SIAM J. APPL. MATH.
Vol. 42, No. 2, April 1982


0036-1399/82/4202-0013 $01.00/0

# CONTRIBUTIONS TO THE EIGENVALUE PROBLEM FOR SLOWLY VARYING OPERATORS*

L. SIROVICH†‡ AND B. W. KNIGHT†

**Abstract.** We consider Hermitian operators with kernels which exhibit a slow variation, $K(x-y, \varepsilon(x+y)/2)$. The Wigner transform of such a kernel is $\tilde{K}(p, q) = \int K(u, q) \exp(-iup)\, du$. It has been shown (Sirovich and Knight, Quart. Appl. Math., 38, (1981), pp. 469–488) that eigenvalues follow an area rule: $\lambda_n$ is approximately an eigenvalue if the locus of $\lambda_n = \tilde{K}(p, q)$ encloses an area $(2n+1)\pi\varepsilon$. In this paper we illustrate this construction along with the associated eigenfunctions for a class of operators which model retinal organization. In addition, the eigenvalue problem for small values of the index $n$ (principal eigenfunctions and eigenvalues) is considered in detail. Modified eigenfunctions are obtained in this limit. Also considered is the splitting of degenerate eigenvalues and related nonuniformities.

**1. Introduction.** In a previous paper [1] we developed a method for treating the eigenvalue problem

$$\int_{-\infty}^{\infty} K\{x, y\}\psi(y)\, dy = \lambda\psi(x) \tag{1.1}$$

for symmetric operators which exhibit slow variation. That investigation gave an asymptotic approximation for the calculation of eigenvalues which we called the area rule. In brief this is given as follows: If $K\{x, y\} = K(x-y, \varepsilon(x+y)/2) = K(u, q)$ is the kernel of an operator and

$$\tilde{K}(p, q) = \int_{-\infty}^{\infty} K(u, q) \exp(-iup)\, du \tag{1.2}$$

the indicated transform, then $\lambda_n$ approximates the $n$th eigenvalue if

$$\lambda_n = \tilde{K}(p, q) \tag{1.3}$$

is a simple closed curve in the $(p, q)$-plane which contains an area

$$\mathscr{A}(\lambda_n) = (2n+1)\pi\varepsilon. \tag{1.4}$$

In all of this $\varepsilon$ represents the small parameter of slowness.

Both the methods and results of [1] are related to the classical WKB theory [2] and may be regarded as an extension of that theory to operators more general than the second order differential operators usually considered. As will be seen in the cases treated here, the strongly geometrical flavor to the area rule also provides a valuable qualitative tool in assessing the spectrum. A comment on the form of the transform (1.2) is also in order. Our use of it arose quite naturally in the development given in [1]; however, it was introduced much earlier in another context by Wigner [3], and we therefore refer to it as the Wigner transform. Moyal [4], Groenewald [5] and later Bruer [6] have also used it in an alternate formulation of quantum mechanics. In [7] (this issue, pp. 378–389) we treat the Wigner transform in some detail and show that it leads to a variety of exact results for both spectrum and eigenfunctions. It also leads to the exact solution of a wide class of operators.

* Received by the editors April 9, 1980, and in revised form March 3, 1981. This research was supported in part by the National Science Foundation under grant MCS-77-08598 and the U.S. Eye Institute under grants EY188, EY1428 and EY1472.

† Laboratory of Biophysics, Rockefeller University, New York, New York 10021.

‡ Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912.

Mention should also be made of the deep generalizations of WKB theory to higher dimensions for differential operators made by Keller [8] (see also Keller and Rubinow [9] and Keller and McLaughlin [10]), Maslov [11] and Arnold [12] (see also Arnold [13, Appendix 11]). The treatment which we give is restricted to one-dimensional cases, and as a result avoids the elaborate mathematical structure found in these papers.

In the present paper we treat a number of cases which both illuminate and extend the analysis presented in [1]. The examples which we give are drawn from models of visual information processing, which was the original impetus for these investigations. We note, however, that for purposes of illustration these models have been considerably simplified. Some of the examples in the following section have been chosen to illustrate features of the area rule. The majority of examples, however, illustrate breakdown of this procedure and the need for refinements. In particular, the discussion of principal eigenvalues given in [1] will be extended in § 3. The occurrence of eigenvalue degeneracy and subsequent splitting will be treated in § 4. Both of these effects are examples of nonuniformity. Another related nonuniformity will be treated in § 5.

Some insight into the area rule follows from consideration of (1.1) in the form

$$\int K\left(x-y, \frac{\varepsilon(x+y)}{2}\right)\psi(y)\,dy = \lambda\psi(x). \tag{1.5}$$

As shown in [1], an asymptotic analysis as $\varepsilon \downarrow 0$ implies that (1.1) has solutions which to lowest order have the form

$$\psi \sim A(\varepsilon x)\exp\left[\frac{i}{\varepsilon}\int^{\varepsilon x} p(s)\,ds\right] \tag{1.6}$$

with amplitude $A$ and phase rate $p$ to be determined. We substitute (1.6) into (1.5) and note that $K$ is a peaking function of its first argument relative to the second argument, so that we may approximate $\varepsilon(x+y)/2$ by $q=\varepsilon x$. Thus to lowest order

$$\begin{aligned}&\int K\left(x-y, \frac{\varepsilon(x+y)}{2}\right)A(\varepsilon y)\left[\exp\left(\frac{i}{\varepsilon}\int_{\varepsilon x}^{\varepsilon y} p(s)\,ds\right)\right]dy\\ &\qquad \sim \int K(x-y,q)A(q)[\exp(-ip(q)(x-y))]\,dy,\end{aligned} \tag{1.7}$$

from which (1.3) directly follows.

In most cases of interest (1.3) generates a family of closed loops in the $(p, q)$ plane. For a fixed value of $\lambda$ the real roots $p(q)$ of $\tilde{K}(p,q)=\lambda$ are given by the branches of the closed loop, and the required eigenfunction is a sum with contributions from each branch. Pairs of branches coalesce at branch points (points of vertical tangency) in the neighborhood of which the form (1.6) fails. In such cases a more sensitive analysis may be given, but for present purposes it suffices to remark that this results in the demand that the phases of the two local contributions to the eigenfunction must differ by $\pi/2$. Thus, if we denote the leftmost branch point by $Q_0$ and phase by $\phi$, the phase condition near $Q_0$ is

$$\Delta\phi = \frac{1}{\varepsilon}\int_{Q_0}^{q} \Delta p(s)\,ds - \frac{\pi}{2}, \tag{1.8}$$

where $\Delta p(q)$ represents the difference in the two branches near $Q_0$. As the point $q$

moves to the right, real roots of (1.3) appear or disappear as branch points are reached. At each branch point the phase difference is either augmented or decremented by $\pi/2$, and a lobe of area is either added or subtracted from (1.8). Near the final branch point $Q^0$ the additional phase factors have alternately canceled, the integral in (1.8) is near the area in the closed curve (1.3) and the phase condition on $\Delta\phi$ is that it become $\pi/2$ at $Q^0$. In particular,

$$\Delta\phi = \frac{1}{\varepsilon}\int_{Q_0}^{Q^0} \Delta p(s)\, ds - \frac{\pi}{2} = \frac{\pi}{2}. \tag{1.9}$$

The integral in (1.9) is $\mathscr{A}$ and (1.9) is (1.4) modulo $2n\pi$, which is just the indeterminacy in phase.

**2. Examples.** We start with a discussion of the kernel

$$K\{x, y\} = \frac{A(q)}{2l(q)} \exp\left[-\frac{|x-y|}{l(q)}\right], \tag{2.1}$$

where

$$q = \frac{\varepsilon(x+y)}{2}.$$

For background we mention that (2.1) is a simple model of a visual system for which $l$ measures the neutral spread and $A/2l$ the density of photoreceptors.

Under the Wigner transform (1.2), this becomes

$$\tilde{K} = \frac{A(q)}{1+p^2 l^2(q)}. \tag{2.2}$$

In spite of the simplicity of (2.1) (or (2.2)) it gives rise to a rich variety of features, some of which are unusual. We note that if $\varepsilon = 0$ (2.1) becomes a convolution kernel, with a continuous spectrum which lies in the interval

$$A(0) \geqq \lambda > 0,$$

and sinusoidal eigenfunctions.[1] We first consider a case for which the spectrum becomes purely discrete under $\varepsilon$-perturbation.

*Pure point spectrum.* In [1] we considered an exactly solvable kernel. We now consider another solvable case, but one which is of the general type considered in this section,

$$K\{x, y\} = \frac{1}{2[1+(\varepsilon(x+y)/2)^2]^{1/2}} \exp\left[-|x-y|\left[1+\left(\frac{\varepsilon(x+y)}{2}\right)^2\right]^{1/2}\right], \tag{2.3}$$

so that $l(q) = (1+q^2)^{-1/2}$ and $A = (1+q^2)^{-1}$. Hence

$$\tilde{K} = \frac{1}{1+q^2+p^2} = \tilde{K}(p^2+q^2). \tag{2.4}$$

The $\lambda$-curves are circles as indicated in Fig. 1, and the area rule yields

$$\lambda_n = \frac{1}{1+(2n+1)\varepsilon}. \tag{2.5}$$

[1] Here and in the following the term eigenfunction can correspond to both discrete and continuous spectra.

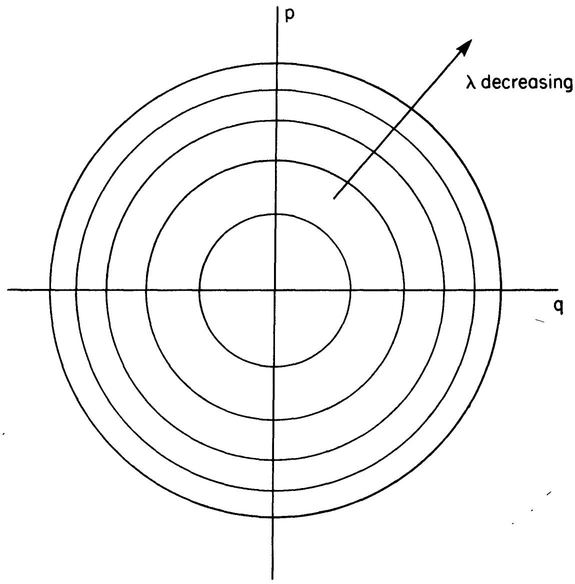


FIG. 1. *λ-curves for* $\tilde{K} = 1/(1+p^2+q^2)$. *The radii of the circle are proportional to* $\sqrt{n}$.

As shown in [7], this is one of a class of solvable cases for which the exact eigenfunctions are

$$\psi_n = H_n(\varepsilon x)\exp\left(-\frac{\varepsilon^2 x^2}{2}\right), \tag{2.6}$$

where $H_n$ represents the Hermite polynomial of degree $n$ [14]. The corresponding eigenvalue is exactly given by

$$\Lambda_n = (-1)^n \int_0^\infty L_n(2x)\tilde{K}(\varepsilon x)\exp(-x)\,dx,$$

where $L_n$ denotes the Laguerre polynomial of degree $n$ [15]. For the case at hand

$$\Lambda_n = (-1)^n \int_0^\infty \frac{L_n(2x)\exp(-x)}{1+\varepsilon x}\,dx. \tag{2.7}$$

Under the limit $n$ fixed, $\varepsilon \downarrow 0$,

$$\Lambda_n \sim (-1)^n \int_0^\infty L_n(2x)(1-\varepsilon x)\exp(-x)\,dx = 1-(2n+1)\varepsilon, \tag{2.8}$$

which is seen to agree with (2.5) under the same limit. The other limit of interest, $\varepsilon$ fixed $n \uparrow \infty$, is more troublesome. To deal with this case, we first observe that (2.7) may be written as

$$\Lambda_n = (-1)^n \int_0^\infty ds\, e^{-s} \int_0^\infty dx\, L_n(2x)\exp[-x(1+\varepsilon s)].$$

The second integration is of a standard form [15], and gives

$$\Lambda_n = \int_0^\infty \left(\frac{1-y}{1+y}\right)^n \frac{1}{\varepsilon}\exp\left[-\frac{y}{\varepsilon}\right]\frac{dy}{1+y}.$$

The limit $n \uparrow \infty$, $\varepsilon$ fixed, now follows from application of Laplace's formula [16], and we find

$$\Lambda_n \sim \frac{1}{1+(2n+1)\varepsilon}.$$

Thus comparison of this and (2.8) with (2.5) shows that the area rule is uniformly valid. This is in contrast with the example given in [1], where the area rule failed in the limit just considered. An explanation of this point is given in [7]. It is perhaps of interest to note in passing that the kernel (2.3) is not of trace class.

*Point and continuous spectra.* Without loss of generality we may normalize $A(q)$ in (2.1) such that $A(0)=1$. (The case $A(0)=0$ will be included in later discussion.) Also in the interest of simplicity we take $A(-q)=A(q)$, and for the moment we assume $A(q)$ decreases monotonically from the origin. The $\lambda$-curves

$$\tilde{K}(p,q)=\lambda=\frac{A(q)}{1+p^2l^2(q)} \tag{2.9}$$

are therefore fully specified by their form in the first quadrant.

We recall from [1] that the asymptotic eigenfunctions are composed of linear combinations of

$$\psi_w^{\pm}=\frac{\exp\left[\pm(i/\varepsilon)\int_0^q p(q')\,dq'\right]}{\sqrt{\tilde{K}_p(p(q),q)}}, \tag{2.10}$$

which we term WKB solutions because of their role in that theory. From (2.9) we see that

$$\tilde{K}_p=\frac{-2pl^2A}{(1+p^2l^2)^2}=-\frac{2pl^2}{A}\lambda^2 \tag{2.11}$$

and

$$p^2=\frac{1}{l^2}\left(\frac{A}{\lambda}-1\right), \qquad p=\pm\frac{1}{l}\sqrt{\frac{A}{\lambda}-1}. \tag{2.12}$$

From the assumptions made on $A(q)$, (2.9) implies $\lambda$ is restricted to the range of $\tilde{K}$, viz. $0<\lambda<1$.

Except under very special circumstances discrete eigenvalues only occur when the corresponding $\lambda$-curve is a closed loop.[2] In this case it follows that $\tilde{K}_p$ must vanish at some point of the curve, and from (2.11) that $p=0$ at such a point. If we write $A_\infty=A(\infty)$, then (2.12) implies that the discrete eigenvalues are restricted to the interval $A_\infty<\lambda<1$. Therefore if $A_\infty>0$, the full interval

$$0<\lambda<A_\infty \tag{2.13}$$

constitutes the continuous spectrum of the operator, and if $A_\infty=0$ the spectrum is entirely discrete.

To examine the discrete spectrum we consider the area rule, which states in this case

$$(2n+1)\pi\varepsilon=4\int_0^{q_0}\frac{1}{l(q')}\left(\frac{A(q')}{\lambda}-1\right)^{1/2}dq', \tag{2.14}$$

[2] It is possible to have unclosed $\lambda$-curves extending to infinity which still enclose finite area.

where $A(q_0)=\lambda$. Under an obvious variable change (2.14) becomes

$$(2n+1)\pi\varepsilon=-4\int_\lambda^1 \frac{1}{l\,dA/dq}\left(\frac{A}{\lambda}-1\right)^{1/2} dA.$$

A particularly simple form results for the subset of cases where $A$ and $l$ are related by

$$l\frac{dA}{dq}=-1,$$

in which case we obtain

$$(2n+1)\pi\varepsilon=\frac{8}{3}\lambda\left(\frac{1-\lambda}{\lambda}\right)^{3/2}. \tag{2.15}$$

For this form it is clear that if $A_\infty>0$ then the operator $K$ possesses only a finite number of discrete eigenvalues in addition to the continuous spectrum (2.13). On the other hand, if $A_\infty=0$, we obtain a pure point spectrum with $\lambda=0$ as an accumulation point.

*Infinite point spectrum plus continuous spectrum.* It is also possible within this framework to have an infinite point spectrum in addition to a continuous spectrum. For this and other reasons we introduce the following simple model:

$$K\{x,y\}=\frac{1}{2}\frac{1+k|q|}{1+|q|}\exp\left[-|x-y|\right] \tag{2.16}$$

with $1>k\geqq 0$. In this case

$$\tilde{K}=\frac{1+k|q|}{1+|q|}\frac{1}{1+p^2}=\lambda.$$

The point spectrum is restricted to the interval

$$k<\lambda<1, \tag{2.17}$$

while the continuous spectrum lies in the interval

$$0<\lambda<k. \tag{2.18}$$

As before we may restrict attention to the first quadrant. Then, as is easily seen,

$$q_0=\frac{1-\lambda}{\lambda-k} \tag{2.19}$$

is a turning point and

$$p^2=\frac{\lambda-k}{\lambda}\frac{q_0-q}{1+q}.$$

A sketch of the $\lambda$-curves is shown in Fig. 2. Closed contours are obtained for (2.17) while for (2.18) we obtain open contours starting at $p=\pm\sqrt{(1-\lambda)/\lambda}$ when $q=0$ and asymptoting to $\pm\sqrt{(1-k)/\lambda}$ when $|q|\uparrow\infty$. If $k=0$ only closed contours are obtained.

Application of the area rule is straightforward in this case, and from it we find

$$(2n+1)\pi\varepsilon=\frac{4}{\sqrt{\lambda_n}}\left[-\sqrt{1-\lambda_n}+\frac{1-k}{\sqrt{\lambda_n-k}}\tan^{-1}\frac{\sqrt{1-\lambda_n}}{\sqrt{\lambda_n-k}}\right].$$

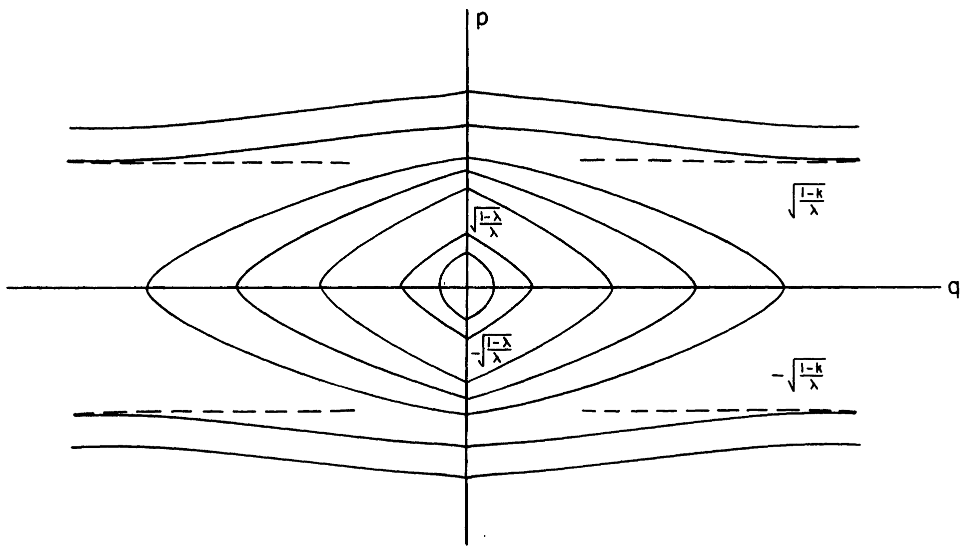


FIG. 2. *$\lambda$-curves for $\tilde{K} = (1+k|q|)/(1+|q|)(1+p^2)$. Curves meet p-axis at $\pm\sqrt{(1-\lambda)/\lambda}$. Closed contours result when $k<\lambda<1$. Open contours result for $0<\lambda<k$, in which case the curves asymptote at $\pm\sqrt{(1-k)/\lambda}$.*

In the limit $\varepsilon$ fixed, $n \uparrow \infty$, we find

$$\lambda_n \sim k + \frac{4(1-k)^2}{\varepsilon^2(2n+1)^2 k}, \tag{2.20}$$

which clearly shows the infinite point spectrum with an accumulation at $k$ in addition to the continuous spectrum (2.18). For later purposes we also give the expression for the eigenvalues which lie close to unity ($n$ fixed and $\varepsilon \downarrow 0$)

$$\lambda_n \sim 1 - [\tfrac{3}{8}(2n+1)\pi\varepsilon(1-k)]^{2/3}. \tag{2.21}$$

This is the set which includes the principal eigenvalues which need more careful examination, as we will see.

*Difference of absolute exponentials.* A model which gives a more realistic portrayal of retinal organization involves the difference of kernels of the form (2.1),

$$K\{x, y\} = \frac{A_1(q)}{2l_1(q)} \exp\left[-\frac{|x-y|}{l_1(q)}\right] - \frac{A_2(q)}{2l_2(q)} \exp\left[-\frac{|x-y|}{l_2(q)}\right] \tag{2.22}$$

which under the Wigner transformation becomes

$$\tilde{K} = \frac{(A_1 - A_2) + p^2(A_1 l_2^2 - A_2 l_1^2)}{(1+p^2 l_1^2)(1+p^2 l_2^2)}. \tag{2.23}$$

Qualitatively, the kernel $K\{x, y\}$ should show a central region of excitation followed by a broad surround of inhibition. To achieve this we must therefore have

$$\frac{A_1}{l_1} > \frac{A_2}{l_2}, \qquad l_2 > l_1.$$

A broad range of possible $\lambda$-curves exist for (2.23). For simplicity we isolate some of these by choosing

$$A_1 = A = \frac{A_2}{k}, \qquad l_1 = \frac{l_2}{c} = 1$$

with $A$ monotonically decreasing from the origin and the constants $c$ and $k$ such that $c>1$, $c>k$. Under these conditions we obtain

$$\tilde{K}=A(q)\frac{(1-k)+p^2(c^2-k)}{(1+p^2)(1+c^2p^2)}. \tag{2.24}$$

For large values of $p$ or $q$ the picture obtained resembles the cases already considered. In particular, if $A(q)$ is bounded away from zero the spectrum is part continuous. The main new features of the $\lambda$-curves of (2.24) are determined by the behavior of $\tilde{K}$ in the neighborhood of the origin. We can assume

$$A(q)\sim 1-q^2$$

in the neighborhood of $q=0$, which inserted into the expansion of (2.24) yields

$$\tilde{K}\sim(1-k)\left\{1-\frac{1-kc^2}{1-k}p^2-q^2\right\}. \tag{2.25}$$

Thus if $1>kc^2$ the $\lambda$-curves in the neighborhood of the origin have the form of ellipses, while if the inequality is in the other direction we obtain hyperbolas locally. A global sketch in the first instance is shown in Fig. 3, while Fig. 4 indicates the $\lambda$-curves for the second case. We also indicate in Fig. 5 the case $k=1$ (which is also of physiological interest).

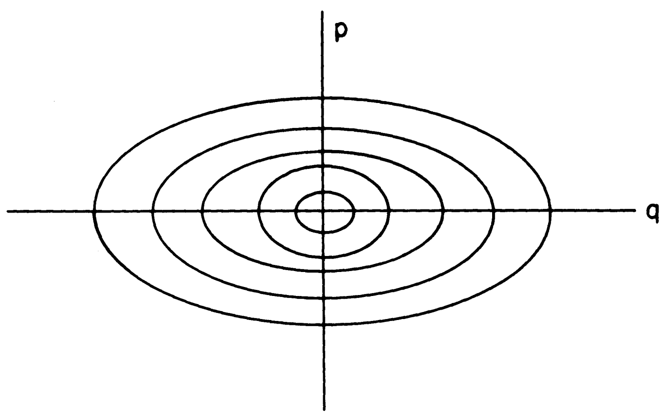


FIG. 3. *$\lambda$-curves for $\tilde{K}(p,q)$ as given by* (2.24), *when* $1>kc^2$.

There are several features to note in regard to Fig. 4. We observe that the loops contained within the figure eight occur in pairs one in the upper and the other in the lower half plane. The area rule implies a degeneracy in the eigenvalues since the corresponding eigenfunctions occur in pairs. (This is also true for all eigenvalues obtained in cases corresponding to Fig. 5.) Further, it is clear that for the figure eight $\lambda$-curve itself $\tilde{K}_p=0$ at the origin and therefore neighboring $\lambda$-curves will have correspondingly small values of $\tilde{K}_p$ near the origin. Thus, as inspection of the WKB solutions (2.10) shows, this indicates a nonuniformity in the procedure which therefore requires special treatment (see § 5).

Although the above features have appeared in the context of a special form, (2.24), the presence of the figure eight and the corresponding eigenvalue degeneracy follows directly from the appearance of two hills (or two valleys)—and therefore a saddle—in the topography of the surface $\tilde{K}(p,q)$. Thus we see that qualitative features

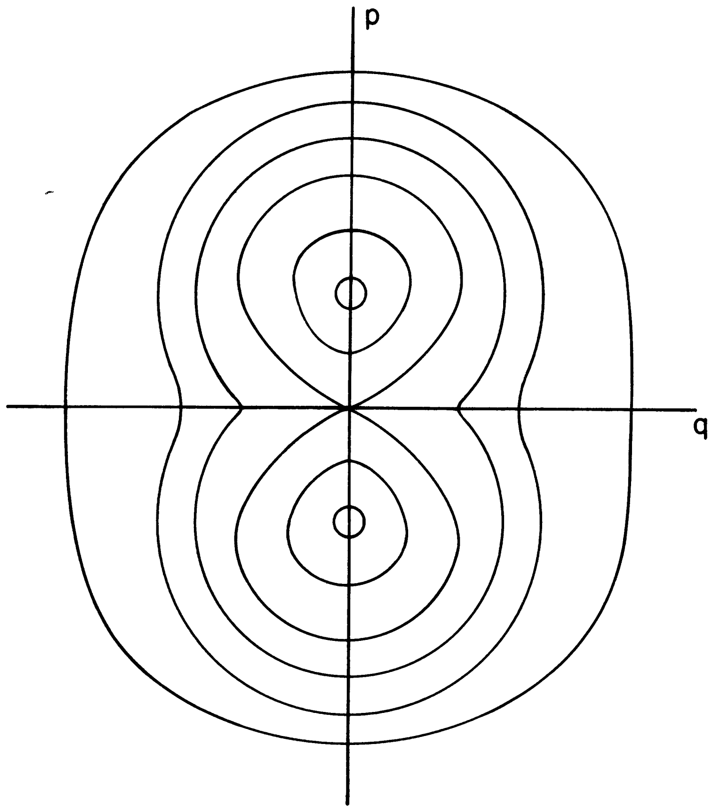


FIG. 4. *λ-curves for* $\tilde{K}(p, q)$ *as given by* (2.24) *when* $1 < kc^2$.

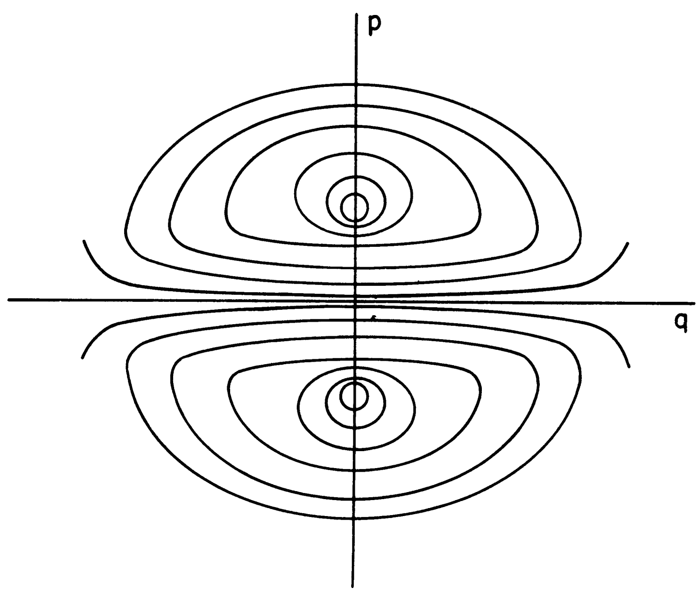


FIG. 5. *λ-curves for* $\tilde{K}(p, q)$ *as given by* (2.24) *when* $kc^2 = 1$.

of the spectrum easily emerge from the topography of a $\tilde{K}(p, q)$ surface, an observation which remains true in more elaborate situations.

*Model with two foveae.* In our discussion of (2.2) (and (2.24)) we restricted attention to coefficients $A/l$ which had a single maximum. For a real retina this corresponds to the region of the fovea where photoreceptors are most densely packed. Because of an interesting connection with the case just considered we now discuss the less common situation of two foveae.[3] To model this case we continue to use (2.2) but permit $A(q)$ to have more than one maximum. A specific example is given by

$$A=\frac{1}{1+(q^2-b^2)^2}, \qquad l=\frac{1}{\sqrt{1+(q^2-b^2)^2}},$$

for which we can say that the foveae are centered at $q=\pm b$. The corresponding form of $\tilde{K}$ is

$$\tilde{K}=\frac{1}{1+(q^2-b^2)^2+p^2}, \tag{2.26}$$

and the corresponding $\lambda$-curves are sketched in Fig. 6.

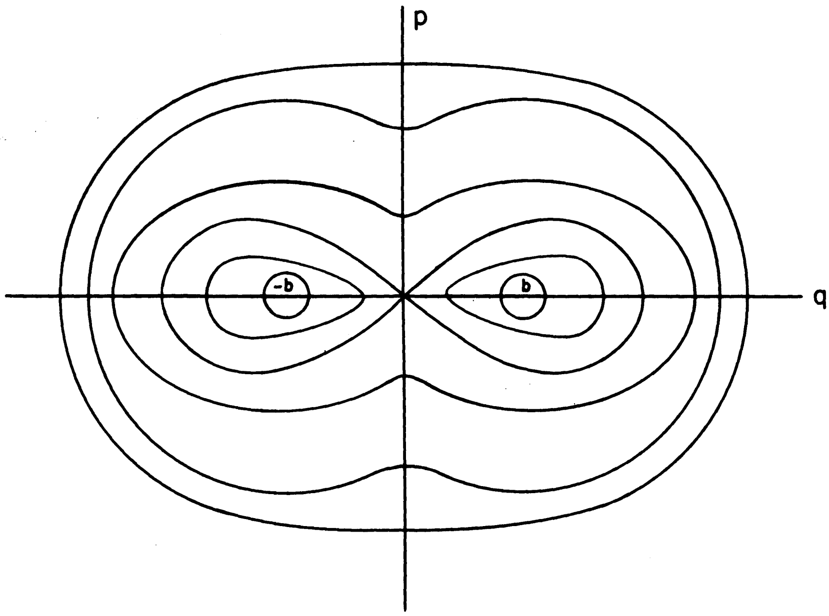


FIG. 6. *$\lambda$-curves for a model with two foveae*; $\tilde{K}=1/(1+(q^2-b^2)^2+p^2)$.

The sketch is qualitatively similar to that of Fig. 4 but rotated through ninety degrees. The topography of the surface (2.26) exhibits two peaks located at $(p, q)=(0, \pm b)$, with a saddle lying between them at the origin. Eigenvalue degeneracy is again present in this case as is the nonuniformity in the neighborhood of the origin. In §§ 4 and 5 we analyze these two features in the context of the $\lambda$-curves, as shown in Fig. 6. This discussion will also serve for the case described by Fig. 4. A fuller explanation of this point is given later.

[3] Two foveae are found in a number of birds, amphibians, and fish.

**3. Principal eigenvalues and eigenfunctions.** In [1] we indicated a failure in the argument leading to the area rule for the case of small index $n$, i.e. for the principal eigenvalues. In such cases the area enclosed by a $\lambda$-curve is relatively small and in terms of the WKB solutions (2.10), the phase and amplitude factors approach one another in order of magnitude. The more sensitive analyses given in [1] indicates that the area rule remains valid for the principal eigenvalues but that the principal eigenfunctions must be modified. For purposes of comparison we briefly summarize these results. For a coordinate system with origin at the center of the principal $\lambda$-curves, such that

$$\frac{\partial \tilde{K}^0}{\partial p} = 0 = \frac{\partial \tilde{K}^0}{\partial q}$$

(zero superscript indicates evaluation at origin) and

$$D^2 = \tilde{K}^0_{pp}\tilde{K}^0_{qq} - (\tilde{K}^0_{qp})^2 > 0,$$

then

$$\lambda_n \sim \tilde{K}^0 + \frac{\varepsilon D}{2}(2n+1) + O(\varepsilon^2),$$

as would be predicted by the area rule. In addition,

$$\psi_n \sim h_n(x\sqrt{\varepsilon D/\tilde{K}^0_{pp}}) \exp\left[-\frac{i\varepsilon x^2 \tilde{K}^0_{pq}}{2D}\right],$$

where $h_n$ is the Hermite function of order $n$ (see [1, (58)]).

The underlying reason for the validity of the area rule is that $\lambda$-curves for small $p$ and $q$ are fit by ellipses. For the case considered earlier, (2.16), it is clear from Fig. 2 that this is not true. In fact we see that $\tilde{K}_q(p, 0)$ does not exist and it is therefore necessary to examine this case further.

If we insert (2.21) into (2.19) we obtain

$$q_0 \sim \frac{1}{1-k}\left[\frac{3}{8}(2n+1)\pi\varepsilon(1-k)\right]^{2/3} \tag{3.1}$$

so that $q = O(\varepsilon^{2/3})$. This implies that a new scaling is required for the argument of the principal eigenfunctions, namely that $\psi$ is a function of $q/\varepsilon^{2/3}$ or equivalently of $x\varepsilon^{1/3}$. Thus to consider the principal eigenvalues we return to the eigenfunction problem (1.1) in the form

$$\int K\{x, y\}\psi(\varepsilon^{1/3}y)\, dy = \lambda\psi(\varepsilon^{1/3}x). \tag{3.2}$$

In the present instance the kernel is given by (2.16). If this is substituted and as well the variable change

$$\varepsilon^{1/3} = \delta, \quad x\delta = r, \quad y\delta = s$$

is made, we find

$$\int \frac{1}{2\delta} \frac{1 + k\delta^2\left|\frac{r+s}{2}\right|}{1+\delta^2\left|\frac{r+s}{2}\right|} \exp\left[-\frac{|r-s|}{\delta}\right]\psi(s)\, ds = \lambda\psi(r).$$

To solve we expand in powers of $\delta$, e.g.

$$\psi(s)=\psi(r)+\delta\left(\frac{s-r}{\delta}\right)\psi'(r)+\frac{\delta^2}{2}\left(\frac{s-r}{\delta}\right)^2\psi''(r)+\cdots.$$

The arrangement of coefficients is implied by the peaking form of the exponential in the integrand. At the lowest order we obtain

$$\frac{d^2\psi}{dr^2}-(1-k)|r|\psi=-\frac{1-\lambda}{\delta^{2/3}}\psi. \tag{3.3}$$

Under the further transformation

$$\rho=(1-k)^{1/3}r, \qquad \alpha=\frac{1-\lambda}{\delta^2(1-k)^{2/3}} \tag{3.4}$$

this becomes

$$\frac{d^2\psi}{d\rho^2}-|\rho|\psi=-\alpha\psi. \tag{3.5}$$

The eigentheory for this operator is as follows:

*even*: $\psi=\mathrm{Ai}\,(|\rho|-\mu_n)$,

$\mathrm{Ai}'\,(-\mu_n)=0$,

$\alpha_n=\mu_n$;

*odd*: $\psi=(\operatorname{sgn}\rho)\,\mathrm{Ai}\,(|\rho|-\nu_n)$,

$\mathrm{Ai}\,(-\nu_n)=0$,

$\alpha_n=\nu_n$.

Here Ai $(x)$ represents the Airy function [14] and Ai$'$ $(x)$ its derivative. The eigenvalues $\alpha_n$ are therefore the collection of zeros of Ai and Ai$'$, both of which are tabulated [14]. If we denote the exact eigenvalues of (3.2) by $\Lambda_n$, then the right-hand side of (3.3) implies

$$\Lambda_n\sim 1-\alpha_n(1-k)^{2/3}\varepsilon^{2/3},$$

which is in contrast to the area rule calculation (2.21), which can be written

$$\lambda_n\sim 1-\tilde{\alpha}_n(1-k)^{2/3}\varepsilon^{2/3},$$
$$\tilde{\alpha}_n=[\tfrac{3}{8}(2n+1)\pi]^{2/3}.$$

In Table 1 we compare $\alpha_n$ and $\tilde{\alpha}_n$. From this table we see that except for the first coefficient—which is off by 9.5%—all coefficients as calculated by the area rule are off by less than 1%.

TABLE 1

| $n$ | $\tilde{\alpha}_n$ | $\alpha_n$ |
|---|---|---|
| 0 | 1.11546 | 1.01879 |
| 1 | 2.32025 | 2.33811 |
| 2 | 3.26162 | 3.24820 |
| 3 | 4.08181 | 4.08795 |
| 4 | 4.82632 | 4.82010 |
| 5 | 5.51716 | 5.52056 |

**4. Splitting of degenerate eigenvalues.** We now present a somewhat heuristic analysis which indicates the splitting of degenerate eigenvalues and hence the removal of degeneracy. The approximate derivation of multiple eigenvalues, as indicated by the situations depicted in Figs. 4 and 6, appears frequently in quantum mechanical problems, as does their "splitting" under a more sensitive treatment. A "splitting" calculation like ours below appears, for example, in [17, problem 7, p. 129].

In the interest of simplicity we consider the case when there is symmetry across both the $p$ and $q$ axes. Also we will work with the case as depicted in Fig. 6 rather than in Fig. 4. There is no loss of generality in this restriction since it is an exact result that eigenvalues are invariant under unimodular transformations of the $(p, q)$-plane [7]. In particular, Fig. 6 and its ninety degree rotation give rise to the same eigenvalue structure. Transformation of the eigenfunctions under this group is also made explicit in [7].

The situation we consider is depicted in Fig. 7. According to the area rule we find a proper $\lambda$-curve in the right half plane and its mirror image in the left half plane.

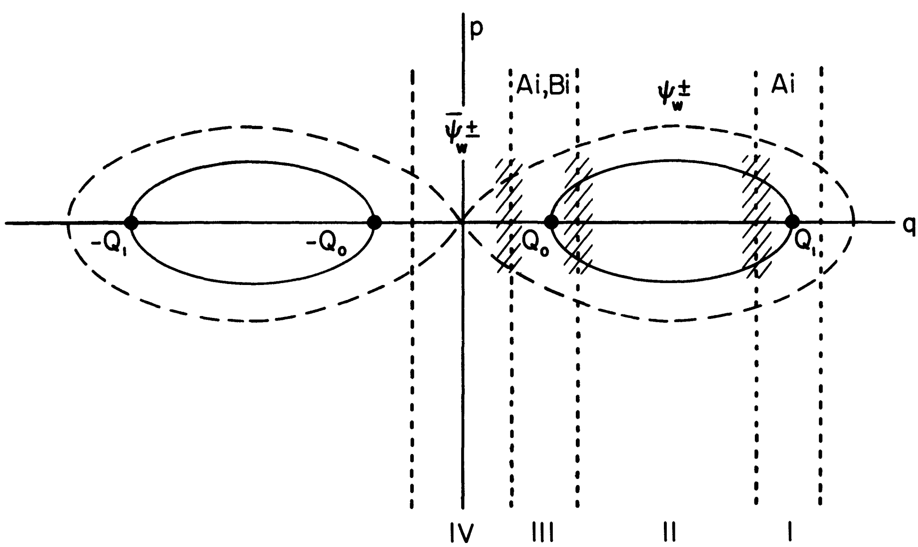


FIG. 7. *Sketch describing the splitting of degenerate eigenvalues. Eigenfunctions in each of the four regions are described by the indicated functions (see text). Dashed lines indicate overlap regions.*

Turning points occur at $\pm Q_0$ and $\pm Q_1$, while the origin is a saddle point. Since the presumed eigenvalue is degenerate we may join the eigenfunction pieces in the two half planes as an odd or even function. This we now examine more closely.

We express the WKB solutions with support in $(Q_0, Q_1)$ by

$$\psi_w^\pm = \frac{\exp\left[\pm(i/\varepsilon)\int_{Q_0}^q p(s)\,ds\right]}{|\tilde{K}_p(p(q), q)|^{1/2}}, \tag{4.1}$$

where $p(q)$ is the positive branch of $\tilde{K}(p, q) = \lambda$. We also need the WKB solutions having support in $(-Q_0, Q_0)$, and write these as

$$\bar{\psi}_w^\pm = \frac{\exp\left[\pm(1/\varepsilon)\int_0^q \bar{p}(s)\,ds\right]}{|\tilde{K}_p(\bar{p}(q), q)|^{1/2}}. \tag{4.2}$$

In effect, $\bar{p} = ip$, and is real positive in this interval. Consideration of the WKB solutions in the left half plane will be avoided by means of symmetry. In addition, we will need to specify the solutions in the neighborhoods of the turning points $Q_0$ and $Q_1$.

From our deliberations in [1] we know that the appropriate variable in the neighborhood of $Q_1$ is

$$z_1 = -\left(\frac{2\tilde{K}_q(0, Q_1)}{\tilde{K}_{pp}(0, Q_1)}\right)^{1/3}\left(\frac{Q_1 - q}{\varepsilon^{2/3}}\right), \tag{4.3}$$

and that the solution in this neighborhood is Ai $(z_1)$, where Ai is the Airy function [14]. Similarly, the solution in the neighborhood of $Q_0$ is given in terms of the variable

$$z_0 = -\left(\frac{2\tilde{K}_q(0, Q_0)}{\tilde{K}_{pp}(0, Q_0)}\right)^{1/3}\left(\frac{Q_0 - q}{\varepsilon^{2/3}}\right), \tag{4.4}$$

but now is a linear combination of Ai $(z_0)$ and Bi $(z_0)$, where the latter represents the second solution of the Airy differential equation [14].

In each of the four regions of Fig. 7 a different functional form of the eigenfunction is valid. In region II we write

$$\psi \sim A\psi_w^+ + B\psi_w^-, \tag{4.5}$$

while in region IV

$$\psi \sim C(\bar{\psi}_w^+ \pm \bar{\psi}_w^-), \tag{4.6}$$

so that an eigenfunction crosses the origin as either an even or odd function. In region I the solution is proportional to Ai $(z_1)$. If this and (4.5) are expanded in the overlap region between I and II, we find as in [1] that

$$\frac{A}{B} = \exp\left[i\left(\frac{\pi}{2} - \frac{\mathscr{A}}{\varepsilon}\right)\right], \tag{4.7}$$

where $\mathscr{A}$ is the area contained in the $\lambda$-curve. Next we turn to the overlap region of regions III and IV. For $q$ near $Q_0$ we find that (4.5) has the form

$$\psi \sim \frac{C}{|\tilde{K}_p(\bar{p}(q), q)|^{1/2}}\left\{\exp\left[\frac{P}{\varepsilon} - \frac{2}{3}z_0^{3/2}\right] \pm \exp\left[-\frac{P}{\varepsilon} + \frac{2}{3}z_0^{3/2}\right]\right\}, \tag{4.8}$$

where

$$P = \int_0^{Q_0} \bar{p}(s)\, ds. \tag{4.9}$$

On the other hand, in the same region [14]

$$\text{Ai}\,(z_0) \sim \frac{1}{2\sqrt{\pi}z_0^{1/2}}\exp\left[-\frac{2}{3}z_0^{3/2}\right], \tag{4.10}$$

$$\text{Bi}\,(z_0) \sim \frac{1}{\sqrt{\pi}z_0^{1/2}}\exp\left[+\frac{2}{3}z_0^{3/2}\right]. \tag{4.11}$$

Comparison of (4.10) and (4.11) with (4.8) shows that the solution in region III has the form

$$\psi \sim D\left\{\exp\left(\frac{P}{\varepsilon}\right)\text{Ai}\,(z_0) \pm \frac{\exp(-P/\varepsilon)}{2}\text{Bi}\,(z_0)\right\}. \tag{4.12}$$

It only remains for us to match (4.5) to (4.12) in the overlap of regions II and III.

To this end we first note that in this overlap region [14]

$$\mathrm{Ai}\,(z_0) \sim \frac{1}{\sqrt{\pi}\, z_0^{1/4}} \sin\left(\frac{\pi}{4} + \frac{2}{3} z_0^{3/2}\right), \tag{4.13}$$

$$\mathrm{Bi}\,(z_0) \sim \frac{1}{\sqrt{\pi}\, z_0^{1/4}} \cos\left(\frac{\pi}{4} + \frac{2}{3} z_0^{3/2}\right). \tag{4.14}$$

On the other hand, (4.5) in the overlap region is given by

$$\psi \sim \frac{1}{|K_p(p(q), q)|^{1/4}} \left\{ A \exp\left(i\frac{2}{3} z_0^{3/2}\right) + B \exp\left(-i\frac{2}{3} z_0^{3/2}\right) \right\}. \tag{4.15}$$

If we introduce (4.13) and (4.14) into (4.12) it may be represented by

$$\begin{aligned} \psi \sim \frac{D}{2iz_0^{1/4}\sqrt{\pi}} \Bigg\{ & \left( \exp\left[ i\left(\frac{\pi}{4} + \frac{2}{3} z_0^{3/2}\right)\right] - \exp\left[ -i\left(\frac{\pi}{4} + \frac{2}{3} z_0^{3/2}\right)\right] \right) \exp\left(\frac{P}{\varepsilon}\right) \\ & \pm \frac{i}{2} \left( \exp\left[ i\left(\frac{\pi}{4} + \frac{2}{3} z_0^{3/2}\right)\right] + \exp\left[ -i\left(\frac{\pi}{4} + \frac{2}{3} z_0^{3/2}\right)\right] \right) \exp\left(-\frac{P}{\varepsilon}\right) \Bigg\}. \end{aligned} \tag{4.16}$$

If we write

$$\exp\left(\frac{P}{\varepsilon}\right) + \frac{i}{2} \exp\left(-\frac{P}{\varepsilon}\right) = M e^{i\phi}, \tag{4.17}$$

so that for example

$$\tan \phi = \frac{1}{2} \exp\left(-\frac{2P}{\varepsilon}\right), \tag{4.18}$$

then (4.12) may be written

$$\psi \sim \frac{DM}{2iz_0^{1/4}\sqrt{\pi}} \left\{ \exp\left[ i\left(\frac{2}{3} z_0^{3/2} + \frac{\pi}{4} \pm \phi\right)\right] - \exp\left[ -i\left(\frac{2}{3} z_0^{3/2} + \frac{\pi}{4} \pm \phi\right)\right] \right\}. \tag{4.19}$$

Then if we compare the ratio of exponentials in (4.15) and in (4.19) we obtain

$$\frac{A}{B} = -\exp\left[ i\left(\frac{\pi}{2} \pm 2\phi\right)\right] \tag{4.20}$$

which in combination with the earlier result (4.7) yields

$$\exp\left[ i\left(\pm 2\phi + \frac{\mathscr{A}}{\varepsilon}\right)\right] = -1$$

or

$$\pm 2\phi + \frac{\mathscr{A}}{\varepsilon} = (2n+1)\pi. \tag{4.21}$$

To bring this result into contact with the area rule, we observe that for $\varepsilon \downarrow 0$ (4.18) implies

$$\phi \sim \frac{1}{2} \exp\left(-\frac{2P}{\varepsilon}\right),$$

which on substitution gives

(4.22) $$\mathcal{A} \sim (2n+1)\pi\varepsilon \mp \varepsilon \exp\left(-\frac{2P}{\varepsilon}\right).$$

As a result of this calculation we see that the eigenvalue degeneracy is broken. For the integral operators we have been considering we also see that the larger eigenvalue goes with the even eigenfunction.

In a sense (4.22) is misleading, since the splitting term is not the next order in a systematic calculation. Higher algebraic orders have been neglected in the area rule development. Such terms however would give no indication of the splitting, i.e., they preserve the degeneracy. This will also be brought out by the analysis of the next section.

**5. Saddle point analysis.** While the discussion of the previous section indicates the removal of eigenvalue degeneracy, it does not cover those cases when the $\lambda$-curve passes close to the saddle point. As already noted, in these cases $\tilde{K}_p$ in the neighborhood of the origin becomes small and the development of the WKB solutions (2.10) can no longer be regarded as self consistent. In Fig. 8 we indicate the two cases which must be further examined: when the two closed loops lie close to the figure eight;

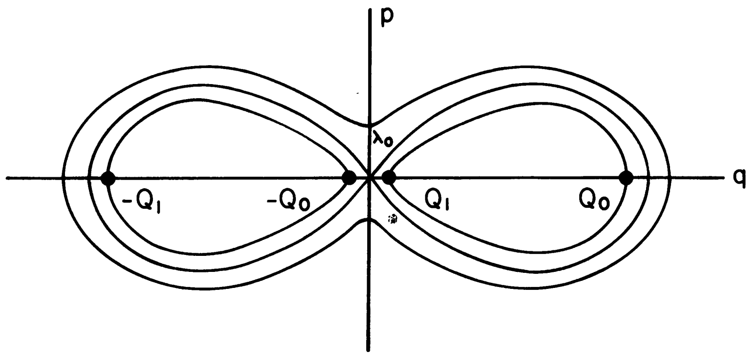


FIG. 8. *Sketch describing behavior in neighborhood of the figure eight curve $\lambda_0$. The origin is a saddle point.*

and when the single closed loop outside the figure eight lies close to it. In the latter case no degeneracy is apparent from the area rule.

As in the previous section we will consider a fully symmetric case. Symmetry is unnecessary for these deliberations; however, we wish to avoid the tedious calculations involved in dealing with the general case.

We begin the analysis with a discussion of $\tilde{K}(p,q)$ in the neighborhood of the saddle at the origin. A Taylor series yields

(5.1) $$\tilde{K} \sim \lambda_0 + \frac{p^2}{2}\tilde{K}^0_{pp} + \frac{q^2}{2}\tilde{K}^0_{qq},$$

where the zero superscript indicates evaluation at the origin. Also we have used $\tilde{K}(p,q) = \lambda_0$ for the figure eight curve itself as indicated in Fig. 8. Since the origin is a saddle, $\tilde{K}^0_{qq}/\tilde{K}^0_{pp} < 0$. The curves in the neighborhood of the origin are described by

(5.2) $$p^2 - q^2\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right| = \frac{2(\lambda-\lambda_0)}{\tilde{K}^0_{pp}}.$$

(For operators of the sort we have been considering $\tilde{K}^0_{pp}<0$; however, this is unnecessary in what follows.)

On formal grounds we take

$$\lambda-\lambda_0=O(\varepsilon). \tag{5.3}$$

Then for $2(\lambda-\lambda_0)/\tilde{K}^0_{pp}<0$ (5.2) defines the turning point

$$Q_0=\sqrt{(2(\lambda-\lambda_0)/\tilde{K}^0_{qq})}, \tag{5.4}$$

which also implies $Q_0=O(\sqrt{\varepsilon})$. At this stage it is appropriate to return to the integral equation (1.1) and reexamine it under the present scaling. From the above we then consider

$$\int K\{x,y\}\psi(\sqrt{\varepsilon}\,y)\,dy=\left[\lambda_0+\varepsilon\left(\frac{\lambda-\lambda_0}{\varepsilon}\right)\right]\psi(\sqrt{\varepsilon}\,x)$$

and can proceed by a formal perturbation analysis as we did in § 3. We avoid this course with the observation that (5.2) is the $(p,q)$-plane image of the differential equation which would result from the perturbation analysis. This is

$$\left(\frac{d^2}{dx^2}+\varepsilon^2x^2\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|\right)\psi(\sqrt{\varepsilon}\,x)=\frac{2(\lambda_0-\lambda)}{\tilde{K}^0_{pp}}\psi(\sqrt{\varepsilon}\,x), \tag{5.5}$$

where we have introduced the already mentioned slow dependence. Under the variable change

$$z=\sqrt{2\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/4}x,\qquad a=\frac{\lambda_0-\lambda}{\varepsilon\tilde{K}^0_{pp}}\left|\frac{\tilde{K}^0_{pp}}{\tilde{K}^0_{qq}}\right|^{1/2} \tag{5.6}$$

this becomes

$$\left(\frac{d^2}{dz^2}+\frac{1}{4}z^2\right)\psi=a\psi, \tag{5.7}$$

which is one of the standard forms of the differential equation for parabolic cylinder functions [15]. Standard solutions to (5.7) are denoted by

$$W_1=W(a,z),\qquad W_2=W(a,-z). \tag{5.8}$$

In what follows we match solutions based on (5.8) to WKB solutions, and we will therefore require the following asymptotic forms [15]:

$$\begin{aligned} W_1&\sim\sqrt{\frac{2k}{z}}\cos\left(\frac{z^2}{4}-a\ln z+\frac{\pi}{4}+\frac{\phi_2}{2}\right),\\ W_2&\sim\sqrt{\frac{2}{kz}}\sin\left(\frac{z^2}{4}-a\ln z+\frac{\pi}{4}+\frac{\phi_2}{2}\right)\end{aligned} \tag{5.9}$$

for $z\uparrow\infty$. In these

$$\phi_2=\arg\Gamma(\tfrac{1}{2}+ia),\qquad k=(1+\exp 2\pi a)^{1/2}-\exp\pi a \tag{5.10}$$

and $\Gamma$ is the gamma function.

In view of the symmetry which we have imposed we can assume that the solutions are either odd or even at the origin,

$$\psi^{\pm}\propto W_1\pm W_2. \tag{5.11}$$

The constant of proportionality, $C$, is introduced through

$$\psi^{\pm} \sim \frac{2C}{\sqrt{q}}\left[\sqrt{k}\cos\left(\frac{z^2}{4} - a\ln z + \frac{\pi}{4} + \frac{\phi_2}{2}\right) \pm \frac{1}{\sqrt{k}}\sin\left(\frac{z^2}{4} - a\ln z + \frac{\pi}{4} + \frac{\phi_2}{2}\right)\right]. \tag{5.12}$$

Next we define

$$\sqrt{k} + i\sqrt{\frac{1}{k}} = \left|k + \frac{1}{k}\right|^{1/2} \exp(i \mp \phi_3), \tag{5.13}$$

so (5.12) can be written

$$\begin{aligned}\psi^{\pm} \sim \frac{C}{\sqrt{q}}\left|k + \frac{1}{k}\right|^{1/2}\Big\{&\exp\left[i\left(\frac{z^2}{4} - a\ln z + \frac{\pi}{4} + \frac{\phi_2}{2} \mp \phi_3\right)\right] \\ &+ \exp\left[-i\left(\frac{z^2}{4} - a\ln z + \frac{\pi}{4} + \frac{\phi_2}{2} \mp \phi_3\right)\right]\Big\}.\end{aligned} \tag{5.14}$$

To proceed further we distinguish between the two cases $a \gtrless 0$.

*Case* 1. $a > 0$. In this case $Q_0$ as given by (5.4) is a turning point, and we write the WKB solution as

$$\psi_w \sim \frac{1}{|K_p(p(q), q)|^{1/2}}\left\{A \exp\left[\frac{i}{\varepsilon}\int_{Q_0}^{q} p(s)\,ds\right] + B\exp\left[-\frac{i}{\varepsilon}\int_{Q_0}^{q} p(s)\,ds\right]\right\}. \tag{5.15}$$

The match of this form to Ai in the neighborhood of $Q_1$, Fig. 8, leads as before to (4.7). In order to accomplish the match in the neighborhood of $Q_0$, $q > Q_0$, we note from (5.2)

$$p \sim \left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\sqrt{q^2 - Q_0^2}, \tag{5.16}$$

and consider the integral

$$\begin{aligned}\frac{1}{\varepsilon}\int_{Q_0}^{q} p(s)\,ds &\sim \frac{1}{\varepsilon}\int_{Q_0}^{q}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\sqrt{s^2 - Q_0^2}\,ds \\ &= \frac{1}{\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\frac{1}{2}\left[q\sqrt{q^2 - Q_0^2} - Q_0^2\ln\frac{q + \sqrt{q^2 - Q_0^2}}{Q_0}\right] \\ &\sim \left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\frac{1}{2\varepsilon}\left[q^2 - Q_0^2\ln\frac{2q}{Q_0}\right] - \frac{a}{2}.\end{aligned} \tag{5.17}$$

In order to obtain the last expression we have taken $q/Q_0 \gg 1$, and made use of

$$a = \frac{Q_0^2}{2\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}, \tag{5.18}$$

which follows from (5.4) and (5.6). If this and a like expansion of $\tilde{K}_p(p(q), q)$ are inserted into (5.15), we obtain

$$\begin{aligned}\psi_w \sim \frac{1}{\sqrt{q|K^0_{qq}/K^0_{pp}|}}\Big\{&A\exp\left[\frac{i}{\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\frac{1}{2}\left(q^2 - Q_0^2\ln\frac{2q}{Q_0}\right) - i\frac{a}{2}\right] \\ &+ B\exp\left[-\frac{i}{\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\frac{1}{2}\left(q^2 - Q_0^2\ln\frac{2q}{Q_0}\right) + i\frac{a}{2}\right]\Big\}.\end{aligned} \tag{5.19}$$

In order to match this to (5.14) we note that from (5.6)

$$\frac{z}{\sqrt{2}}=\frac{q}{\sqrt{\varepsilon}}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/4}, \tag{5.20}$$

from which we can write

$$\frac{z^2}{4}-a\ln z=\frac{1}{2\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\left(q^2-Q_0^2\ln\frac{2q}{Q_0}\right)-\frac{a}{2}\ln a. \tag{5.21}$$

This permits us to rewrite (5.13) in the form

$$\begin{aligned}\psi^{\pm}\sim\frac{C}{\sqrt{q}}\left|k+\frac{1}{k}\right|^{1/2}\Bigg\{&\exp\left[\frac{i}{2\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\left(q^2-Q_0^2\ln\frac{2q}{Q_0}\right)+i\left(\frac{\pi}{4}+\frac{\phi_2}{2}\mp\phi_3-\frac{a}{2}\ln a\right)\right]\\ &+\exp\left[-\frac{i}{2\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\left(q^2-Q_0^2\ln\frac{2q}{Q_0}\right)-i\left(\frac{\pi}{4}+\frac{\phi_2}{2}\mp\phi_3-\frac{a}{2}\ln a\right)\right]\Bigg\}.\end{aligned} \tag{5.22}$$

The forms (5.19) and (5.22) are now easily matched. Since the ratio of the two sinusoidal dependences is invariant, we have

$$\frac{A}{B}\exp[-ia]=\exp\left[i\left(\phi_2\mp2\phi_3-a\ln a+\frac{\pi}{2}\right)\right]. \tag{5.23}$$

On eliminating $A/B$ from this by means of (4.7), we obtain

$$\exp\left[i\left(-\frac{\mathscr{A}}{\varepsilon}-a\right)\right]=\exp[i(\phi_2\mp2\phi_3-a\ln a)],$$

which implies

$$\frac{\mathscr{A}}{\varepsilon}+\phi_2\mp2\phi_3+a-a\ln a=2\begin{bmatrix}n\\ n+1\end{bmatrix}\pi,$$

or

$$\frac{\mathscr{A}}{\varepsilon}=a\ln a-a-\phi_2+(2n+1)\pi\mp(\pi-2\phi_3). \tag{5.24}$$

This is a nonlinear equation for the eigenvalue $\lambda_n$, which enters through $a$, $\phi_2$, $\phi_3$ as well as the area $\mathscr{A}$ of the $\lambda$-curve. In addition it is required that the indexing integer $n$ be such that the resulting $\lambda$-curve lie close to the $\lambda_0$-curve in Fig. 8.

To compare (5.24) with the results of the last section as well as with the area rule we consider the limit $a\uparrow\infty$. In this case

$$\begin{aligned}\phi_2&=-a+a\ln|a|+O\left(\frac{1}{a}\right),\\ \phi_3&=\frac{\pi}{2}-\frac{\exp(-\pi a)}{2}+O(\exp(-3\pi a)),\end{aligned} \tag{5.25}$$

both of which follow from simple asymptotic expansions of (5.13) and (5.10). If these are substituted in (5.24), then

$$\frac{\mathscr{A}}{\varepsilon}=(2n+1)\pi\mp\exp(-\pi a)+O\left(\frac{1}{a}\right). \tag{5.26}$$

To identify this with (4.22), observe that in the present limit

$$\frac{2P}{\varepsilon} \sim \frac{2}{\varepsilon} \left| \frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}} \right|^{1/2} \cdot \int_0^{Q_0} (Q_0^2 - q^2)^{1/2} \, dq = \frac{\pi Q_0^2}{2\varepsilon} \left| \frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}} \right|^{1/2} = a\pi.$$

Observe that the term which breaks degeneracy is of exponentially small order, while the error estimate in (5.26) is only algebraically small. It is clear from the derivation that these algebraic terms play no role in breaking the eigenvalue degeneracy, which thus justifies this unusual way of writing (5.26).

We may also consider the limit $a \downarrow 0$. In this case (5.10) and (5.13) become

$$\phi_2 = O(a), \qquad \phi_3 = \frac{3\pi}{8} + O(a),$$

and with these inserted in (5.24) we obtain

$$\frac{\mathscr{A}}{\varepsilon} \sim (2n+1)\pi \mp \frac{\pi}{4}. \tag{5.27}$$

This can be contrasted with (4.21), which under the same limit gives

$$\frac{\mathscr{A}}{\varepsilon} \sim (2n+1)\pi \mp 2 \tan^{-1} \frac{1}{2}. \tag{5.28}$$

Since $\tan^{-1} \frac{1}{2} \approx 26.5°$, the term in question, (5.28) is off by roughly 18%.

*Case* 2. $a < 0$. With only small changes this case follows the previous one, and we present it in outline form. In the neighborhood of the origin the two solutions (5.11) are still valid, as are their asymptotic developments (5.14). The latter are to match to a WKB solution which we now write as

$$\psi_w = \frac{1}{|K_p(p(q), q)|^{1/2}} \left\{ A \exp\left[\frac{i}{\varepsilon} \int_0^q p(s) \, ds\right] + B \exp\left[-\frac{i}{\varepsilon} \int_0^q p(s) \, ds\right] \right\}. \tag{5.29}$$

If we define

$$\hat{Q}_0^2 = \frac{2(\lambda_0 - \lambda)}{\tilde{K}^0_{qq}} = -2a\varepsilon \left| \frac{\tilde{K}^0_{pp}}{\tilde{K}^0_{qq}} \right|^{1/2}, \tag{5.30}$$

then in analogy with (5.17) we can write

$$\frac{1}{\varepsilon} \int_0^q p(s) \, ds \sim \frac{1}{2\varepsilon} \left| \frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}} \right|^{1/2} \left[ q^2 + \hat{Q}_0^2 \ln \frac{2q}{\hat{Q}_0} \right] - \frac{a}{2}. \tag{5.31}$$

This in turn allows us to write, instead of (5.29),

$$\begin{aligned} \psi_w \sim \frac{1}{\sqrt{q |\tilde{K}^0_{qq}/\tilde{K}^0_{pp}|}} \Bigg\{ & A \exp\left[\frac{i}{2\varepsilon} \left| \frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}} \right|^{1/2} \left( q^2 + \hat{Q}_0^2 \ln \frac{2q}{\hat{Q}_0} \right) - \frac{ia}{2} \right] \\ & + B \exp\left[-\frac{i}{2\varepsilon} \left| \frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}} \right|^{1/2} \left( q^2 + \hat{Q}_0^2 \ln \frac{2q}{\hat{Q}_0} \right) + \frac{ia}{2} \right] \Bigg\}. \end{aligned} \tag{5.32}$$

In the present instance we can write

$$\frac{z^2}{4} - a \ln z = \frac{1}{2\varepsilon} \left| \frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}} \right|^{1/2} \left( q^2 + \hat{Q}_0^2 \ln \frac{2q}{\hat{Q}_0} \right) - \frac{a}{2} \ln |a|,$$

which allows us to write (5.13) as

$$\psi^{\pm} \sim \frac{C}{\sqrt{q}}\left|k+\frac{1}{k}\right|^{1/2}\left\{\exp\left[\frac{i}{2\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\left(q^2+\hat{Q}_0^2\ln\frac{2q}{\hat{Q}_0}\right)+i\left(\frac{\pi}{4}+\frac{\phi_2}{2}\mp\phi_3-\frac{a}{2}\ln|a|\right)\right]\right.$$

$$\left.+\exp\left[-\frac{i}{2\varepsilon}\left|\frac{\tilde{K}^0_{qq}}{\tilde{K}^0_{pp}}\right|^{1/2}\left(q^2+\hat{Q}_0^2\ln\frac{2q}{\hat{Q}_0}\right)-i\left(\frac{\pi}{4}+\frac{\phi_2}{2}\mp\phi_3-\frac{a}{2}\ln|a|\right)\right]\right\}. \tag{5.33}$$

Comparison of (5.32) with (5.33) gives

$$\exp\left[i\left(\frac{\pi}{2}+\phi_2\mp 2\phi_3-a\ln|a|\right)\right]=\frac{A}{B}\exp(-ia),$$

and with (4.7) this gives

$$\exp\left[-i\left(\frac{\mathscr{A}}{2\varepsilon}+a\right)\right]=\exp\left[i(\phi_2\mp 2\phi_3-a\ln|a|)\right]. \tag{5.34}$$

Note that $\mathscr{A}/2$ appears since in this case only half the area is being computed in (4.7). Instead of (5.24) we now obtain

$$\frac{\mathscr{A}}{\varepsilon}+2a-2a\ln|a|+2\phi_2\mp 4\phi_3=4n\pi. \tag{5.35}$$

We can now discuss limits of (5.35) comparable to those discussed for (5.24).

$|a|\uparrow\infty$. In this case

$$\phi_2=-a+a\ln|a|+O\left(\frac{1}{a}\right),\qquad \phi_3\sim\frac{\pi}{4},$$

so that (5.35) with a redefinition of $n$ becomes

$$\mathscr{A}\sim(2n+1)\pi\varepsilon, \tag{5.36}$$

which is in agreement with the area rule. In this case no degeneracy was present on the area rule, and the two branches which appear in (5.35) simply correspond to different values of $n$ in (5.36). Odd and even eigenfunctions produce alternating eigenvalues.

$|a|\downarrow 0$. Under this limit, as before,

$$\phi_2=O(a),\qquad \phi_3=\frac{3\pi}{8}+O(a).$$

If this is inserted into (5.36) the result is

$$\frac{\mathscr{A}}{\varepsilon}\sim\left(4n\pi\pm\frac{3\pi}{2}\right). \tag{5.37}$$

It will be informative if we simultaneously discuss the limiting forms (5.27) and (5.37). Both of these apply when the corresponding $\lambda$-curves lie close to the figure eight in Fig. 8, and both of these give the correction to the area rule calculation. In each of these cases the formulas can furnish an invalid area. It may turn out that the upper value of (5.27) is a proper area, in that it is less than a lobe area of the figure eight or $\lambda_0$-curve, but that the lower value (5.27) is not a valid area in this sense. Similarly it is possible for the upper value in (5.37) to be a proper value but the lower not proper. It is of interest to observe that in such instances the results of the two formulas cross over to one another. For example, if the lower value of (5.37) is improper then on heuristic grounds we would expect that one half of the area to be a proper area,

$$\frac{\mathcal{A}}{2\varepsilon} \sim 2n\pi - \frac{3\pi}{4} = (2n-1)\pi + \frac{\pi}{4},$$

which is clearly of the form of (5.27). In a similar fashion we go from the lower value of (5.27) to a lower value of (5.37). Note that the parity of the associated eigenfunctions is maintained in this crossover.

## REFERENCES


[1] L. SIROVICH AND B. W. KNIGHT, *On the eigentheory of operators which exhibit a slow variation*, Quart. Appl. Math., 38 (1981), pp. 469–488.

[2] J. HEADING, *An Introduction to Phase Integral Methods*, Methuen, London, 1962.

[3] E. P. WIGNER, *On the quantum correction for thermodynamic equilibrium*, Phys. Rev., 40 (1932), pp. 749–759.

[4] J. E. MOYAL, *Quantum mechanics as a statistical theory*, Proc. Camb. Phil. Soc., 45 (1949), pp. 99–124.

[5] H. J. GROENEWALD, *On the principles of quantum mechanics*, Physica, 12 (1946), pp. 405–460.

[6] J. BRUER, *The classical limit of quantum theory*, Ph.D. thesis, Rockefeller University, New York, 1978.

[7] B. W. KNIGHT AND L. SIROVICH, *The Wigner transform and some exact properties of linear operators*, this Journal, this issue, pp. 378–389.

[8] J. B. KELLER, *Corrected Bohr–Sommerfeld quantum conditions for nonseparable systems*, Ann. Phys., 4 (1958), pp. 180–188.

[9] J. B. KELLER AND S. I. RUBINOW, *Asymptotic solution of eigenvalue problems*, Ann. Phys., 9 (1960), pp. 24–75.

[10] J. B. KELLER AND D. W. MCLAUGHLIN, *The Feynman integral*, Amer. Math. Monthly, 82 (1975), pp. 451–465.

[11] V. P. MASLOV, *Théorie des perturbations et methods asymptotiques*, Dunod, Paris, 1972.

[12] V. I. ARNOLD, *On a characteristic class entering in the quantization conditions*, Funct. Anal. Appl., 1 (1967), pp. 1–13.

[13] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York, 1978.

[14] M. ABRAMOVITZ AND I. A. STEGUN, *Handbook of Mathematical Functions*, U.S. Government Printing Office, Washington, DC, 1964.

[15] W. MAGNUS, F. OBERHETTINGER AND R. SONI, *Formulas and Theorems for Special Functions of Mathematical Physics*, Springer-Verlag, New York, 1966.

[16] L. SIROVICH, *Techniques of Asymptotic Analysis*, Springer-Verlag, New York, 1971.

[17] G. BAYM, *Lectures on Quantum Mechanics*, Benjamin, New York, 1969.