

It is my utmost hope that inventive mathematical physicists may exploit my remarks here in ways I cannot guess. In particular, the eigenvalue problem which is "easy" in some respects might become useful in exploring hard parts of Hamiltonian theory.

My intent below is to help insight, not to establish proof. I will sketch a line of reasoning which stands by itself, but the general familiarity of mathematical physicists with classical and quantum mechanics will be used freely to underscore key ideas.

2. BACKGROUND

A Laboratory of Biophysics is where I work. We study the dynamics of real neural networks, particularly those of the eye and the visual part of the brain because visual stimulation furnishes us a firm handhold on our experimental material. In our work reported here there is no clear border between theory and experiment, but I would roughly classify my closest colleagues as follows. In early experimental work: Frederick Dodge, Keffer Hartline, Floyd Ratliff, Jun Toyoda. In more recent experimental work: Scott Brodie, Stevan Dawis, Floyd Ratliff, Robert Shapley, Lawrence Sirovich. In recent theory Lawrence Sirovich has been my close collaborator.

This work was sheltered and encouraged in what was traditionally Keffer Hartline's laboratory; his recent death at 79 we feel keenly.

By the organizers' mandate, I will now give a lightning-like review of the biophysical background of the present theoretical problem. Figure 1 is a microscope picture of a neural network. This is a bit of the retina of the

HAMILTONIAN DYNAMICAL THEORY

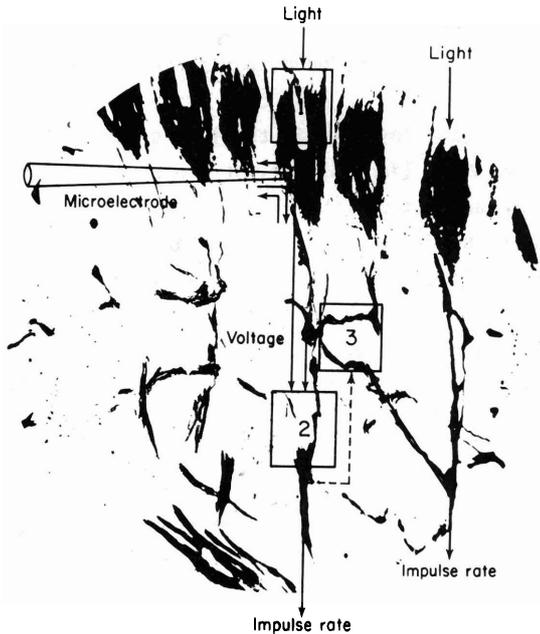


Figure 1. A bit of the neural network in the horseshoe crab's retina, with schematic of transducers and information pathways superimposed. Micrograph by William H. Miller.

horseshoe crab, which retina plays a role in real-network dynamics similar to that of Kepler motion in classical mechanics, or the hydrogen atom in quantum theory; in a well-defined sense it is an exactly solvable system.

The black object marked "1" is a visual cell which contains, within an insulating membrane, biophysical machinery of molecular size which produces voltage in response to light. From the visual cell proceeds a nerve fiber, a narrow tube of conducting fluid bounded by insulating membrane. The region marked "2" generates relaxation-oscillations, nerve impulses which proceed (downward in the picture) to the brain. The rate of impulse generation at "2" is modulated by the level of voltage generated at "1". The function of the box marked "3" is to convert impulse rate back into voltage which, as shown, is

done both for the cell's own signal (a feedback loop) and for signals arriving from nearby cells. We are looking at a typical neural information-processing network which performs a quite profound transformation on the information it receives.

Through a microelectrode inserted in the living visual cell we may measure the way voltage responds to light at "1", or conversely we may mandate a voltage at the microelectrode and observe the response in impulse rate at "2"; or we may vary the light on the nearby cell and measure through the microelectrode the back-conversion of impulse rate to voltage at "3". We may also drive "3" by ourselves sending a train of impulses backwards up the optic nerve and measuring the consequent voltage transduction at "3". **By doing these manipulations** in combination we may obtain a very detailed understanding of the dynamics of this network.

Figure 2 shows some laboratory data [from (3)]. The signal input in each case was sinusoidal in time, and in each case the output is sinusoidal as well. Here we are characterizing the performance of the signal transducers of Figure 1. Each transducer is characterized by the amplitude and phase of the sinusoid with which it responds to an input sinusoid of standard amplitude and phase.

Returning to Figure 1, if the amplitude-and-phase characterization is appropriate, we should be able to predict how the optic nerve fiber responds to sinusoidal light: simply multiply the measured amplitudes for transducers "1" and "2", and algebraically add their measured phases. Figure 3 [from (3)] justifies this observation. **Amplitude and phase are plotted versus frequency.** On the left the filled circles show amplitude and phase data for transducer "1" and open circles for transducer "2". Multiplication of amplitudes, and addition of phases predicts the solid curves on the right, which are in nice agreement with the measured data points.

We note the combined transduction of Figure 3 is by no means an "identity" transduction: higher frequencies are more emphasized than low ones, and this corresponds to the creature's need to detect sudden changes in visual environment.

A second type of experiment (14) shows that when the network responds to simultaneous inputs, it responds simply with the algebraic sum of what the inputs would have yielded individually.

The above information in fact tells us how to mathematically combine the component transducer responses of the neural network of Figure 1 (or any network of generally similar organization) into rules that predict the dynamics of the total system. The prescription is simply the natural formalization of

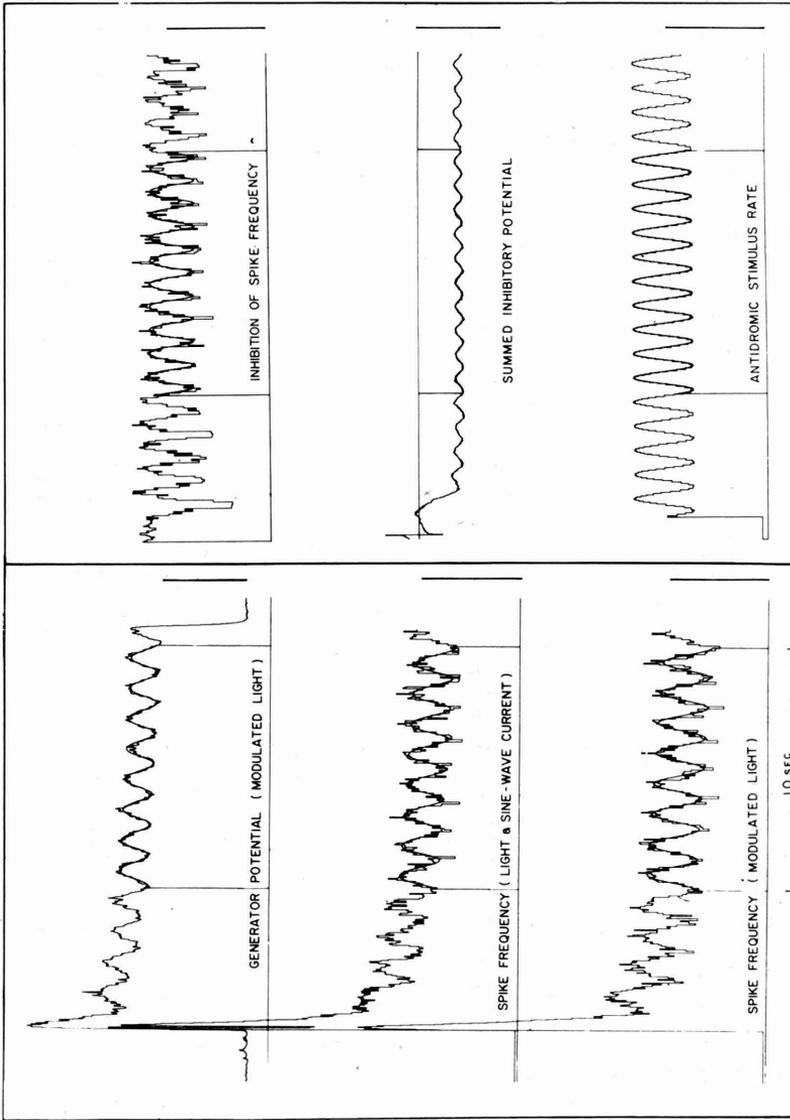


Figure 2. Experimental raw data for various input-output pairs. Left frames relate to the three component experiments of Figure 3.

what we said above. Proceed thus: start with sinusoidal input and represent the amplitude and phase of each transducer as a complex number. **The neuro-**anatomy of the network, with known directions of information flow, may be depicted as a net (or "graph") of directed lines, with a transducer at each vertex. Assign a symbol (a "signal variable") to each line. Add together the symbols of lines that converge on a vertex, multiply their sum by the

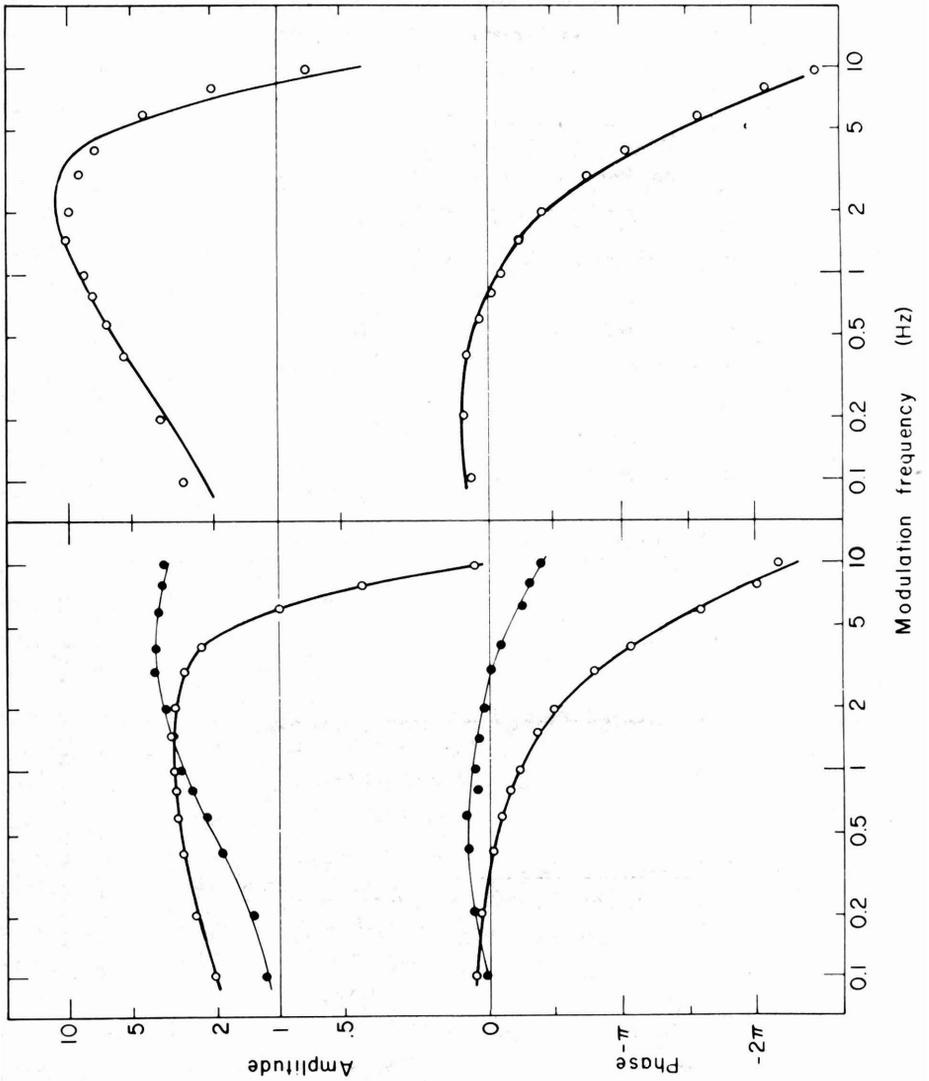


Figure 3. How responses of transducers in sequence combine. Left (bottom) frames show responses of component transducers. Open circles light-to-voltage, solid circles voltage-to-impulse rate. Right frames: curve predicted from left frame, circles observed light-to-impulse rate.

complex number which represents that transducer, and set the result equal to the output signal variable of the vertex. The result is simultaneous equations which may be solved algebraically for any signal variable in the network. Each signal variable so found is a complex number, whose amplitude and phase relate directly to the amplitudes and phases of the network's sinusoidal inputs. Finally, the neural network's dynamical response to arbitrary input may be calculated by using Fourier analysis to represent that input as a sum of temporal sinusoids. This prescription stems from familiar mathematical physics or signal theory.

It is notable that the rules we advance are appropriate for a system of linear components, and our biological network conforms well to these rules not by any necessity of underlying physical laws but rather by what looks like careful design. The same may be said, for example, about a public address system, and presumably similar demands of careful signal processing in both cases underlying this linearity. If the horseshoe crab's retina is confronted with a huge dynamic range of input, in fact it shows a nonlinearity in the form of a clever gain control; and in vertebrate retinas far more sophisticated gain controls are found. (Parenthetically, published discussion of nerve networks, which refers to feedback and nonlinearity as the same thing, perpetrates an unwitting confusion of the theoretical fundamentals).

Our prescription above in a mathematical sense furnishes a procedure for exactly predicting the outputs of a network of linear transducers. However, the horseshoe crab's eye contains more than 10^3 transducers interconnected by 10^5 signal channels, and the human eye contains 10^8 transducers. A procedure with the two goals of insight and efficiency must go further and deal systematically with these large numbers by use of further knowledge about the network's organization. It is an easy step to anticipate the replacement of sums by integrals, which will give the dynamical rules in terms of integral equations with complex-valued kernels. To explicitly state such an integral formulation requires knowledge of the density of signal paths connecting separated points in the network. But it is fairly evident that we can state the general form of the rule which relates stimulus input at points y on the retina to response of a transducer type at point x on the retina: if the stimulus input (sinusoidal in time at frequency $\omega/2\pi$) is $s(y)e^{i\omega t}$ where $s(y)$ is a complex number that specifies how both amplitude and phase depend on y , then the consequent transducer response $r(x)e^{i\omega t}$ will be related by a linear integral map of the form

$$r(x) = \int dy K\{x, y\}s(y) \quad (2.1)$$

where the (complex valued) two-point function K depends only on the nature of the nerve network and on the frequency $\omega/2\pi$. In particular, in (2.1) the response $r(x)$ can be the signal output on an optic nerve fiber at point x ; thus

the network's dynamical input-output relationship will be determined once we determine the corresponding K in (2.1).

Now Figure 1 suggests the anatomy of the horseshoe crab's retina is invariant under translations in the retina's plane, namely, that in equation (2.1) we have

$$K(x, y) = K(x - y) \quad (2.2)$$

This is in fact the case, and in consequence a stimulus which is spatially a sinusoidal plane wave which depends only on $p \cdot x$ (where p is the wave vector) will induce a likewise sinusoidal spatial dependence in every signal variable in the network. **In result the relations among transductions lose their dependence on position, reduce to simple algebra (parametric in p) and we may fully solve the network's dynamics in simple terms. Experimentally, we stimulate the eye with a spatial sinusoid; mathematically this corresponds to letting $s = s_0 \exp i p \cdot y$ and $r = r_0 \exp i p \cdot x$ in (2.1) which with (2.2) easily relates response to stimulus by**

$$r_0 = \tilde{K}(p, \omega) s_0 \quad (2.3)$$

where $\tilde{K}(p, \omega)$ is the spatial Fourier transform of $K(x)$.

Table 1 [adapted from (7)] shows what this combination of theory and experiment yields: a full characterization of the network's dynamics in terms of component transductions, with parameters evaluated. **Detailed inspection of these expressions shows that the effect of the network upon the spatial structure of the image is "Laplacian-like": changes of gradient are emphasized, and thus so are parts of the visual field where transitions in light intensity occur.**

To test whether our theory works we have used it to predict how the horseshoe crab's retina responds to a moving step of light intensity. This is shown in the left frame of Figure 4. Predicted responses to steps moving at four velocities are shown vertically. The center column shows how the center of the retina should respond: we have resolved the step into running sinusoids and multiplied each by the top expression of Table 1. **The left and right columns show what we predict at the left and right terminations of the neural network, by applying (7,15) the classical Wiener-Hopf theory for truncated translation kernels to our situation. The right frame of Figure 4 shows the result of the corresponding experiments. The twelve theory-versus-experiment pairs exhibit trends in numerous distinguished features, all of which are nicely captured by the dynamical theory.**

The review above summarizes the work of about 15 years by a very small club, and in which society made a dollar investment which alternatively could buy about 7/100 of an MX missile.

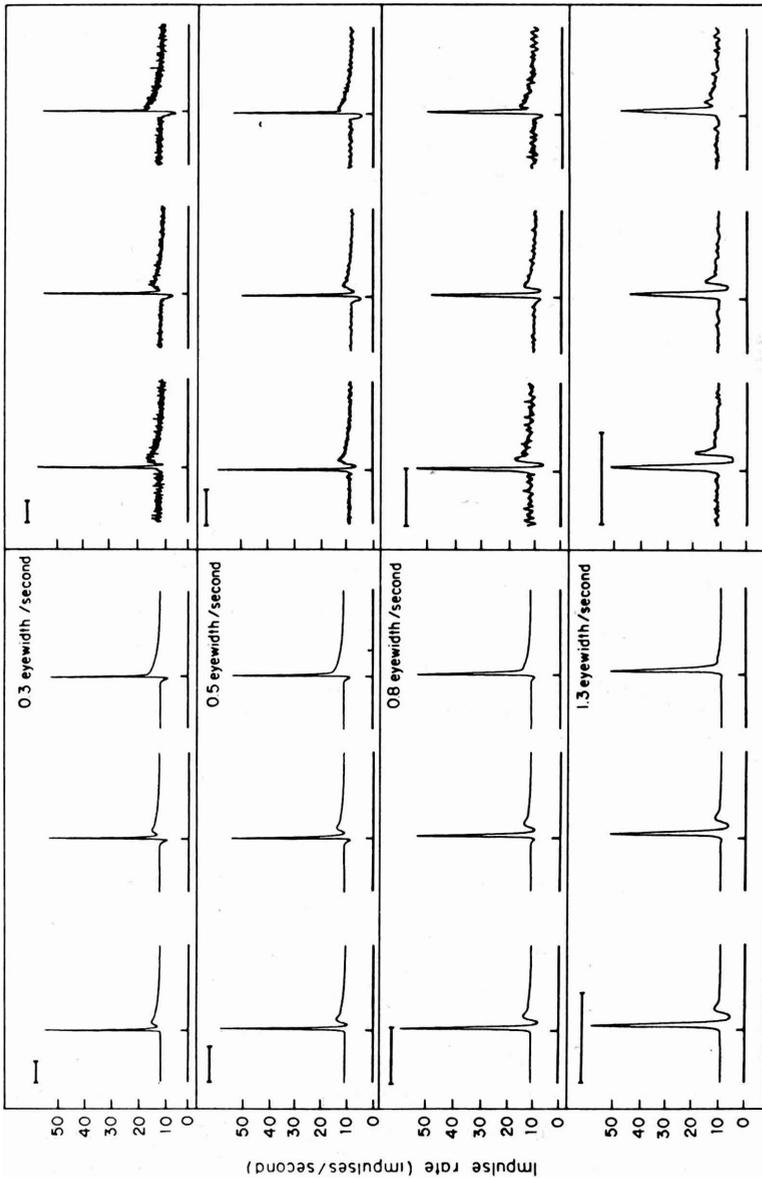


Figure 4. Response of horseshoe crab retinal network to a moving step in light intensity, theory and experiment. **Middle columns:** response in central area predicted from direct use of Table 1. **Left and right columns:** response at two edges from Table 1 used in Wiener-Hopf theory.

Table 1

Description	Equation
Spatiotemporal transfer function	$\tilde{K}(p, \omega) = \frac{E(\omega) G(\omega) \tilde{W}(p)}{1 + E(\omega) T_L(\omega) \tilde{H}(p)}$
Generator potential	$G(\omega) = e^{-i\omega t_\ell} \cdot \left(\frac{1}{1+i\tau_d\omega}\right)^{n_d} \cdot \left(\frac{1}{1+i\tau_b\omega}\right)^{n_b} \cdot \left(\frac{R}{1+i\tau_a\omega}\right) \cdot \left(\frac{i\tau_a\omega}{1+i\tau_a\omega}\right)^{n_p}$
Encoder	$E(\omega) = \frac{1}{1 + \frac{k}{1+i\tau\omega}}$
Lateral inhibition	$T_L(\omega) = \frac{1}{1-C} \left(\frac{1}{1+i\tau_1\omega} \cdot \frac{1}{1+i\tau_2\omega} - \frac{C}{1+i\tau_3\omega} \right) \cdot \frac{1}{1+i\tau_4\omega}$
Fourier transform of inhibitory kernel	$\tilde{H}(p) = K \cdot \frac{1-(p/a)^2}{(p/b)^4 + 2(p/c)^2 + 1}$
Point spread	$\tilde{W}(p) = \exp(-s^2 p^2 / 4)$

Parameter	Dimension	Value
t_ℓ	seconds	0.038
t_d	seconds	0.0076
n_d		3
t_b	seconds	0.017
R		0.75
t_a	seconds	0.030
n_p		0.25
T	seconds	1.5
T_1	seconds	0.40
T_2	seconds	0.036
T_3	seconds	0.055
T_4	seconds	0.036
C		0.019
K		0.1
a	Rad/eye width	1.0
b	Rad/eye width	17.56
c	Rad/eye width	23.61
s	eye widths	24.83
n_b		0.00951
		3

3. KERNELS WITH A SLOW SECOND DEPENDENCE

The kernel of equation (2.1) in the exercise above had the special feature that it was independent of its second ("mean position") argument in equations (1.3). In consequence its eigenfunctions in equation (1.2) were spatial sinusoids (and eigenvalues given by the Fourier transform $\tilde{K}(p)$ of $K(x)$). The full solution of the dynamical response problem was achieved by representing the stimulus as a superposition of those eigenfunctions of (1.2). Evidently the same method would work for more general $K\{x,y\}$ once we were in possession of its eigenfunctions and eigenvalues; in this sense the solution of (1.2) would solve the dynamical response problem for less ideal neural networks.

A mammalian retina (such as our own or a cat's) has a region of most acute resolution at its center, and this is reflected in the transduction kernel K , as shown in Figure 5 (shown schematically for $\omega=0$ whence K is without temporal phase shifts and is real). The departure of K from translational invariance is on a space scale ϵ slow compared to the space-scale of K 's transduction profile. This suggests that we insert a "parameter of slowness", ϵ in equation (1.2) and study our problem for small ϵ :

$$\int dy K(x-y, \frac{x+y}{2}) \psi(y) = \lambda \psi(x) \tag{3.1}$$

Asymptotic solution of (3.1) in detail will be discussed in section five below, but first we gain some good insight about (3.1) by studying K in a manner patterned after the previous section (where (3.1) had $\epsilon=0$). Define fast and slow variables

$$\begin{aligned} u &= x-y \\ q &= \epsilon(x+y)/2 \end{aligned} \tag{3.2}$$

and as above Fourier transform in the fast variable, to now obtain

$$\tilde{K}(p,q) = \int du \exp(-ip \cdot u) K(u,q) = WK\{x,y\}, \tag{3.3}$$

which is an evidently invertable linear transform of K introduced by Wigner (16) to quantum statistical mechanics long ago. The Wigner transform \tilde{K} reveals much about its inverse transform K .

Laboratory measurements (still underway) suggest^{17,18,19,20} that some mammalian retinas (cat or man) have a transfer kernel K whose Wigner transform is fairly well represented by the three-parameter expression

$$\tilde{K}(p,q) = \frac{1+q^2}{1+\frac{1}{\kappa}q^2} \{ e^{-(1+q^2)p^2} - z e^{-a(1+q^2)p^2} \}. \tag{3.4}$$

Here the frequency dependence stems entirely from the complex number z , which is measured at each temporal frequency. We have divided another overall

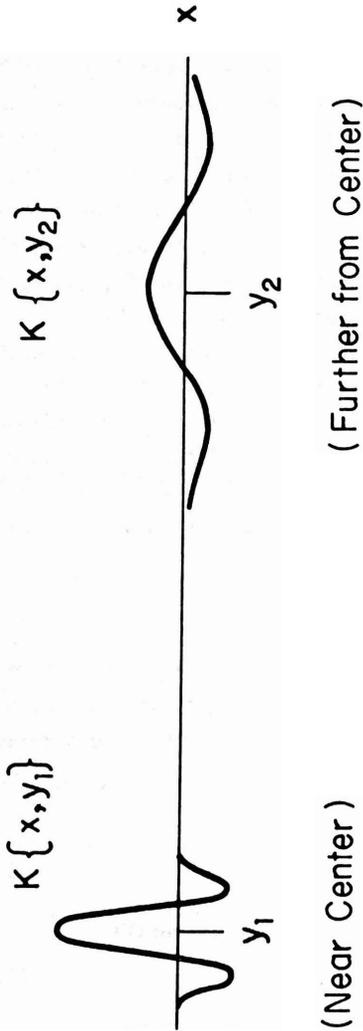


Figure 5. Slow change in response profile, as in retina of cat or human.

complex number from (3.4) and have scaled q to give another parameter the value unity. The eigenfunctions and the spectrum (which proves part discrete and part continuous) implied by (3.4) are our immediate goals^{21,22} but our approach is general.

The features of \tilde{K} in the p, q "phase space" may be related to the spectrum of K as follows. If K has the generic spectral representation

$$K\{\underline{x}, \underline{y}\} = \sum_n \lambda_n \psi_n(\underline{x}) \hat{\psi}_n(\underline{y}) \equiv \sum_n \lambda_n E_n\{\underline{x}, \underline{y}\} \tag{3.5}$$

(where the $\hat{\psi}$ are the adjoint eigenfunctions bi-orthonormal to the ψ) then we may define the "major projection operator"

$$F_c = \frac{1}{2\pi i} \int_c ds (s-K)^{-1} = \lambda_n \sum_n c E_n \tag{3.6}$$

where the inverse operator may be represented as

$$R_s = (s-K)^{-1} = \sum_n \frac{1}{s-\lambda_n} E_n \tag{3.7}$$

and (3.6) follows at once by the Cauchy residue theorem with a contour c as shown in Figure 6. We will estimate how the image $WF_c = \tilde{F}_c$ appears in phase space; first we note from equation (3.3) that

$$W\delta(\underline{x}-\underline{y}) = 1. \tag{3.8}$$

Also we anticipate that formally for the operator product AB of two kernels $A\{\underline{x}, \underline{y}\}$ and $B\{\underline{x}, \underline{y}\}$ the phase-space image will be

$$W(AB) = \tilde{A}(\underline{p}, \underline{q}) \tilde{B}(\underline{p}, \underline{q}) + o(\epsilon) \tag{3.9}$$

as in the limit $\epsilon \rightarrow 0$, AB is exactly a convolution of two translation kernels, and $W(AB) = \tilde{A}\tilde{B}$ is exactly the product of their Fourier transforms. If we let $A = s - K$ and $B = R_s$ as in (3.7), then $W(AB) = 1$ and formally to $O(\epsilon)$ equation (3.9) gives

$$R_s(\underline{p}, \underline{q}) = \frac{1}{s - \tilde{K}(\underline{p}, \underline{q})} \tag{3.10}$$

Take the Wigner transform of (3.6) and use (3.10) to evaluate the integrand of the residue integral. For the s -contour integration, $\tilde{K}(\underline{p}, \underline{q})$ is simply a parameter, and we see

$$F_c(\underline{p}, \underline{q}) = \frac{1}{2\pi i} \int_c ds \frac{1}{s - \tilde{K}(\underline{p}, \underline{q})} \tag{3.11}$$

approximately (while we note [from (3.6)] that

$$\tilde{F}_c(\underline{p}, \underline{q}) = \lambda_n \sum_n c \tilde{E}_n(\underline{p}, \underline{q}) \tag{3.12}$$

is exact). Equation (3.11) gives at once

$$F_c(\underline{p}, \underline{q}) = \begin{cases} 1 & \text{if } \tilde{K}(\underline{p}, \underline{q}) \text{ in } c \\ 0 & \text{otherwise} \end{cases} \tag{3.13}$$

Thus [to the extent K lets us ignore $O(\epsilon)$ and obtain (3.10)] the major projective operator $F_c\{\underline{x}, \underline{y}\}$, which leaves intact those $\psi_n(x)$ whose λ_n are enclosed in the contour and which annihilates their complementary subspace in

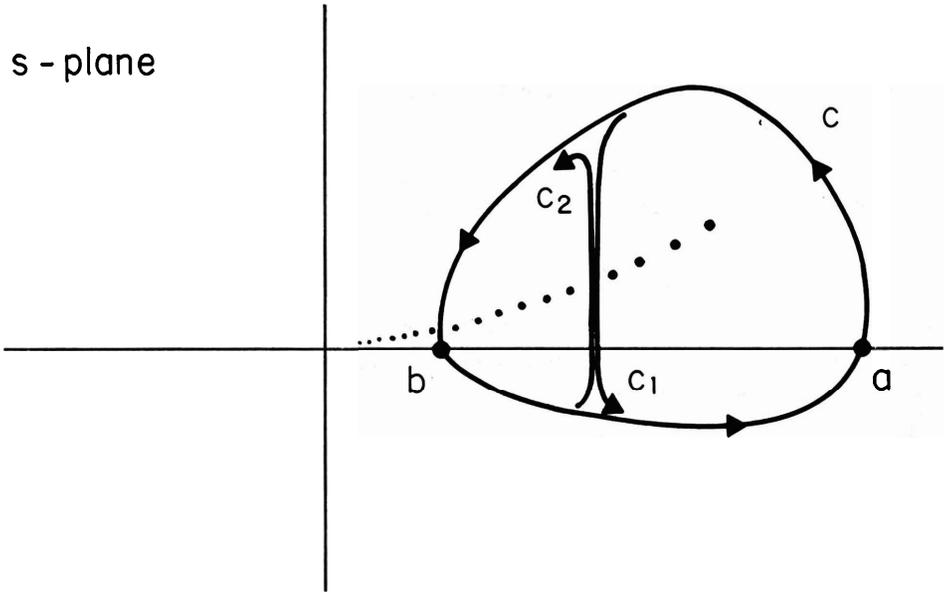


Figure 6. The contours c, c_1, c_2 on the s -plane, with the real points a, b . The set of isolated points depict the discrete spectrum.

function space, has a Wigner image which is a "characteristic function" that assigns unity to a subset of the p, q phase space and assigns zero to its complementary subset. Subdivision of the contour c (as into c_1 and c_2 in Figure 6) subdivides function space ($F_c = F_{c_1} + F_{c_2}$) and subdivides the characteristic function in phase space ($\tilde{F}_c = \tilde{F}_{c_1} + \tilde{F}_{c_2}$).

The exact result (3.12) states that \tilde{F}_c depends only on point spectrum enclosed and is otherwise independent of how c is routed. From this we conclude that if the approximation (3.11) is to be fairly good then $\tilde{K}(p, q)$ can take on only values near those of the point spectrum of K . In particular, choose a contour which consists of small circles which surround the enclosed spectrum; we see that each $\tilde{E}_n(p, q)$ can be substantial only near those points (p, q) where $\tilde{K}(p, q) = \lambda_n$.

We can furthermore measure the point set which supports the characteristic function $F_c(p, q)$. From (3.3) it is straightforward to show in general [from the δ -function's Fourier representation, and using (3.2)] that

$$\int dq dp \tilde{K}(p, q) = (2\pi\epsilon)^D \int dq K(q, q) = (2\pi\epsilon)^D \text{Tr}(K) \tag{3.14}$$

where D is the dimensionality of the vector q . As $\text{Tr}(E_n) = 1$, we have from (3.12) the measure of the set where $F_c(p, q) = 1$:

$$\int dq dp \tilde{F}_c(p, q) = (2\pi\epsilon)^D \cdot N \tag{3.15}$$

where N is the number of spectral points enclosed by c .

If $K\{x,y\}$ is Hermitian then its λ_n are real and its E_n Hermitian, and a further result emerges. Complex conjugation of (3.3) shows directly that $\tilde{K}(p,q)$ is real in this case. A "surface in phase space" where $\tilde{F}_c(p,q)$ descends from 1 to 0 in this case must occur at points p,q where $\tilde{K}(p,q)$ has the real value assumed where c crosses the real axis. Suppose such real crossing values are called "a" and "b". Then from (3.15) we can write the number of eigenvalues between "a" and "b" as

$$N = \frac{1}{(2\pi\epsilon)^D} \times \{\text{Volume of phase space between } K(p,q) = a \text{ and } K(p,q) = b\}. \tag{3.16}$$

This is a weaker (but more generally applicable) form of the "area rule" given below. From (3.16) we see that the Wigner image of (3.5),

$$\tilde{K}(p,q) = \sum_n \lambda_n \tilde{E}_n(p,q) \tag{3.17}$$

is a sum of terms which decompose phase space into "layers of an onion" with the n^{th} layer contributing only near where $\tilde{K} = \lambda_n$.

We now see how $\tilde{K}(p,q)$ may be used to find whether the operator K has a region of continuous spectrum, and estimate the eigenvalue where the continuum begins. The equation

$$\tilde{K}(p,q) = \lambda \tag{3.18}$$

specifies a "surface in phase space" (parametric in λ) which may either be closed or may extend all way to infinity, and a critical λ value which divides between these two conditions typically will divide between finite volume and infinite volume in equation (3.16). Infinite volume in (3.16) corresponds to eigenvalues packed with indefinite closeness, whence the critical λ from (3.18) corresponds to the start of the continuous spectrum. Our explicit \tilde{K} of equation (3.4), for real z between 0 and 1, and $0 < \kappa < 1$, easily yields such a critical λ . (It is $\lambda = \kappa(1 - \frac{1}{a}) / (az)^{\frac{1}{a-1}}$ with the continuum below.) For complex \tilde{K} a proper analysis of the continuum's edge must look deeper.

4. THE END OF THE POINT SPECTRUM AND THE SUMMIT OF \tilde{K}

The considerations above should lead us to expect (at least if \tilde{K} is real and ϵ is small) that if two different kernel transforms \tilde{K} both yield, over some range of λ , very similar phase-space surfaces for the solution of $\tilde{K}(p,q) = \lambda$, then over that range of λ their eigenvalue spectra λ_n will be close and so will be their corresponding projection operators $\psi_n(x) \hat{\psi}_n(y) = W^{-1} \tilde{E}_n(p,q)$ and hence their eigenfunctions. This stems, most briefly, from the partition of phase space by equation (3.17).

In particular, near its "summit" at (p_s, q_s) , \tilde{K} may be expanded through quadratic terms in $p - p_s$, $q - q_s$, giving an approximate \tilde{K}_S whose largest eigenvalues and corresponding eigenfunctions should well approximate those for the exact \tilde{K} . Now it was already well known to Wigner¹⁶ that any \tilde{K} which is a (real) polynomial in p corresponds to a $K = W^{-1}\tilde{K}$ which is a specific (Hermitian) differential operator. (For example, notice that the generalized function

$$K\{x, y\} = -\frac{1}{2}\delta''(x-y) + V(\epsilon \frac{x+y}{2}) \delta(x-y) \quad (4.1)$$

when substituted in (3.1) gives

$$K\psi(x) = (-\frac{1}{2} \frac{d^2}{dx^2} + V(\epsilon x)) \psi(x) \quad (4.2)$$

which is Schrödinger's Hamiltonian operator with coordinate $\epsilon x = q$ and ϵ taken as Planck's constant, while substitution in (3.3) gives

$$K = \frac{1}{2}p^2 + V(q), \quad (4.3)$$

the corresponding classical Hamiltonian function. For each of the polynomial terms in our case the demonstration is exactly similar. In summary, we have a well defined (Hermitian) second-order differential operator K_S , and near the top part of the spectrum, the eigenvalues and eigenvectors of our original integral operator K are well approximated by the solution of the eigenvalue differential equation

$$K_S \psi_n(\underline{x}) = \lambda_n \psi_n(\underline{x}) \quad (4.4)$$

where K_S consists only of constants and linear, bilinear, and quadratic terms in the components of \underline{x} and of $\partial/\partial \underline{x}$. For such an operator (4.4) can be solved exactly. Its spectrum is composite of D (the dimension of q) separate equally spaced component spectra with different spacings determined by a diagonalization of the coefficients of \tilde{K}_S and each starting a half-step down from the maximum $\tilde{K}(p_s, q_s)$. The eigenfunctions ψ_n are in closed form in terms of Hermite functions. (In short, (4.4) has a solution like that of a quantum harmonic oscillator in D dimensions.) The spectrum and eigenfunctions will be good up to that n [and λ_n from (4.4)] where the two equations $\tilde{K}(p, q) = \lambda_n$ and $\tilde{K}_S(p, q) = \lambda_n$ yield appreciably different p, q surfaces. We note that as agreement deteriorates with increasing n , the value of λ_n descends from the maximum of \tilde{K} toward smaller values. In a sense we have solved the "important" part of the eigenvalue problem, in that the answer of the question "what is the action of K on $f(\underline{x})$ " by eigenfunction expansion, depends less critically on estimating those ψ_n which go with small λ_n . (In our application to the retina the largest λ_n corresponds to a $\psi_n(\underline{x})$ which is the spatial pattern to which the retina is most sensitive, see refs. 21 and 22.)

For the special case of an integral equation in one dimension ($D=1$) phase space reduces to the (p,q) phase plane, on which the equation

$$\tilde{K}_S(q,p) = \lambda \quad (4.5)$$

yields a (λ -dependent) family of concentric ellipses each enclosing area $A(\lambda)$, and (4.4) is an ordinary differential equation which features the same coefficients¹⁰ as does (4.5). Exact solution for the λ_n shows that when $\lambda = \lambda_n$ in (4.5) then the elliptical (q,p) locus encloses

$$A(\lambda_n) = (n + \frac{1}{2}) \cdot 2\pi\epsilon \quad (4.6)$$

a stronger version of (3.16) which in fact (as we will discuss in the next section) holds, to $o(\epsilon)$ also for the eigenvalues of the exact K and area $A(\lambda)$ enclosed by the locus of $\tilde{K}(p,q) = \lambda$. [Equation (4.6) we call the "area rule".]

If, instead of expanding about a summit point of \tilde{K} , we obtain \tilde{K}_S by freely assigning real values to its various coefficients, then (4.5) may yield a second generic locus pattern which is concentric hyperbolas. For every λ the (p,q) locus extends to infinity and by the previous section we anticipate a continuous spectrum for the differential equation (4.4) which is confirmed by its exact solution in terms of parabolic cylinder functions.

Between ellipse and hyperbola, a contrived choice of real coefficients in (4.5) yields a locus-family of congruent parabolas symmetric about a horizontal line. Continuous spectrum is again indicated, and in this case equation (4.4) solves exactly in terms of the Airy Ai function, a fact important to our asymptotic considerations below.

Finally, suppose \tilde{K} is not Hermitian so that $\tilde{K}(p,q)$ is not real.^{12,13} We may still find a (complex p,q) solution to the equations

$$\partial\tilde{K}/\partial p = 0, \quad \partial\tilde{K}/\partial q = 0. \quad (4.7)$$

and develop a \tilde{K}_S through quadratic terms around that point in complex phase space. From this \tilde{K}_S by our same procedures as above we may obtain the approximate eigenvalue equation (4.4) (now no longer Hermitian) which is still formally solved by the same closed analytic expressions for ψ_n , λ_n which solved before, whence a discrete spectrum of complex λ_n emerges formally. Because \tilde{K} is analytic in p,q , a point distinguished by (4.7) is generically a saddle point in complex phase space. The discrete spectrum applies if two conditions are fulfilled: (1) the functions $\tilde{K}(p,q)$ and $\tilde{K}_S(p,q)$ should agree over a region of complex phase space which includes a "section" on which p,q are real, and (2) the formal eigenfunctions solving (4.4) must go to zero as we go to infinity for real χ . If the second condition is not fulfilled, then on the "real section" we have a case like that of the "hyperbolic locus" above, and a continuous spectrum is indicated.

These are explicit rules for estimating the more significant end of the discrete spectrum. For example, in our particular case (3.4) if our original kernel variable x (and hence p, q) was one dimensional, there are saddles where $q_s = 0$ and $p_s = \pm [(1n az)/(a-1)]^{\frac{1}{2}}$ which is complex along with z , and the estimated spectrum is

$$\lambda_n(z) = (1 - \frac{1}{a})(az)^{\frac{-1}{a-1}} \{1 - \epsilon(2n+1) \sqrt{\frac{1}{\frac{\kappa}{a^2} - a} - 1} (1n az)^{\frac{1}{2}}\} \quad (4.8)$$

The corresponding eigenfunctions are given in ref. 10 and they go to zero as we take x to ∞ provided $(1n az)^{\frac{1}{2}}$ lies in the right half of the complex plane.

The "cliff" between $\tilde{F}_c = 1$ and $\tilde{F}_c = 0$ in fact has a width roughly of $O(\epsilon/n_{1ast})$. More detailed development^{23,10} of (3.9) gives

$$W(AB) = \tilde{A}\tilde{B} - \epsilon \frac{i}{2} \sum_m \left(\frac{\partial A}{\partial p_m} \frac{\partial B}{\partial q_m} - \frac{\partial A}{\partial q_m} \frac{\partial B}{\partial p_m} \right) + O(\epsilon^2 \times 2^{nd} \text{ derivatives}). \quad (4.9)$$

Thus neglect of ϵ^1 appears justified on the basis of (3.13) except near the cliff where such neglect suddenly becomes very bad, and a more sensitive examination in "stretched" coordinates appears appropriate. The isomorphic quantum-mechanical question, "what is the Wigner transform, in semi-classical approximation, of a minimal Von Neuman density" has been elegantly elucidated by Berry.²⁴ Further results for near the cliff, which emerge from the framework discussed above, are in preparation by Lawrence Sirovich and myself.

5. THE EIGENFUNCTIONS AND HAMILTONIAN MECHANICS

For the structure of ordinary or partial differential equation eigenvalue problems with a slow scale, such as (4.2) and (4.4), a revealing asymptotic analysis has been achieved by Keller²⁵ and by Keller and Rubinow.²⁶ A solution is assumed in the form

$$\psi(x) = \sum_b \psi_b(x) = \sum_b \exp \frac{i}{\epsilon} S(b)(\epsilon x, \epsilon) \quad (5.1)$$

where each (generally complex) S is of the form

$$S(q, \epsilon) = S_0(q) + \epsilon S_1(q) + O(\epsilon^2) \quad (5.2)$$

and the sum in (5.1) is over "branches". Typically each term of (5.1) gives a breakdown by going infinite on some subset of x , and a second form of asymptotic expansion valid near that set is matched to the terms of (5.1). Consistency of all matches [single-valuedness of ψ in (5.1)] proves possible only for certain discrete eigenvalues.

Our integral equation (3.1) proves to be a straightforward extension for this procedure.¹¹ Let $\epsilon x = q$ as in (5.1), (5.2), let $x - y = u$ and substitute (informally--in "hope") one term of (5.1) in (3.1) to find

$$\int du K(u, q - \frac{\epsilon}{2}u) \exp \frac{i}{\epsilon} \{S(q - \epsilon u) - S(q)\} = \lambda. \tag{5.3}$$

In the limit $\epsilon \rightarrow 0$ the difference-quotient in the exponent becomes a derivative in the $-u$ direction and we have by comparison with (3.3)

$$\tilde{K}(\partial S_0(q)/\partial q, q) = \lambda. \tag{5.4}$$

There is a standard reduction (due to Monge; see Fritz John,²⁷ also ref. 11) of this nonlinear first order partial differential equation to the solution of a system of ordinary differential equations. Introduce the gradient of S_0 as an auxiliary vector field

$$p = \partial S_0 / \partial q \equiv S_{0q} \tag{5.5}$$

and solve the ordinary differential equations (1.4) for $p(t), q(t)$. Then along the line $q(t)$ integrate

$$dS_0/dt = p(t) \cdot (dq(t)/dt) \tag{5.6}$$

and S_0 along that line will be embedded in a solution of (5.4). Thus \tilde{K} plays the role of a Hamiltonian in classical mechanics; our frequently useful $\tilde{K}(p, q) = \lambda$ above chooses an energy surface to which trajectories are confined, the continuous spectrum's end value is the energy at which classical motion can become unbounded, and the spectrum-ending "summits" given by (4.7) are singular points of (1.4). If (1.4) had been given to us, (5.4) would be our Hamilton-Jacobi equation.

By developing (5.3) through ϵ^1 we find the following:¹¹ Let $\exp iS_1(q) = A(q)$ (S_1 proves imaginary generally for real K) then

$$(\partial/\partial q) \cdot (\tilde{K}_p(S_{0q}, q) A^2(q)) = 0, \tag{5.7}$$

a zero-divergence equation for the indicated vector field, which is colinear with the vector field \tilde{K}_p which directs $dq(t)/dx$ according to (1.4). While (5.7) devolves naturally from the usual partial differential equations which contain an exact conservation law, it is remarkable for our non-local operator which does not. Equation (5.7) tells us that where \tilde{K}_p goes to zero, (or more generally at "caustics" where different flow lines $q(t)$ touch) A becomes infinite, and that is the breakdown of (5.1). As differential operators like (4.2) are subsumed in our approach, we may (proceeding with sensible caution) apply the results of the program stated by Keller to our broader context here.

For a 1-dimensional integral equation, we may solve $K(p, q) = \lambda$ for $p = p_b(q, \lambda)$ as shown in Figure 7. In this case the differential equations (5.6), (5.7) integrate at once, yielding to $O(\epsilon)$

$$\psi(q) = \sum_b \psi_b(q) = \sum_b (\tilde{K}_p(p_b(q)))^{-\frac{1}{2}} \exp \frac{i}{\epsilon} \int^q p_b(q) dq. \tag{5.8}$$

Evidently (5.8) allows us to "follow the phase" of any ψ_b : the phase advances like $1/\epsilon$ times the area under the $p_b(q, \lambda)$ curve up to given q . Where the level line $\tilde{K}(p, q) = \lambda$ becomes vertical, \tilde{K}_p vanishes and (5.8) breaks down. But in this region $p(q, \lambda)$ becomes a horizontal parabola for which we have (as noted in the last section) an explicit solution for ψ in terms of the Airy function. This solution matches to (5.8) and shows explicitly that the phase of S changes by $+$ (or $-$) $\pi/2$ as we round a parabola that turns our path rightward (or leftward). In this way we find the phase change around the total path, which must be $2\pi n$ if ψ is to be single valued, and the area rule (4.6) (see refs. 8,9,10) follows at once: if $A(\lambda) = 2\pi\epsilon (n + \frac{1}{2})$ then the value of K on the level line $\tilde{K}(p, q) = \lambda$ is the eigenvalue removed n th away from a singular point, distinguished by (4.7), which the level line encloses.

The area rule continues to hold good beyond the "equally-spaced-spectrum" estimate of the last section. For ordinary differential operators it reduces to the familiar W.K.B. result. Not surprisingly, it also applies (with proper technical care) for multi-dimensional kernels which correspond to separable Hamiltonians \tilde{K} . Our example (3.4), which has two ignorable angles (as does the classical "central force" Hamiltonian), is a case in point.

Now the solution $p(q, \lambda)$ to $\tilde{K}(p, q) = \lambda$ is locally analytic in λ , and analytic in q also except at branch points where $\tilde{K}_p = 0$. The closed curve of Figure 7 corresponds to a prescription for choosing a contour γ (avoiding branch cuts) on the complex q plane to evaluate the area:

$$A = \int_{\gamma} p(q; \lambda, z) dq \tag{5.9}$$

where the further locally analytic "z" dependence, as in (3.4), has been explicitly recognized. In principle we may locally invert the analytic relation (5.9):

$$\lambda = \Lambda(A, z). \tag{5.10}$$

A thoughtful review of the area rule (4.6) reveals that it is a condition for single-valued $\psi(q)$ on the complex q -plane, with no demand that K be Hermitian nor that the isolated branch points of $p(q, \lambda)$ be at real q .^{12,13}; see also 28,29 Thus for small ϵ , (5.10) and the area rule asymptotically solve our eigenvalue problem:

$$\lambda_n(z) = \Lambda(2\pi\epsilon(n + \frac{1}{2}), z) \tag{5.11}$$

which gives complex λ_n as a function of complex z . **10**

The technical challenge of inverting (5.9) to (5.20) may be met in several practical ways. We conclude with one: the infamous Birkhoff transformation.^{30,31} (Infamous in classical mechanics for $D \geq 2$ and non-separable Hamiltonian, in which case it is generically divergent.). The Birkhoff transformation, limited

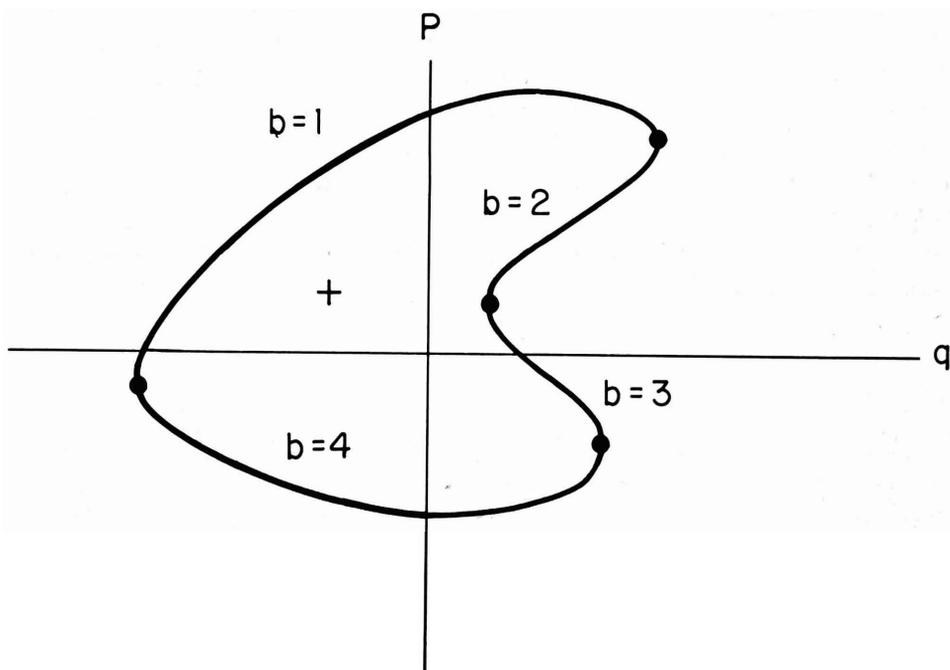


Figure 7. Locus of $K(p,q) = \lambda$ on the phase plane, showing 4 branch points and 4 branches of solutions $p_b(q,\lambda)$. The cross depicts the enclosed extremum of \tilde{K} .

to the phase plane,

$$p = P(\hat{p}, \hat{q})$$

$$q = Q(\hat{p}, \hat{q}) \tag{5.12}$$

proceeds, by an algebraic algorithm, to develop (5.12) as power series which converge, if at all, to an area preserving map which takes level lines of $\tilde{K}(p,q)$ (as in Figure 7) to concentric circles on the (real) \hat{p}, \hat{q} plane. Substitution of (5.12) into $\tilde{K}(p,q)$ gives

$$\lambda = \tilde{K}(p,q; z) = \Gamma(\hat{p}^2 + \hat{q}^2, z) \tag{5.13}$$

as a power series in $\hat{p}^2 + \hat{q}^2$ with functions of z as coefficients. As (5.12) is area preserving, this implements the inversion (5.10) and (5.11) becomes explicitly

$$\lambda_n(z) = \Gamma(\epsilon(2n+1), z) , \tag{5.14}$$

a series in $\varepsilon(2n+1)$ which (for a 1-dimensional or separable Hamiltonian) converges up to $\varepsilon(2n+1) = 0(1)$. Our example (4.8) in fact gave this series through its first two terms; the next term, elaborate but straightforward (carried out in ref. 22), tells us how the complex point spectrum first departs from uniform spacing.

The question (which is not a pressing one for our laboratory's present application), of what the same procedure tells us about the true discrete spectrum near a singular point of a non-separable $\tilde{K}(p,q)$, is an intriguing question that connects the eigenvalue problem with the really difficult part of Hamiltonian mechanics.

REFERENCES

- 1) F Ratliff, H K Hartline, and W H Miller, Spatial and temporal aspects of retinal inhibitory interaction. *J. Opt. Soc. Amer.* 53 (1963) 110-120.
- 2) H K Hartline, Visual receptors and retinal interaction, *Science* 164 (1969) 270-278.
- 3) B W Knight, J Toyoda and F A Dodge, A quantitative description of the dynamics of excitation and inhibition in the eye of Limulus. *J. Gen. Physiol.* 56 (1970) 421-437.
- 4) F Ratliff, B W Knight, F A Dodge and H K Hartline, Fourier analysis of dynamics of excitation and inhibition in the eye of Limulus, amplitude, phase, and distance, *Vision Res.* 14 (1974) 1155-1168.
- 5) S E Brodie, B W Knight and F Ratliff, The response of the Limulus retina to moving stimuli: a prediction by Fourier synthesis. *J. Gen. Physiol* (1978) 129-166.
- 6) S E Brodie, B W Knight and F Ratliff, The spatiotemporal transfer function of the Limulus lateral eye. *J. Gen. Physiol.* 72 (1978) 167-202.
- 7) L Sirovich, S E Brodie, and B W Knight, Effects of boundaries on the response of a neural network. *Biophys. J.* 28 (1979) 423-446.
- 8) L Sirovich and B W Knight, On the eigentheory of operators which exhibit a slow variation. *Quart. Appl. Math.* 38 (1981) 469-488.
- 9) L Sirovich and B W Knight, Contributions to the eigenvalue problem for slowly varying operators, *SIAM J. Appl. Math.* 42 (1982) 356-377.
- 10) B W Knight and L Sirovich, The Wigner transform and some exact properties of linear operators. *SIAM J. Appl. Math.* 42 (1982) 378-389.
- 11) L Sirovich and B W Knight, Two scale integral eigenfunction problems in several dimensions (submitted P.N.A.S.).
- 12) B W Knight and L Sirovich, W.K.B. method for non-Hermitian integral operators. I. Generalized Bohr-Sommerfeld-Keller rule (in print).
- 13) L Sirovich, W.K.B. method for non-Hermitian integral operators. II. Complex Hamiltonian theory (in print).

- 14) F Ratliff, B W Knight and N Milkman, Superposition of excitatory and inhibitory influences in the retina of Limulus. Proc. Nat. Acad. Sci. 67 (1970) 1558-1564.
- 15) L Sirovich, Boundary effects in neural networks. SIAM J. Appl. Math. 39 (1980) 142-160.
- 16) E P Wigner, On the quantum correction for thermodynamic equilibrium. Phys. Rev. 40 (1932) 749-759.
- 17) C Enroth-Cugell and J G Robson, The contrast sensitivity of retinal ganglion cells of the cat. J. Physiol. 187 (1966) 517-552.
- 18) H R Wilson and S C Giese, Threshold visibility of frequency gradient patterns. Vision Res. 17 (1977) 1177-1190.
- 19) R A Linsenmeier, L J Frishman, H G Jakiela and C Enroth-Cugell, Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurements. Vision Res. 22 (1982) 1173-1183.
- 20) S Dawis, E Kaplan, R Shapley and D Tranchina, The receptive field organization of X-cells in the cat: spatiotemporal coupling and asymmetry. Vision Res. (in print).
- 21) L Sirovich and B W Knight, Characteristic responses of models of an inhomogeneous retina (submitted for publication).
- 22) B W Knight and L Sirovich, Dynamic response spectrum and stimulus eigenfunctions of a model inhomogeneous retina (in preparation).
- 23) J E Moyal, Quantum mechanics as a statistical theory. Proc. Camb. Phil. Soc. 45 (1949) 99-124.
- 24) M V Berry, Semi-classical mechanics in phase space: a study of Wigner's function. Phil. Trans. Roy. Soc. Lond. 287 (1977) 237-271.
- 25) J B Keller, Corrected Bohr-Sommerfeld quantum conditions for nonseparable systems. Ann. Phys. 4 (1958) 180-188.
- 26) J B Keller and S I Rubinow, Asymptotic solution of eigenvalue problems. Ann. Phys. 9 (1960) 24-75.
- 27) F John, Partial Differential Equations. (Springer-Verlag, New York, 1971).
- 28) W H Furry, Two notes on phase-integral methods. Phys. Rev. 71 (1947) 360-371.
- 29) M V Berry and K E Mount, Semiclassical approximations in wave mechanics, Repts. Prog. Phys. 35 (1972) 315-397.
- 30) G D Birkhoff, Dynamical Systems. Amer. Math. Soc. 1927, revised edition 1966.
- 31) J Moser, Lectures on Hamiltonian systems, 1968, in: American Mathematical Society Memoirs 81, Amer. Math. Soc., Providence, R.I., 1968.