ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
Classe di Scienze

H. P. MC KEAN
J. C. SCOVEL

Geometry of some simple nonlinear differential operators

Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série, tome 13, nº 2 (1986), p. 299-346

<http://www.numdam.org/item?id=ASNSP_1986_4_13_2_299_0>

Geometry of Some Simple Nonlinear Differential Operators.

H. P. MCKEAN (*) - J. C. SCOVEL (**)

1. – Introduction.

The very simplest nonlinearity is the map $x \to x^2$ folding the real line in half. We study the effect of this nonlinearity when it is combined with the simplest of differential operators $D = d/dx$ and D^2; more precisely, we study the geometry of the maps $A: f \to Df + f^2$ on the space of functions of period 1 and $B: f \to D^2 f + f^2/2$ on the space of functions vanishing at $x = 0$ and $x = 1$. A is a fold, *i.e.* there are coordinates on the domain and on the range, so that A is expressed as $(x_1, x_2, x_3, ...) \to (x_1^2, x_2, x_3, ...)$. B is not so simple: indeed it presents local folds in co-dimension 1, cusps in co-dimension 2, and a whole series of higher singularities, though, being an analytic map, the degree of the singularity is always finite. The number of preimages of a point is finite too. The singular set of B is comprised of sheets $M_n = \{f: \lambda_n(f) = 0\}$ in which $\lambda_1(f) < \lambda_2(f)$, *etc.* is the spectrum of $F = -D^2 + f$ subject to Dirichlet boundary conditions. The first sheet is a convex surface. The others lie one below the other and have each one more principal direction of negative curvature relative to the ambient space. The chief tool is the simple geometrical observation that Bf_1 and Bf_2 coincide if and only if $f = \frac{1}{2}(f_1 + f_2)$ lies on a singular sheet and $e = \frac{1}{2}(f_1 - f_2)$ is proportional to the singular direction at f. That is why this particular map is so tractable. B maps M_1 1:1 onto a convex surface and the region above the latter is both the full range of B and the 1:1 image of the region above M_1. As the image Bf rises, its preimages proliferate; in particular, if f_0 lies below the n-th sheet, then $Bf = Bf_0$ has at least $2n$ solutions. This proliferation can be followed in detail for $Bf = k$ as $k \uparrow \infty$ since the equation can be in-

(*) The work presented in this paper was performed at the Courant Institute of Mathematical Sciences, New York, with the support of the National Science Foundation under Grant NSF-MSC-76-07039.

(**) Much of the work presented consists of a revision of the author's thesis (1983) completad at the Courant Institute of Mathematical Sciences, New York.

Pervenuto alla Redazione il 13 Febbraio 1985.

tegrated explicitly in terms of elliptic functions. This leads to the following estimate: if f_0 is fixed, then the number N of preimages of $Bf_0 + k$ obeys the rule $N \geqslant [C + o(1)]^4\sqrt{k}$ in which C is a certain constant expressible by elliptic integrals. It is our belief that the estimate is sharp for most large values of k, but the proof escapes us.

This study was inspired by a result of Ambrosetti-Prodi [1] and Berger-Church [3]. They proved that if Ω is a domain in R^d $(d \geqslant 2)$, if $\lambda_1 < \lambda_2$ are the first two eigenvalues of $-\Delta$ subject to Dirichlet boundary conditions and if K is a convex function of 1 variable with derived function rising from $-\infty < K'(-\infty) < \lambda_1$ to $\lambda_2 < K'(+\infty) < \infty$, then the map $f \to -\Delta f + K(f)$ is a fold. We decided to see what happens when $K'(f)$ crosses all the energy levels of $-\Delta$ in dimension 1, the map B being the simplest candidate. A good deal of the present paper carries over to $-\Delta f + f^2/2$ and variants of it in $d \geqslant 2$ dimensions. This will be reported upon by Scovel in another publication [9].

2. – Variations on the Riccati operators.

The Ricatti operator (1) $Af = Df + f^2$ provides a nice model problem: it presents *a global fold*, as will be proved in a moment. Let H^0 be the (real) function space $L^2(0, 1)$ with standard inner product and let H^1 be the subspace of H^0 of periodic functions of period 1 with $Df \in H^0$ and norm $\|f\|_1^2 = \int_0^1 |Df|^2 + \int_0^1 |f|^2$. A is a *proper map* from H^1 to H^0, meaning that the inverse map preserves compactness.

PROOF. Let $Af = Df + f^2 = v$ be controlled in H^0. Then $\int_0^1 v^2 \geqslant \int_0^1 |Df|^2 + \int_0^1 f^4$ so that f is controlled in H^1. Since we are on the unit interval f^2 is controlled in H^1, which is compactly imbedded in H^0. The rest is standard.

Singular set and range

The singular set of A is the class M of points $f \in H^1$ where the differential $dA = D + 2f$ has a nontrivial null vector $e \in H^1$; off the singular set the cheap implicite function theorem guarantees that A is a local diffeomorphism, dA being a Fredholm map of index zero and so (boundedly) invertible there.

PROPOSITION 1. *M is the plane* $(f, 1)_1 = \int_0^1 f = 0$.

(1) D signifies differentiation $Df = f'$.

PROOF. $dAe = De + 2fe = 0$ is solved by multiples of $e_1 = \exp\left(-2\int_0^x f\right)$ and periodicity of e is equivalent to $\int_0^1 f = 0$.

PROPOSITION 2. *$A(M)$ is the class of function $v \in H^0$ such that the Hill's operator for $-D^2 + v$ has vanishing lowest periodic eigenvalue; in particular, it is a smooth convex surface dividing H^0 into two connected pieces.*

PROOF. Let $f \in M$ and $e = \exp\left(-2\int_0^x f\right)$ as above, and write $e = w^{-2}$. Then $f = w^{-1}Dw$ and $Af = D(w^{-1}Dw) + (w^{-1}Dw)^2 = w^{-1}D^2w$. This shows that $v = Af$ has groundstate $w > 0$ $[-D^2w + vw = 0]$ so that $-D^2 + v$ has vanishing lowest eigenvalue. The converse is just as easy: if $-D^2 + v$ has lowest eigenvalue 0, then its groundstate w is a positive solution of $-D^2w + vw = 0$, and it is an elementary exercise to check that $f = w^{-1}Dw \in H^1$ satisfies $\int_0^1 f = 0$ and $Af = v$. The rest is common knowledge. Let $\lambda_0(v)$ be the lowest eigenvalue of $-D^2 + v$ so that $A(M) = H^0 \cap [v : \lambda_0(v) = 0]$. The convexity of $A(M)$ is equivalent to the concavity of λ_0: $\lambda_0(\frac{1}{2}v_1 + \frac{1}{2}v_2) > \frac{1}{2}\lambda_0(v_1) + \frac{1}{2}\lambda_0(v_2)$ *unless* $v_1 = v_2$, while the smoothness of $A(M)$ follows from the fact that the gradient of λ_0 does not vanish along $A(M)$: in fact, the gradient is $\nabla\lambda_0(v) = \omega^2\left(\int_0^1 w^4\right)^{-\frac{1}{2}}$; see art. 4 for review of such matters.

PROPOSITION 3. *A is* 1:1 *on M.*

PROOF. Let f_1 and $f_2 \in M$ have the same image. Then $0 = Af_1 - Af_2 = De_0 + (f_1 + f_2)e_0$ with $e_0 = f_1 - f_2$, and either $e_0 \equiv 0$ or else it is of one signature. The latter possibility contradicts $\int_0^1 e_0 = \int_0^1 f_1 - \int_0^1 f_2 = 0$.

PROPOSITION 4. *A maps the half-space above M* 1:1 *onto the region above $A(M)$; in fact, the latter is the whole range of A.*

PROOF. Let $Af = v$. Then, for $w \in H^1$,

$$\int_0^1 [(Dw)^2 + vw^2] = \int_0^1 [(w')^2 + (f' + f^2)w^2]$$

$$= \int_0^1 (w')^2 - 2\int_1^1 fww' + \int_0^1 f^2w^2 = \int_0^1 (w' - wf)^2 > 0\,,$$

and since the minimum of the first expression, subject to $\int_0^1 w^2 = 1$, is the lowest eigenvalue $\lambda_0(v)$, $\lambda_0(Af) \geqslant 0$; moreover, if $\lambda_0(v) = 0$ so that $v \in A(M)$, then $f = w'/w$ at the minimum and $f \in M$, w being root free in its office of ground state. This shows that $A(H^1)$ lies on or above $A(M)$. Now restrict A to M or above: it is *open* above being a local diffeomorphism and also *closed* being proper, so it has to be *onto*. It remains to check that it is 1:1. This proceeds as before: if $Af_- = Af_+ = v$ with f_- and f_+ above M, then $e_0' + 2f_0 e_0 = 0$ with $e_0 = \frac{1}{2}(f_+ - f_-)$ and $f_0 = \frac{1}{2}(f_+ + f_-)$, and either $e_0 \equiv 0$, or else it is a nonvanishing multiple of $\exp\left(-2\int_0^x f_0\right)$ and its periodicity implies that $\int_0^1 f_0 = 0$, forcing one of the f's to lie below M. The proof is finished.

PROPOSITION 5. *The number N of preimages $v \in H^0$ obeys the rule*: $N(v) = 0, 1,$ *or* 2 *according as v lies below, on, or above $A(M)$.*

PROOF. The first count follows from prop. 4, as does the second in view of the fact that $A^{-1}(A(M)) = M$. To confirm the third count, notice that $N(v)$ is constant above $A(M)$, A being a proper, local diffeomorphism there, so that it suffices to compute an example. Let Af be a constant k. Then $0 = \int_0^1 f' Af = \int_0^1 (f')^2 + \int_0^1 f' f^2 = \int_0^1 (f')^2$ so $f' = 0$ and $f = \pm\sqrt{k}$. Now $\lambda_0(k) = k$, and if this number is negative there are no solutions, while if it is 0 there is one, and if it is positive there are 2. The count is finished.

The fold

A is now a prime candidate for a fold: if v lies above $A(M)$, then it has 2 preimages f_+ and f_-; $e_0 = \frac{1}{2}(f_+ - f_-)$ satisfies $e_0' + 2fe_0 = 0$ with $f_0 = \frac{1}{2}(f_+ + f_-) \in M$ and with the right labelling $e_0 > 0$, so that $f_+ = f_0 + e_0$ lies above M, $f_- = f_0 - e_0$ lies below, and $v = f_0' \pm e_0' + f_0^2 \pm 2f_0 e_0' + e_0^2 = f_0' + f_0^2 + e_0^2 = v_0 + e_0^2$ with $v_0 = Af_0 \in A(M)$, as in the figure.

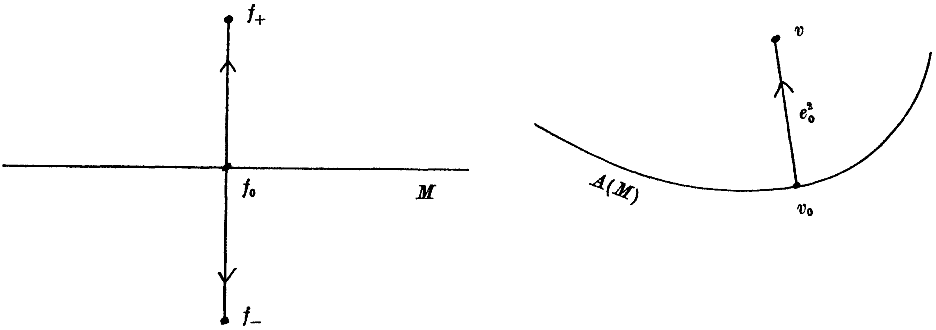


Figure 1

This really looks like a fold of the form $(x_1, x_2, x_3 \ldots) \to (x_1^2, x_2, x_3, \ldots)$: in fact, *it may be reduced to precisely this map by the application of global diffeomorphisms to the right and to the left.*

The proof can be done by general methods but it is nicer to do it by hand; the preceeding computations indicate the route to follow.

STEP 1. Let $e_0 = \exp\left(-2\int_0^x f_0\right)$ for $f_0 \in M$ $[e_0' + 2f_0 e_0 = 0]$; it is to be proved that *the map* $f_0 \to f_0 + ce_0(f_0)$ of $(f_0, c) \in M \times \mathbb{R}$ to H^1 is $1{:}1$ *and onto.*

PROOF. It is required to solve $f = f_0 + ce_0$ for $f_0 \in M$ and $c \in \mathbb{R}$ starting from $f \in H^1$. Let $p = \exp\left(2\int_0^x f\right) = 1/e_0$. Then $2f_0 = p'/p$, so that $p = f_0 + ce_0$ needs $p' = 2fp = -2c$ with $p(0) = 1$, of which the explicit solution is $p = \exp\left(2\int_0^x f\right) - 2c\int_0^x \exp\left(2\int_y^x f\right) dy$, c being determined by the periodicity of p: $1 = p(0) = p(1) = \exp\left(2\int_0^1 f\right) - 2\int_0^1 \exp\left(2\int_x^1 f\right) dx$. Now p is root-free and so positive $[p(0) = 1]$: otherwise, it has two consecutive roots (or a double root) at which its slope $p' = -2c$ is the same, and that is impossible in a periodic function unless $c = 0$, in which case $p = \exp\left(2\int_0^x f\right)$ is positive anyhow. The upshot is that p is of the form $\exp\left(2\int_0^x f_0\right)$ with $f_0 \in H^1$ and $\int_0^1 f_0 = 0$ by the periodicity of p. The proof is finished.

STEP 2. *The map of Step* 1 *is a global diffeomorphism.*

PROOF. This is plain from the computations of Step 1 expressing the inverse map $f \to (f_0, c)$: it is well-defined and smooth.

STEP 3. The left-hand diffeomorphism $M \times \mathbb{R} \to H^1$ just produced is combined with the map A to obtain the diagram

$$(f_0, c) \to f \xrightarrow{A} f' + f^2 = f_0' + f_0^2 + c^2 e_0^2 .$$

Now the map B of $M \times \mathbb{R}$ to H^0 expressed by the rule

$$(f_0, c) \to \begin{cases} f_0' + f_0^2 + c^2 e_0^2 & \text{if } c \geqslant 0 \\ f_0' + f_0^2 + c & \text{if } c < 0 \end{cases}$$

is a global homeomorphism as will be proved in the next steps. The application of the inverse map to the right of the diagram produces the final fold:

$$(f_0, c) \xrightarrow{A} f \to f_0' + f_0^2 + c^2 e_0^2 \xrightarrow{B^{-1}} (f_0, c^2),$$

in which $f_0 \in M$ is fixed and the additional coordinate c is merely squared.

STEP 4. The forward map B is smooth; it sends $M \times 0$, 1:1 onto $A(M)$, $M_+ = M \times (0, \infty)$ onto the region above $A(M)$, and $M_- = M \times (-\infty, 0)$ onto the region below $A(M)$. The main part of the proof is to check that each of the 3 pieces is a global diffeomorphism. The lower part is trivial and the middle piece is easy from what went before, so only the upper piece is dealt with below.

STEP 5. *B is 1:1 on M^+.*

PROOF. B maps $(f_0, c) \subset M \times (0, \infty)$ to $v = v_0 + ce_0^2$ with $v_0 + Af_0 = f_0' + f_0^2$ and $e_0 = \exp\left(-2\int_0^x f_0\right)$; notice that $\omega = \exp\left(\int_0^x f_0\right)$ is the ground state of v_0: $-\omega'' + v_0\omega = 0$. Now if B is not 1:1 on M_+, then neither is the map $(v_0, c) \in A(M) \times (0, \infty) \to v_0 + ce_0^2 = (\omega''/\omega) + (c/\omega^4)$, and as the map $v \to \omega$ is plainly 1:1, $\omega \to \omega^{-1}\omega'' + c\omega^{-4}$ must be the culprit. Let us absorb the number $c > 0$ into ω and look at the simplified map $\omega \to \omega^{-1}\omega'' + \omega^{-4}$ of arbitrary (smooth) positive functions: the old c can be recovered from the new $\omega(0)$. It is to be proved that *this map is injective*: in fact if it sends ω_1 and ω_2 to the same place, then $\omega_1''/\omega_1 + \omega_1^{-4} = \omega_2''/\omega_2 + \omega_2^{-4}$ gives

$$(\omega_1''\omega_2 - \omega_1\omega_2'')(\omega_1 - \omega_2) = (\omega_1\omega_2)^{-1}(\omega_1^4 - \omega_2^4)(\omega_1 - \omega_2) \geqslant 0,$$

and this is not possible unless $\omega_1 \equiv \omega_2$. The proof is easy: ω_2 cannot be everywhere $\geqslant \omega_1$, or vice versa and keep both $v_1 = \omega_1''/\omega_1$ and $v_2 = \omega_2''/\omega_2$ on the surface $A(M)$ where the lowest eigenvalue vanishes, so there is an interval where $\omega_1 > \omega_2$ with equality at the ends. In that interval, $\omega_1'\omega_2 - \omega_1\omega_2'$ increases from its left hand value $(\omega_1)(\omega_1 - \omega_2)' \geqslant 0$ and this makes $(\omega_1/\omega_2)' > 0$ inside, preventing the equality $\omega_1 = \omega_2$ at the right hand end.

STEP 6. *$B|_{M+}$ is closed.*

PROOF. If $v = f_0' + f_0^2 + ce_0^2$ is controlled in H^0, then $\int_0^1 v^2 \geqslant \int_0^1 (f_0')^2 + \int_0^1 f_0^4$ is also controlled in view of

$$\int_0^1 f_0'(f_0^2 + ce_0^2) = 0 - 2c\int_0^1 f_0 e_0 e_0' = 8c\int_0^1 f_0^2 e_0^2 \geqslant 0,$$

the portion of $\int_0^1 v^2$ not accounted for being positive. The rest is routine.

STEP 7. $B|_{M^+}$ *is a local diffeomorphism.*

PROOF. The differential of B acts by the rule

$$(\dot f_0, \dot c) \in TM \times \mathbb{R} \to D\dot f_0 + 2f_0\dot f_0 + \dot c e_0^2 - 4ce_0^2\int_0^x \dot f_0\,;$$

it is a compact perturbation of the simpler map

$$(\dot f_0, \dot c) \in TM \times \mathbb{R} \to D\dot f_0 + 2f_0\dot f_0 + \dot c - 2\int_0^1 f_0\dot f_0\,.$$

Now the first map (and likewise the second) has only the trivial null-space; indeed, with $p = \exp\left(2\int_0^x f_0\right)$, the vanishing of $dB(\dot f_0, \dot c)$ is expressed as $p^{-1}D(p\dot f_0) = 4cp^{-2}\int_0^x \dot f_0 - \dot c p^{-2}$; this is multiplied by p^2 and differentiated to produce $DpD(p\dot f_0) = 4c\dot f_0$; a subsequent multiplication by $p\dot f_0$ and integration from $x = 0$ to $x = 1$ yields $-\int_0^1 p[D(p\dot f_0)]^2 = 4c\int_0^1 p(\dot f_0)^2$, and as the integrals right and left are of opposite signature, they must vanish, forcing $p\dot f_0$ to be constant, and in fact vanish in view of $p > 0$ and $\int_0^1 \dot f_0 = 0$; then $\dot c = 0$, too, as required. It remains to prove that the second simplified map is boundedly invertible from H^0 to $TM \times \mathbb{R}$, which is easy: $h = D\dot f_0 + 2f_0\dot f_0 + \dot c - 2\int_0^1 f_0\dot f_0$ determines $\dot c = \int_0^1 h$ at once; then you can compute $\dot f_0$ in terms of $\dot f_0(0)$, $\int_0^1 f_0\dot f_0$, $\dot c$, and h by elementary integration as in Step 1: $\dot f_0 = e_0\left[\dot f_0(0) - 2\int_0^x e_0^{-1}(y)\,dy\right] +$ known function, and finally determine 1) $\int_0^1 f_0\dot f_0$ by periodicity of $\dot f_0$, and 2) $\dot f_0(0)$ itself by $\int_0^1 \dot f_0 = 0$. Both $\dot f_0$ and $\dot c$ depend boundedly upon h by inspection. The proof is finished.

STEP 8 is merely to collect all the pieces to conclude that B is a diffeomorphism of M^+. The fold is confirmed.

VARIATION 1. The nonlinearity $K_0(f) = f^2$ can be replaced by any strictly convex function K with $K''(f) \geqslant c > 0$ near $f = \pm\infty$. The result is the same: $A \equiv D + K$ *is a global fold.* The case $K(f) = f^4/4 - f$ is typical. The singular set is now the cubic $M: \int_0^1 f^3 = 1$. The proof is similar, but less concrete. See Scovel [8] for details.

VARIATION 2. If K is strictly increasing with $K'(f) \geqslant c_2 > 0$ near $\pm\infty$, or if it is strictly decreasing with $K'(f) \leqslant -c_2 < 0$ near $\pm\infty$, then $A = D + K$ *is a global homeomorphism*; it is even a diffeomorphism where $K'(f)$ does not vanish. The case $K(f) = f^3/3$ is typical. The singular set is now the isolated point $f \equiv 0$.

VARIATION 3. Let $K(f) = f^3 + af^2 + bf$ be such that neither $K'(f) = 3f^2 + 2af + b$ nor $K''(f) = 6f + 2a$ is of one signature; the second stipulation is unnecessary, while the first requires $-\infty < b < a^2/3$. The singular set of $A = D + K$ is the locus M where $\int_0^1 K'(f) = \int_0^1 (3f^2 + 2af + b)$ vanishes which is to say $\int (f + a/3)^2 = \frac{1}{3}(a^2/3 - b) = r^2$ in short, it is a sphere of radius r about $-a/3$ in the H^0 format; in the H^1 format it appears as an ellipsoid. The suggestion is that some kind of ellipsoidal coordinates could be useful in the study of this map, but that has not been carried out.

3. – The operator $-D^2 f + f^2/2$.

The map $B: f \to -f'' + f^2/2$ occupies the rest of the paper; it is defined first for (real) functions $f \in C^2[0, 1]$ with $f(0) = f(1) = 0$. Let $\boldsymbol{G}$ be the Green's operator for $-D^2$ in that class:

$$\boldsymbol{G}f(x) = (1 - x)\int_x^x y f(y)\, dy + x\int_0^1 (1 - y) f(y)\, dy .$$

Then the modified map $A = \boldsymbol{G}B: f \to f + \boldsymbol{G}f^2/2$ extends naturally to a map of H^1 to itself, where H^1 is now the space of (real) functions $f \in C[0, 1]$ with $f(0) = f(1) = 0$ and $\|f\|_1^2 = \int_0^1 |f'|^2 < \infty$. A is preferred to B for technical reason but the passage back and forth is easily made and often helpful. Note that the differential $dA = I + \boldsymbol{G}f$ is of the form *identity* + *compact* and so invertible when it has a trivial null space. It is important that *A is a proper map*.

PROOF. Let H^0 be the space $L^2[0, 1]$, as before with $\|f\|_0^2 = \int_0^1 f^2$. It is enough to prove that f is bounded in H^0 if Af is bounded in H^1, $\boldsymbol{G}: L^1 \to H^1$ being compact, as is self evident from

$$D\boldsymbol{G}f = -\int_0^x y f(y)\, dy + \int_x^1 (1 - x) f(y)\, dy .$$

Let $\|Af\|_1$ be bounded by $c_1 < \infty$; it is to be proved that $\|f\|_0$ is limited by some other number. c_2 depending only on c_1. Let $f = ce$ with $\int_0^1 e^2 = 1$, so that $c = \|f\|_0$ and note

$$(1) \qquad e + cGe^2/2 = Af/c \quad \text{is of length } \leqslant c_1/c \text{ in } H^1 .$$

Now $G1 = \frac{1}{2}\, x(1-x)$, so

$$(2) \qquad \tfrac{1}{4}\int_0^1 x(1-x)\, e^2 = \int_0^1 (e^2/2)\, G1 = \int_0^1 G(e^2/2) = -\int_0^1 e/c + \int_0^1 Af/c^2 \leqslant c_3/c$$

with $c_3 = 1 + c_1/c$. In view of $\int_0^1 e^2 = 1$ and $|Af| \leqslant \|Af\|_1$. The proof is finished by estimating the 3 pieces of

$$(3) \qquad 1 = \int_0^1 e^2 = \int_0^{c^{-\frac{3}{4}}} e^2 + \int_{c^{-\frac{3}{4}}}^{1-c^{-\frac{3}{4}}} e^2 + \int_{1-c^{-\frac{3}{4}}}^1 e^2$$

separately for large $c = \|f\|_0$. 2) implies that the middle piece is over estimated by

$$c_3/c \cdot \left[\frac{1}{4} \cdot \min_{c^{-\frac{3}{4}} \leqslant x \leqslant 1-c^{-\frac{3}{4}}} x(1\ x)\right]^{-1} \leqslant (1 + c_1)\, 5c^{-\frac{1}{4}}$$

if $c^{-\frac{3}{4}} < \frac{1}{5}$. Now by 1), $|e| \leqslant cGe^2/2 + c_1/c$ with

$$Ge^2 = (1-x)\int_0^x ye^2 + x\int_x^1 (1-y)\, e^2 \leqslant x(1-x)\int_1^1 e^2 ,$$

so that $|e| \leqslant c^{\frac{1}{4}}/2 + c_1/c$ for $0 \leqslant x \leqslant c^{-\frac{3}{4}}$, and the first piece is over estimated by $c^{-\frac{3}{4}}(c^{\frac{1}{4}} + c_1/c)$. The same estimate applies for $x \geqslant 1 - c^{-\frac{3}{4}}$, so that by 3),

$$1 \leqslant 2c^{-\frac{3}{4}}(c^{\frac{1}{4}} + c_1/c)^2 + 5(1 + c_1)\, c^{-\frac{1}{4}} = 0(c^{-\frac{1}{4}}) ,$$

providing an explicit limitation of c which is useless at this moment to spell out. The proof is finished.

4. – Singular set.

The singular set of A comprises the points of H^1 where its differential $dA = I + Gf$ has a non-trivial null vector e, so that $e + Gfe = 0$, which is to say that $e \in C^2[0, 1]$ satisfies $-e'' + fe = 0$ with $e(0) = e(1) = 0$. Let F

denote the operator $-D^2+f$ subject to these boundary conditions. The spectrum of F consists of an infinite seguence of (necessarily simple) eigenvalues $\lambda_1(f)<\lambda_2(f)<\lambda_3(f)<\dots$ increasing to $+\infty$ like $\pi^2<4\pi^2<9\pi^2<\dots$ with unit perpendicular eigenfunctions $e_1(f)$, $e_2(f)$, $e_3(f)$, ... attached [2]. The *singular set* of A is now recognized as the union of disjoint sheets $M_n=H^1\cap[f\colon\lambda_n(f)=0]$, $n=1,2,3,\dots$, their geometry occupies much of the sequel.

ASIDE 1. The differential $dA=I+Gf$ satisfies $(dAe,e)_1=(Fe,e)_0$, so the min-max principle of Courant-Hilbert [4] implies that if dA has n eigenvalues <0 then so does F and vice versa; the numerical eigenvalues mostly fail to agree unless they have the special value 0.

ASIDE 2. The geometrical study of the singular sheets in particular, and of A generally, is most conveniently done in the H^0 format in preference to that of H^1. This must be kept in mind: for example, the H^0 gradient of $\lambda_n(f)$ is $\nabla\lambda_n=e_n^2$ while its H^1 gradient is Ge_n^2. The formula is standard as is the variational formula $\dot e_n=(F-\lambda_n)^{-1}(-\pi e_n\dot f_n)$, in which π is the co-projection $I-e_n\otimes e_n$ onto the annihilator of e_n, $(F-\lambda_n)^{-1}$ is construed as a self-map of the annihilator, and $\dot f$ is an infinitesimal variation of f. The simple proof is reproduced.

PROOF. The left hand side of $F\dot e_n=-\dot f e_n+\dot\lambda_n e_n$ is perpendicular to e_n, so $\dot\lambda_n=\dot\lambda_n\int_0^1 e_n^2=\int_0^1\dot f e_n^2$, and $\dot e_n=(F-\lambda_n)^{-1}\left(-\dot f e_n+e_n\int_0^1\dot f e_n^2\right)$ in which you recognize the co-projection of $-\dot f e_n$. The variational formula for $\dot\lambda_n$ is just the statement that $\nabla\lambda_n=e_n^2$ in the H^0 format. Let us also note for future use that *$\lambda_n(f)$ is a (real) analytic function of f as is $e_n(f)$ provided you adopt the normalization $e_n'(0)=1$ instead of the customary but ambiguous $\int_0^1 e_n^2=1$.*

PROOF. Let $y_2(x,\lambda)$ be the solution of $Fy_2=\lambda y_2$ with $y_2(0)=0$ and $y_2'(0)=1$. It satisfies

$$y_2(x,\lambda)=x+\int_0^x d\xi\int_0^\xi d\eta(f-\lambda)y_2(\eta,\lambda)\qquad(0\leqslant x\leqslant 1)$$

and can be represented by the corresponding Neumann series, from which it is seen that $y_2(1,\lambda)$ is *a*) an entire function of λ of order $\frac{1}{2}$ and *b*) an analytic

(2) The format is H^0, so $\int_0^1 e^2=1$, etc.

function of f. The roots of $y_2(1, \lambda) = 0$ are the eigenvalues $\lambda_n(f)$, and as these roots are simple, so they themselves are analytic in f, by a routine application of Cauchy's formula. Then $y_2(x, \lambda_n(f))$ is the eigenfunction $e_n(f)$ with the proposed renormalization $e_n'(0) = 1$, and its analyticity in f is read off of the Neumann series.

AMPLIFICATION. The fact that $\nabla\lambda_n = e_n^2$ cannot vanish shows that the n-th singular sheet M_n *is a smooth manifold of codimension* 1 *in* H^1. It is even *connected*; indeed, it is the graph of a smooth function over a plane. To see this, view H^1 as the direct sum of the plane $\left(f_0 \colon \int_0^1 f_0 = 0\right)$ and the line in the direction $G1 = \frac{1}{2}x(1-x)$. Then $f \in H^1$ can be expressed as $f_0 + cG1$ and $\lambda_n(f)$ is a strictly increasing function of c with exactly one root $c = c_n(f_0)$ having a nice gradient $\nabla c_n = -e_n^2\left[\int_0^1 e_n^2 G1\right]^{-1}$. In particular M_1, M_2, M_3, etc. appear in descending order as in fig. 2.

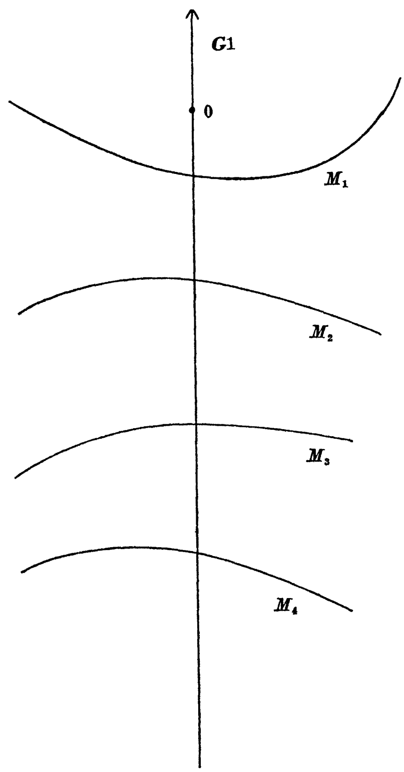


Figure 2

5. – Inverse images are finite.

Let $g \in H^1$ be fixed. It is to be proved that $g = Af = f + Gf^2/2$ has (at most) a finite number of solutions $f \in H^1$. The proof (and much else) depends upon a simple piece of

BASIC GEOMETRY. If $Af_1 = Af_2$, then $0 = Af_1 - Af_2 = e_0 + Gf_0 e_0$ with $e_0 = f_1 - f_2$ and $f_0 = \frac{1}{2}(f_1 + f_2)$ which is to say $e_0 \in C^2[0, 1]$, $e_0(0) = e_0(1) = 0$, and $F_0 e_0 = -e_0'' + f_0 e_0 = 0$. To spell it out, $Af_1 = Af_2$ *only if* $f_0 = \frac{1}{2}(f_1 + f_2)$ *lies on a singular sheet* $M_n = H^1 \cap [f\colon \lambda_n(f) = 0]$ *and* $e_0 = f_1 - f_2$ *is proportional to the corresponding eigenfunction* $e_n(f_0)$. The converse is also true: *if* $f_0 = \frac{1}{2}(f_1 + f_2)$ *is singular and if* $e_0 = f_1 - f_2$ *is proportional to the corresponding eigenfunction, then* $Af_1 = Af_2$. The relations $f_1 = f_0 + e_0/2$, $f_2 = f_0 - e_0/2$ prompt use to speak of f_1 and f_2 as being *reached from* M_n and also to speak of f_2 as a *reflection of* f_1 *across* M_n, but more of that later. See fig. 3.

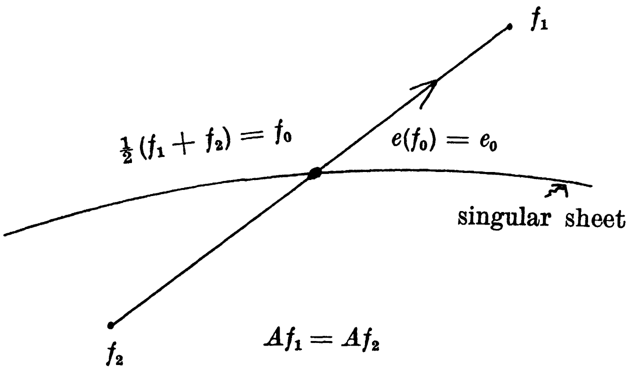


Figure 3

The rest of the proof is divided into 5 steps: it will be seen that the presence of an infinite number of distinct solutions imply that $A^{-1}g$ contains an indefinitely extensible simple curve, this will be contradictory.

STEP 1. Let f_n $(n = 1, 2, 3, \ldots)$ be distinct solutions of $Af = g$. A is proper so it is permissible to suppose that f_n tends to f_∞ in H^1 as $n \uparrow \infty$. This point lies on a singular sheet: indeed, $Af_\infty = g$ so that $\frac{1}{2}(f_\infty + f_n)$ is a singular point, by basic geometry, and as it tends to f_∞ the latter is singular, too. Let $f_\infty \in M_2$ for definiteness. The other sheets are at some distance from f_∞, so it is permissible to require that the singular points $v_n = \frac{1}{2}(f_\infty + f_n)$ lie on M_2 for every $n \geqslant 1$. Let $\frac{1}{2}(f_\infty - f_n) = c_n e_2(v_n)$ in which $e_2(x)$ is temporarily standardized by $e_2'(0) = 1$.

STEP 2. Let f be any point of $A^{-1}g$ with $v = \frac{1}{2}(f_\infty + f) \in M_2$ and put $\frac{1}{2}(f_\infty - f) = ce_2(v)$ as for $f = f_n$ ($n \geqslant 1$) in Step 1. It is to be proved that *the correspondence* $f \to c$ *is* 1:1; in particular, *the numbers* c_n *of Step* 1 *are distinct* (and tend to 0 as $n \uparrow \infty$).

PROOF. Let f_- and f_+ be two such functions so that $f = \frac{1}{2}(f_\infty + f_\pm) + \frac{1}{2}(f_\infty - f_\pm) = v + c_\pm e_2(v_\pm)$ and suppose that c_- and c_+ have the common value c. Then $-e'' + f_\infty e = -e'' + (v + ce)e = ce^2$ for $v = v_\pm$ and $e = e_2(v_\pm)$ with initial conditions $e(0) = 0$ and $e'(0) = 1$. The solution of this problem is unique whence $e_2(v_-) = e_2(v_+)$, $v_+ = v_+$, and $f_- = f_+$, as was to be proved.

STEP 3 is to confirm the existence of an arc of such points $f = v - ce_2(v)$ (faithfully) indexed by small values of c and lying wholly on M_2. This is equivalent to solving $f_\infty = v + ce_2(v)$ for $v \in M_2$ and small c. Note first that the map $(v, c) \in H^1 \times \mathbb{R} \to v + ce_2(v)$ is smooth and that its differential in v is the identity at $v = f$ and $c = 0$. Then, the implicit function theorem guarantees the (unique) existence of a small arc (in H^1) $v = v(c)$ solving $v + ce_2(v) = f_\infty$; the dependence of v upon c is even analytic thanks to the present standardization $e_2^1(0) = 1$; compare art. 4. Now the numbers c_n of Step 1 tend to zero as $n \uparrow \infty$, so by Step 2, the points v_n of Step 1 are nothing but the points $v(c_n)$ of the arc just constructed for $n \uparrow \infty$ at any rate. It follows that the arc lies wholly on M_2 since $\lambda_2(v)$ is an analytic function of v, $v = v(c)$ is analytic in c, and $\lambda_2(v_n) = 0$, while $c_n = o(1)$ for $n \uparrow \infty$; similarly, $f = v - ce_2(v)$ maps to g by A along the whole arc since it does so for $c = c_n = o(1)$.

STEP 4. The arc of Step 3 is now continued to all values $-\infty < c < \infty$: in detail, if continuation is possible for $c < c_* < \infty$, then $f = v - ce_2(v) \in A^{-1}g$ can be made to converge in H^1 by choice of $c_1 < c_2 < \ldots \uparrow c_*$ in view of the properness of A, and the process of Step 3 can be repeated starting at this point so as to provide a continuation of the arc past $c = c_*$.

STEP 5 is to elicit a contradiction from the fact that $f_\infty - f = 2ce_2(v)$: A is proper, so $A^{-1}g$ is compact; this controlls f and so also $v = \frac{1}{2}(f_\infty + f)$ and the size of e_2'' near $x = 0$. It follows from $e_2'(0) = 1$ that $\int_0^1 e_0^2$ cannot be too close to 0 and that $\|f_\infty - f\|_0^2 = 4c^2\int_0^1 e_2^2$ cannot be balanced as $c \uparrow \infty$.

6. – M. and above.

The top sheet M_1 is a smooth surface of co-dimension 1 in H^1, with upward-pointing normal proportional to $\nabla\lambda_1 = e_1^2$ in the H^0 format; it is even *convex* since λ_1 is *concave*:

$$\lambda_1(\tfrac{1}{2}f_1 + \tfrac{1}{2}f_2) > \tfrac{1}{2}\lambda_1(f_1) + \tfrac{1}{2}\lambda_1(f_2) \qquad \textit{unless } f_1 = f_2 .$$

Let M_1^+ be the half space $\lambda_1(f) \geqslant 0$ above M_1.

PROPOSITION 1. *A is 1:1 on M_1^+.*

PROOF. $Af_1 = Af_2$ implies that $f_0 = \frac{1}{2}(f_1 + f_2)$ is singular by basic geometry, in particular, $\lambda_1(f_0) \leqslant 0$, violating the convexity of M_1 if f_1 and f_2 are distinct points of M_1^+.

PROPOSITION 2. *$A(M_1)$ is a smooth (connected) surface of co-dimension 1.*

PROOF. The differential $dA = I + \boldsymbol{G}f$ has null vector $e_1 = e_1(f) \geqslant 0$, it cannot be tangent to M_1 at f in view of $(e_1, \nabla\lambda_1)_0 = (e_1, e_1^2) = \int_0^1 e_1^3 > 0$. The rest is routine using the fact that dA = *identity* + *compact* and the (cheap) inverse function theorem.

PROPOSITION 3. *A is a diffeomorphism of M_1^+.*

PROOF. It is 1:1 and dA has only trivial null space up there (when $\lambda_1(f) = 0$, Ker (dA) is not in TM_1).

PROPOSITION 4. *$A(M_1^+)$ covers the half-space to one side of $A(M_1)$.*

PROOF. The map is open; it is also closed, A being proper.

PROPOSITION 5. *$A(M_1)$ is, itself, a convex surface and $A(M_1^+)$ lies above it.*

PROOF. Let $g_\pm$ be distinct points of $A(M_1^+)$, let $g = (1-c)g_- + cg_+$ $(0 \leqslant c \leqslant 1)$ be the segment joining them, and let $f \in M_1^+$ be the preimage of g. Then (3) $0 = \ddot{g} = e + \boldsymbol{G}(ef + \dot{f}^2)$ with $e = \ddot{f}$. This function is $\leqslant 0$ because it satisfies $-e'' + fe = -(\dot{f})^2 \leqslant 0$ and if it were positive on an open interval $0 \leqslant a < x < b \leqslant 1$ with $e(a) = e(b) = 0$, then you would have

$$\lambda_1(f) \leqslant \frac{\int_a^b [(e')^2 + fe^2]}{\int_a^b e^2} = \frac{\int_a^b e(-e'' + fe)}{\int_a^b e^2} \leqslant 0 ,$$

(3) the spot means $\partial/\partial c$.

and one or both of the inequalities would be strict unless $a = 0$, $b = 1$, and $\dot{f} \equiv 0$, violating $\dot{f} + Gf\dot{f} = \dot{g} \neq 0$. It follows that f lies above the point $(1 - c)\, f_- + cf_+$, so that $\lambda_1(f) > 0$: in short g lies properly above $A(M_1)$.

PROPOSITION 6. *The upward pointing normal to $A(M_1)$ is proportional to $-\, e_1''$ in the H^0 format, or e_1 itself in the H^1 format.*

Warning: $e_1'(0)$ is taken positive here and below, so that $e_1 > 0$ for $0 < x < 1$.

PROOF. Let $\dot{f}$ be a tangent vector to M_1 at f: $\int_0^1 e_1^2 \dot{f} = 0$. Then the corresponding tangent vector to $A(M_1)$ is $\dot{g} = \dot{f} + Gf\dot{f}$ and the normal to $A(M_1)$ must satisfy

$$0 = \int_0^1 n(\dot{f} + Gf\dot{f}) = \int_0^1 (n + fGn)\dot{f}\,,$$

whence $n + fGn = ce_1^2$. Now

$$c\int_0^1 e_1^3 = \int_0^1 e_1(n + fGn) = \int_0^1 n(e_1 + Gfe_1) = 0$$

and $\int_0^1 e_1^3 > 0$, so $c = 0$ and $n + fGn = 0$ implies that $n = -\, e''$; indeed, this is a solution and there cannot be another, $I + Gf$ being of index 0. It remains to fix the signature of the *upward*-pointing normal. To do this, note that $f + e_1 \in M_1^+$ for $f \in M_1$ so that $A(f + e_1) - A(f) = e_1 + Gfe_1 + Ge_1^2/2 = Ge_1^2/2$ points upward from Af into $A(M_+^1)$. Then you have only to check that $n = -\, e''$ has the proper signature: $\int_0^1 (-\, D^2 e_1)\, Ge_1^2 = \int_0^1 e_1^3 > 0$.

7. – Legendre duality.

It is a self-evident fact that M_1^+ can be described as the class of functions $f \in H^1$ with

$$(1) \qquad \tfrac{1}{4}\int_0^1 (e')^2/e + (f, e)_0 \geqslant 0$$

in which $e = e_1^2$ is the square of any ground state with its natural normalization $\int e_1^2 = 1$: in fact, the integral is nothing but $\int_0^1 (e_1')^2$, so that 1) is the quadratic form $Q[e_1]$ in a light disguise and this is least for $e_1 = e_1(f)$. To recapitulate: if $\lambda_1^*(e) = \frac{1}{4}\int_0^1 e^{-1}(e')^2$ for positive function e, then *for fixed* $f \in H^1$, $\lambda_1(f)$ *is the minimum of* $\lambda_1^*(e_1^2) + (f, e_1^2)_0$ *taken over the class of ground states* e_1. Indeed, there is a simple duality here: for fixed e_1, $\lambda_1^*(e_1^2)$, *is the maximum of* $\lambda_1(f) - (f, e_1^2)_0$ *taken over* $f \in H^1$ *because* $Q[e_1] = \lambda_1(f)$ *only for* $e_1 = e_1(f)$ and is larger otherwise, so that

$$\lambda_1(f) - (f, e_1^2)_0 \leqslant Q[e_1] - (f, e_1^2)_0 \equiv \int_0^1 (e_2')^2 = \lambda_1^*(e_1^2)$$

with equality only if $e_1 = e_1(f)$. The reason for reproducing these trivialities will appear in a moment. To begin with, 1) is equivalent to the simple geometrical fact that $f \in M_1^+$, *precisely when the angle between* $f - f_1$ *and the upward normal* e_1^2 at $f_1 \in M_1$ *is less than* $90°$. See fig. 4. To spell it out:

$$(2) \qquad 0 \leqslant \int_0^1 (f - f_1)\, e_1^2 = \int_0^1 f e_1^2 - \int_0^1 e_1 e_1'' \qquad [e_1'' = fe]$$
$$= \int_0^1 [(e')^2 + f e_1^2] = Q[e_1]\,.$$

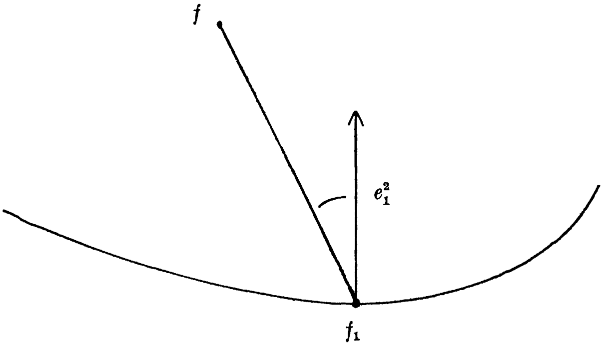


Figure 4

Warning. The class of functions e_1 is narrower than that admitted into the competition before, e_1 being the ground state of a point $f_1 \in M_1$, but it is easy to see that the *infimum* of $Q[e_1]$ is unchanged: it suffices to approximate the general e_1 by a function $e > 0$ in $0 < x < 1$ with both e and e''/e vanishing at $x = 0$ and $x = 1$, which is easy to do.

The same idea is now applied to $A(M_1)$ (with more profit). This is a convex surface and $A(M_1^+)$ fills out what lies above it. Now in the H^1 format, the upward pointing normal to $A(M_1)$ at $g_1 = Af_1$ is proportional $e_1 = e_1(f_1)$ so $g \in A(M_1^+)$ precisely when the inner product between $g - g_1$ and e_1 is $\geqslant 0$ for every $g_1 \in A(M_1)$. To spell it out,

$$\begin{aligned} 0 \leqslant (g - g_1, e_1) &= (g, e_1)_1 + \int_0^1 g_1 e_1'' \\ &= (g, e_1)_1 + \int_0^1 (f + Gf_1^2/2)\, e_1'' \quad [g_1 = Af_1] \\ &= (g, e_1)_1 + \int_0^1 f_1^2 e_1 - \tfrac{1}{2}\int_0^1 f_1^2 e_1 \quad [e_1'' = e_1 f_1] \\ &= \tfrac{1}{2}\int_0^1 (e_1'')^2/e_1 + (g, e_1)_1 \equiv 3 \end{aligned}$$

in which form the criterion shows a remarkable resemblance to 1) in its narrower form: $\frac{1}{4}\int_0^1 (e')^2/e + (f, e)_0 \geqslant 0$ *for every* $e = e_1^2$ *from* M_1. Notice that if $g = Af$ for any $f \in H^1$, then

$$\begin{aligned} 2) &= \int_0^1 (g - g_1)(-e_1'') \\ &= \int_0^1 [f - f_1 + G(f^2 - f_1^2)/2](-e_1'') \\ &= \int_0^1 (f_1 - f) f_1 e_1 + \tfrac{1}{2}\int_0^1 (f^2 - f_1^2)\, e_1 \quad [e_1'' = f_1 e_1] \\ &= \int_0^1 (f - f_1)^2 e_1 \geqslant 0 \end{aligned}$$

since e_1 is positive, in which a remarkable similarity to 2) is seen, but the chief point is that this inequality *identifies, via* 3), $A(M_1^+)$ *as the complete range of* A; *in particular, if* f_- *lies properly below the topmost singular sheet* M_1, *then* $g = Af_-$ *is also the image of some point* f_+ *above* M_1, *i.e., the count* $N(g)$ *of the preimages is* $\geqslant 2$. This is not so on $A(M_1)$: *there,* $N(g) = 1$, *i.e.,* no other point of H^1 has the same image as $f_1 \in M_1$; in particular, the point f_+ lies properly above M_1.

PROOF. Let $f_1 \in M$ and let $g = Af_1$ have another preimage f_2. The latter is properly below M_1, A being 1:1 above. Then $v = \frac{1}{2}(f_1 + f_2)$ is singular and $e = \frac{1}{2}(f_1 - f_2)$ is proportional to the corresponding null vector, by the basic geometry of art. 5. Now $e'' = ve$ implies $f_1 = v + e = e''/e + e$ so that $e_1'' = f_1 e_1$ takes the form $e'' e_1 - e e_1'' = - e^2 e_1 \leqslant 0$. This states that $e' e_1 - e e_1'$ is decreasing for $0 \leqslant x \leqslant 1$, and as it vanishes at $x = 0$ and $x = 1$, it vanishes everywhere and $e \equiv 0$, *i.e.*, $f_1 = f_2$.

AMPLIFICATION 1. Let us investigate a little more the relation between the point f_- below M_1 and the associated point f_+ above M_1 with the same image. Basic geometry states that $v = \frac{1}{2}(f_+ + f_-)$ is singular while $e = \frac{1}{2}(f_+ - f_-)$ is proportional to the corresponding null vector. The preceding argument leads to a contradiction if v lies on M_2 or on any lower sheet; for example, if $v \in M_2$, then e has one root $0 < r < 1$ and e is negative for $r < x < 1$, say, where $e'' e_1 - e e_+'' = - e_+^2 e_+ + \lambda_1(f_+) e e_+$ $[e_+ = e_1(f_+)]$ is negative for $r < x < 1$, so that $e' e_+ - e e_+'$ decreases from its left hand value $e' e_+ < 0$ and cannot vanish at $x = 1$. The only possibility is that $v \in M_1$. Think of $e_1(v)$ as a vector field attached to $v \in M_1$: Then you may speak of $f_- = v - c e_1(v)$ as being *reached* from M_1 by this field; compare fig. 5.

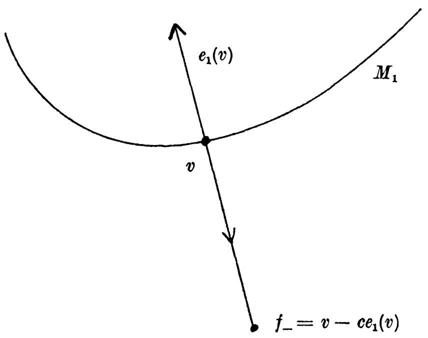


Figure 5

It is easy to see that this can be done in *just one way*: in fact if $v_\pm - c_\pm e_1(v_\pm)$ are distinct reaches from M_1 to f_- then basic geometry implies that $v_+ + c_\pm e_1(v_\pm)$ are (distinct) primages of $v = Af_1$ both above M_1, which is not possible.

The phrase simply reached is used to convey this situation.

AMPLIFICATION 2. The idea of reaching is illustrated by the fact that the normal field to any closed convex surface in $H^0 = L^2[0, 1]$ reaches simply into the whole region below M. The proof is trivial: you have only to find the point f_0 of M closest to f_- below M and to note that $e = f_1 - f_0$ is normal to M at f_0.

AMPLIFICATION 3. It is also true that *the field* e_1 *reaches from* M_1 *to the whole of* M_1^+, *but not* in general) *simply*.

PROOF. Let $f \in M_1^+$; it is required to find $v \in M_1$ so that $f = v + ce_1(v)$ with $c > 0$. This is the same as to say that $-e'' + fe = ce^2$ can be solved for some $c > 0$, by a ground state $e = e_1$ from M_1. To prove this, you minimize $Q[e] = \int_0^1 [(e')^2 + fe^2] \geqslant 0$ in the positive part of H^1, subject to $\int_0^1 e^3 = 1$. $Q[e]$ controls

$$\left|\int_0^1 fe^2\right| \leqslant \left(\int_0^1 |f|^3\right)^{\frac{1}{3}} \left(\int_0^1 e^3\right)^{\frac{2}{3}} \leqslant \|f\|_1 \left(\int_0^1 e^3\right)^{\frac{2}{3}}$$

so $\int (e')^2$ is controlled as you head toward the infimum of $Q[e]$, and the weak compactness of H^1 ensures the existence of an actual minimizer $e = e_0$. At that point, $-e_0'' + fe_0 = ce_0^2$, and $Q[e_0] = c\int_0^1 e_0^3 = c$ makes $c \geqslant 0$; moreover, e_0 has no interior root since it is an eigenfunction of $-D^2 + f - ce_0$ of one signature and so must be proportional to its ground state. A little rescaling does the rest, except to note that *the reach need not be simple*. The point is that 2 (and, indeed, indefinitely many) points f_- can lie below M_1 and have the same image, as will be seen in arts. 8, 9, 12, and 14 below, and these can be reached from M_1: $f_- = v - ce_1(v)$ $(c > 0)$. The reflected points $f_+ = v + ce(v)$ must coincide above M_1 having the same image, so that f_+ can be reached, but in more than one way.

AMPLIFICATION 4. The possibility of reaching from, *e.g.*, M_2 via the field e_2 is also important for the elucidation of $N(g)$. It is much more complicated; see arts. 9 and 10.

AMPLIFICATION 5. The criterion 1) for inclusion in M_1^+ is compared to that for inclusion in $A(M_1^+)$:

$$\frac{1}{2}\int_0^1 \frac{(e_1'')^2}{e_1} + (g, e_1)_1 \geqslant 0 \,. \tag{3}$$

This suggests that for fixed $g \in A(M_1^+)$, the infimum of 3), taken over ground states e_1 from M_1, should have an elegant geometrical meaning. A preliminary scaling by the factor $\left(\int_0^1 e_1\right)^{-1}$ is suggested by the scaling $\int_0^1 e = \int_0^1 e_1^2 = 1$

implicit in 1). Then it is easy to check that 3) has a unique critical point at the ground state $e_1 = e_1(f_0)$ distinguished by the fact that $v_0 = Af_0 \in AM_1$ is the (unique) point of that surface from which g can be reached by the constant field $Gl = \frac{1}{2}x(1-x)$: $g = g_0 + Gc1$. The significance of this is obscure.

AMPLIFICATION 6. The fact that every point f_- below M_1 can be reached from M_1 may be confirmed as follows: the map $(f, c) \to f - ce_1(f)$ of $M_1 \times (0, \infty)$ into the region below M_1, has the differential $(\dot{f}, \dot{c}) \in TM_1 \times \mathbb{R} \to \dot{f} + cP^{-1}(e_1 \dot{f}) - \dot{c}e_1$. This is a compact perturbation of the trivially invertible map $(\dot{f}, \dot{c}) \to \dot{f} - \dot{c}e_1$ and is itself invertible since the vanishing of $\dot{f} + cF^{-1}(e_1\dot{f}) - \dot{c}e_1$ implies that $\dot{f}$ satisfies $-\dot{f}'' + (f + ce_1)\dot{f} = 0$. This is contradictory: it states that $f + ce_1$, which is *above* M_1 is singular and so has lowest eigenvalue $\leqslant 0$. The computation shows that $f \to f - ce_1(f)$ is a local diffeomorphism for $c > 0$. The rest follows from the fact that the map is *proper*, by a general principle; see Berger [2]. This is easy: if $(f, c) \in M_1 \times (0, \infty)$ and if $f_- = f - ce_1(f)$ is bounded in H^1, then $-e_1'' + f_- e_1 = -ce_1^2$ implies that

$$-c\int_0^1 e^3 = \int_0^1 [(e_1')^2 + f_- e_1^2] \geqslant \lambda_1(f_-)$$

is bounded below. This limits the size of $c > 0$ in view of $\int_0^1 e_1^3 \geqslant \left(\int_0^1 e_1^2\right)^{\frac{3}{2}} = 1$ and so also the size of f in H^1 in view of

$$\int_0^1 (e_1')^2 = -c\int_0^1 e^3 \int_0^1 f_- e_1^2 \leqslant |c| \sqrt{\int_0^1 (e_1')^2} + \|f\|_1$$

and $f = f_- + ce_1$. The rest is routine.

AMPLIFICATION 7. The same can be done for the field e_1^p for any $p \geqslant 1$, the case $p = 2$ being the normal field: *it reaches every point above M_1 or below*; moreover, the reach is simple below M_1, which is to say that *$-e'' + fe = -e^{p+1}$ has just one positive solution with $e(0) = e(1) = 0$, if $\lambda_1(f) \leqslant 0$ and at least one negative solution if $\lambda_1(f) > 0$.* This is nothing new in itself, but the geometrical picture is more appealing.

8. – Folds and cusps.

The next topic is the local behavior of A in the vicinity of a singular sheet $M = M_n$. Let $f \in M$ and let $e = e_n(f)$ be the null vector of the differential $dA = I + Gf$. The nature of the singularity is regulated by the degree of contact of e with M, as measured by the vanishing of the successive derivatives of the associated eigenvalue $\lambda = \lambda_n(f)$ in the direction $\dot{f} = e$. Let $M = M_2$ for definiteness, so that $e = e_2(f)$ and temporarily reserve the letters e_+ and λ_+ for the second eigenfunction and eigenvalue of the variable point $f_+ = f + ce$. Then with $\dot{f}_+ = e$, $\cdot = \partial/\partial c$, and π denoting co-projection $I - e_+ \otimes e_+$,

$$\dot{\lambda}_+ = \int_0^1 e_+^2 \dot{f}_+ = \int_0^1 e_+^2 e = \int_0^1 e^3 \quad \text{at } c = 0;$$

$$\ddot{\lambda}_+ = \int_0^1 2e_+ \dot{e}_+ \dot{f}_+ = 2\int_0^1 e_+ \dot{f}(F_+ - \lambda_+)^{-1}(-\pi e_+ \dot{f}_+)$$

$$= 2\int_0^1 e^2 F^{-1}(-\pi e^2) \quad \text{at } c = 0, \text{ which reduces to}$$

$$-2\int_0^1 e^2 F^{-1} e^2 \equiv -2I_2 \quad \text{if} \quad I_1 \equiv \int_0^1 e^3 = 0;$$

similarly

$$\dddot{\lambda}_+ = 6\int_0^1 e[F^{-1}(e^2)]^2 \equiv 6I_3 \quad \text{at } c = 0$$

if $I_1 = I_2 = 0$; and so forth. Not everyone of these successive quantities can vanish at a singular point, so *e has a definite degree of contact.*

Proof. If they did, then the analyticity of λ_2 imply that the whole line $f + \mathbb{R}e$ lies inside M_2 and that cannot be: indeed, with normalization $e'(0) > 0$, e is negative near $x = 1$, so that $f_+ = f + ce \downarrow -\infty$ there as $c \uparrow \infty$ and e_+, which must solve $-e_+'' + f_+ e_+ = 0$ and have one interior root, has in fact an unlimited number of roots. More precisely, if $0 < r < 1$ is the interior root of e, if $r < a < b \leqslant 1$, and if m is the minimum of $-e > 0$ in $a \leqslant x \leqslant b$, then c cannot exceed the root of

$$\sqrt{\mathrm{cm} - \|f\|_\infty}\,(b - a) = 2\pi$$

and still keep $f + ce \in M_2$. Now the map $Af = f + Gf^2/2$ has a standard local form for each of the first few degrees of contact: up to local diffeomorphisms to the left and to the right, it can be expressed in suitable local coordinates $(x_1, x_2, x_3, \dots)$ as *a fold* $(x_1, x_2, x_3, \dots) \to (x_1^2, x_2, x_3, \dots)$ if $I_1 \neq 0$, a *cusp* $(x_1, x_2, x_3, \dots) \to (x_1^3 + x_1 x_2, x_2, x_3, \dots)$ if $I_1 = 0$ but $I_2 \neq 0$, and so forth. Naturally, such standard forms and the means to recognize them are common knowledge; see, for example, Whitney [10: 395-400) or Golubitsky-Guillemin [5: 146-148].

SOME GEOMETRY. The locus $L_1 = M \cap (I_1 = 0)$ is void if $M = M_1$ since $\int_0^1 e_1^3 > 0$ there. *L_i is not void for any lower sheet: in fact, it is a connected submanifold of co-dimension 1 in M with $I_1 > 0$ to one side and $I_1 < 0$ to the other.*

PROOF. M_2 is typical: e_2 has one interior root and M_2 is the class of functions $f = e''/e \in H^1$ produced by such $e = e_2$. The non-emptiness of the locus is seen by taking e anti-symmetric about $x = \frac{1}{2}$; its connection is plain. Now if $\dot{f}$ is tangent to M_2 at f_1 then $\int_0^1 \dot{f} e^2 = 0$,

$$\dot{I}_1 = 3\int_0^1 e^2 \dot{e} = 3\int_0^1 e^2 F^{-1}(-e\dot{f}) = -3\int_0^1 \dot{f} e F^{-1}(e^2)$$

on locus (*), so that $\nabla I_1 = -3eF^{-1}(e^2)$ there, and what must be still proved is the independence of the latter and the normal e^2. But if $c_1 e^2 + c_2 eF^{-1}(e^2)$ vanishes, then so does $c_1 e + c_2 F^{-1}(e^2)$ and application of F produces $c_2 e^2 = 0$, whence $c_1 = c_2 = 0$. The proof is finished.

The pattern repeats itself: *the sublocus $L_2 = L_1 \cap (I_2 = 0)$ is a submanifold of co-dimension 1 in L_1 with $I_2 > 0$ to one side and $I_2 < 0$ to the other*; presumably it is connected but this is not proved.

PROOF for M_2. The first item is that L_2 is not void. Let e be anti-symmetric about $x = \frac{1}{2}$ so that $\int_0^1 e^3 = 0$ and $f \in L_1$. Then $h = F^{-1}(e^2)$ is symmetric about $\frac{1}{2}$; it is to be proved that $I_2 = 2\int_0^{\frac{1}{2}} e^2 h$ can be positive and also negative. The function h satisfies $e^2 = Fh = -h'' + fh = -h'' + e''h/e$, which may be integrated with the aid of $h(0) = 0$ to obtain $h = ce - e\int_0^x d\xi e^{-2}\int_0^\xi e^3$ the constant c being determined by the symmetry of h, expressed as $h'(\frac{1}{2}) = 0$; in particular $\frac{1}{2}I_2 = c\int_0^{\frac{1}{2}} e^3 - \int_0^{\frac{1}{2}} e^3 \int_0 e^{-2} \int_0 e^3$. Let $e = x/\theta$ for $0 \leqslant x < \theta$ and $e = (\frac{1}{2} - x)(\frac{1}{2} - \theta)^{-1}$ for $\theta \leqslant x \leqslant \frac{1}{2}$ with adjustable $0 < \theta < \frac{1}{2}$; this is not a smooth function but no matter. Now it is easy to compute $c = \theta^2/12$

(*) $\dot{e}^2 = \pi e^2$ in view of $\int_0^1 e^3 = 0$.

$+\theta/8-1/16$, and this is negative for $\theta=0^+$ so that $I_2<0$, while for $\theta=\frac{1}{2}^-$, $c=3^{-1}\cdot 2^{-4}$ while

$$\left(\frac{1}{2}\right)I_2=2^{-7}\cdot 3^{-1}-2^4\int_0^{\frac{1}{2}}x^3\int_0 x^{-2}\int_0 x^3=2^{-5}\cdot 3^{-1}\left(\frac{1}{4}-\frac{1}{7}\right)>0\,.$$

This proves that L_2 is not void. Next, compute ∇I_2 along L_2: if $\dot{f}$ is tangent to L_1 and L_2, then $\int_0^1 \dot{f}e^2=0=\int_0^1\dot{f}eF^{-1}(e^2)$, so

$$\begin{aligned}\dot{I}_2&=4\int_0^1 eF^{-1}(-\dot{f}e)F^{-1}(e^2)+\int_0^1 e^2(-F^{-1}\dot{f}F^{-1})e^2\\&=-4\int_0^1\dot{f}eF^{-1}[eF^{-1}(e^2)]-\int_0^1\dot{f}[F^{-1}(e^2)]^2,\end{aligned}$$

with the result that

$$\nabla I_2=-4eF^{-1}[eF^{-1}(e^2)]-[F^{-1}(e^2)]^2\quad\text{ in } T^*L_1\,.$$

It is required to prove the independence of *a*) the normal e^2, *b*) $\nabla I_1=-3eF^{-1}(e^2)$, and *c*) ∇I_2. But if

$$ce^2+c_2eF^{-1}(e^2)+c_3\big(4eF^{-1}[eF^{-1}(e^2)]+[F^{-1}(e^2)]\big)=0\,,$$

and if $c_3\neq 0$, then $h=F^{-1}(e^2)$ vanishes with e, so that if $0\leqslant a<x<b$ are consecutive roots of e with e positive between, then

$$0<\int_a^b e^3=\int_a^b eFh=\int_a^b hFe=0\,.$$

This contradiction makes c_3 vanish. The rest was done before.

FOLDS. $I_1=\int_0^1 e^3\neq 0$ states that the singular direction e is not tangent to the sheet. This is typical of a fold. Let D be a small patch of M about a fixed point f_0 of this type. Then $A\colon D\to D'\equiv A(D)$ is a diffeomorphism. Choose left hand coordinates $(x_1, x_2, x_3, \ldots)$ near D so that $f_0=(0, 0, 0, \ldots)$, $D=(x_1=0)$ and $e(f)=\partial/\partial x_1$ on D and right hand coordinates $(x_1', x_2', x_3^1, \ldots)$ near D' so that $D'=(x_1'=0)$.

The inverse diffeomorphism A^{-1} of D' to D is extended to a diffeomorphism of the coordinate neighborhood about D' to the coordinate neighborhood about D and then applied to the right to reduce the map to the form $(x_1, x_2, x_3, \ldots) \to (x_1', x_2, x_3, \ldots)$, with x_1' a smooth function of the left hand coordinates; in this format, the fact that $e(f_0)$ has contact of degree 1 with D is expressed as $\partial x_1'/\partial x_1 = 0$ and $\partial^2 x_1'/\partial x_1^2 \neq 0$ at $x_1 = 0$. Then $x_1' = x_1^2 h$ with h smooth and positive by choice of the signature of x_1', and introduction of the new coordinate $x_1'' = x_1\sqrt{h}$ combined with the preliminary left hand diffeomorphism $(x_1'', x_2, x_3, \ldots) \to (x_1, x_2, x_3, \ldots)$ produces the diagram of the standard fold:

$$(x_1'', x_2, x_3, \ldots) \to (x_1, x_2, x_3, \ldots) \to \big((x_1'')^2, x_2, x_3, \ldots\big).$$

AMPLIFICATION 1. $\int_0^1 e^3 > 0$ on M, so A is a local fold at each of its points. This amplifies the result of art. 7. *$N(g) = 1$ or $\geqslant 2$ according as $g \in A(M_1)$ or not.*

CUSPS. Now let $e = e_2(f_0)$ have contact of degree 2 with the singular sheet so that $I_1 = \int_0^1 e^3 = 0$ but $I_2 = \int_0^1 e^2 F_0^{-1}(e^2) \neq 0$; it is to be proved that *A is a local cusp.* Now e is tangent to M but not the locus L_1, so the map A is a diffeomorphism of a small patch $D \subset L_1$ about f_0. Choose left-hand coordinates $(x_1, x_2, x_3, \ldots)$ near D so that $f_0 = (0, 0, 0, \ldots)$, $D = (x_1 = x_2 = 0)$, and $e(f_0) = \partial/\partial x_1$ on D, and right hand coordinates $(x_1', x_2', x_3', \ldots)$ near $A(D) = D'$ so that $D' = (x_1' = x_2' = 0)$.

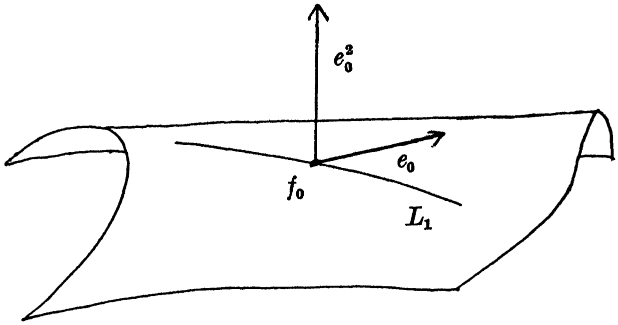


Figure 6

As before the inverse diffeomorphism of D' to D can be used to reduce the map to the form $(x_1, x_2, x_3, \ldots) \to (x_1', x_2', x_3, \ldots)$ with x_1', x_2' smooth functions of the left hand coordinates. The reduction to the standard cusp is indicated for fixed x_3, x_4; it is easy to see that the necessary 2-dimensional coordinate changes depend smoothly on these variable and so define diffeomorphism throughout the coordinate patches.

PROOF. The vanishing of the Jacobian determinant Δ of x_1', x_2' with respect to x_1, x_2 determines a singular curve. This is smooth [grad $\Delta \neq 0$] and the Jacobian annihilates the singular direction $\partial/\partial x_1$ [$\partial x_1'/\partial x_1 = \partial x_2'/\partial x_2 = 0$]; in addition $\partial/\partial x_1$ is tangent to the singular curve at the origin [$\partial\Delta/\partial x_1 = 0$] but has contact only of degree 2 [$\partial^2\Delta/\partial x_1^2 \neq 0$]. These are precisely the conditions of Whitney [10] for a cusp; his proof is elementary but not simple.

The stated conditions lead rapidly to a reduced map of the form $x_1 = x_1' + x_1x_2(1 + \ldots)$, $x_1^3 = x_2$; it is the final reduction to $x_1 = x_2' + x_1x_2$, $x_2' = x_2$ *plain* which is troublesome. Golubitsky-Guilleman [5] shorten the proof, the key point being that x_1 is the root of a cubic form $\mathfrak{T}[x_1', x_2'][x_1]$, but this is a fairly deep fact, so there seems to be no really cheap way.

AMPLIFICATION 2. The discussion indicates that the images of the lower sheets M_2, M_3 etc. in $A(M_1^+)$ are pretty complicated. It is not known how they look in the large, though it might be helpful to compute then on the machine for a finite-dimensional model: for example, you could model $f \to -D^2f + f^2/2$ by

$$(x_1, x_2, \ldots, x_d) \to (2_1 - x_2 + x_1^2, 2x_2 - x_1 - x_3 + x_2^2, \ldots, 2x_d - x_{d-1} + x)\,.$$

This has not been done, except to confirm the presence of cusps in dimension 4 by elementary graphics.

9. – A is not injective on M_2.

The same is true of the lower sheets M_3, M_4, etc., M_2 is just to fix ideas.

PROOF. The locus $L_1 \subset M_2$ where $I_1 = \int_0^1 e^3$ vanishes is a submanifold of co-dimension 1, and the sublocus L_2 where $I_2 = \int_0^1 e^2F^{-1}e^2$ also vanishes is a submanifold of co-dimension 1 in *that*; moreover, $I_2 > 0$ to one side of L_2 in L and $I_2 < 0$ to the other. Fix a point $f_0 \in L_2$ and an ambient ball $\mathcal{O}$ so small that $L_2 \cap \mathcal{O}$ is connected. The maps $f \to f_\pm = f \pm ce_2(f)$ are defined on $\mathcal{O}$; $c \geqslant 0$ is a parameter. Let $e_\pm(f) = e_2(f_\pm)$ and $\lambda_\pm(f) = \lambda_2(f_\pm)$. The goal is to prove

$$\mathcal{O}_- = \mathcal{O} \cap (f\colon f_- \in M_2) = \mathcal{O} \cap (f\colon \lambda_-(f) = 0)\,,$$

$$\mathcal{O}_+ = \mathcal{O} \cap (f\colon f_+ \in M_2) = \mathcal{O} \cap (f\colon \lambda_+(f) = 0)\,,$$

and M_2 itself have a nonvoid intersection for small $c > 0$. Then, the result follows by basic geometry. The steps of the proof are carried out mostly for $\mathcal{O}_+$; they are the same for $\mathcal{O}_-$.

STEP 1 is to compute the gradients

$$\nabla\lambda_\pm = e_\pm^2 \mp c e_2 (F - \lambda_2)^{-1}(\pi e_\pm^2)\,,$$

in which $e_2 = e_2(f)$, $\lambda_2 = \lambda_2(f)$, and π is the co-projection $I - e_2 \otimes e_2$:

PROOF. The response of λ_+ to an infinitesimal variation $\dot{f}$ is

$$\begin{aligned}\dot{\lambda}_+ &= \int_0^1 e_+^2 \dot{f}_+ = \int_0^1 e_+^2 (\dot{f} + c\dot{e}_2)\\ &= \int_0^1 e_+^2 [\dot{f} + c(F - \lambda_2)^{-1}(-\pi e_2 \dot{f})]\\ &= \int_0^1 \dot{f}[e_+^2 - c e_2 (F - \lambda_2)^{-1}(\pi e_+^2)]\,.\end{aligned}$$

The gradient is read off from that.

STEP 2. *$\mathcal{O}_\pm$ is a manifold of codimension* 1 *in* H^1.

PROOF. $\nabla\lambda_+ = 0$ is contradictory. Then $e_+^2 = c e_2 (F - \lambda_2)^{-1}(\pi e_+^2)$ and $1 = \int_0^1 e_+^2 = c\int_0^1 e^2 (F - \lambda_2)^{-1}(\pi e_+^2) = 0$, the range of $(F - \lambda_2)^{-1}$ being the annihilator of e_2.

STEP 3. *$\mathcal{O}_\pm$ intersects M_2 for small $c > 0$.*

PROOF. $\int_0^1 e_2^3$ takes both signs in $M_2 \cap \mathcal{O}$, so that for small $c > 0$, $\lambda_+(f) = \lambda_2(f) + c\int_0^1 e_2^3 + o(c)$ does the same, and as $M_2 \cap \mathcal{O}$ is connected, $\lambda_+(f)$ vanishes someplace there, *i.e.* $M_2 \cap \mathcal{O}_+$ is not void.

STEP 4. *$M_2 \cap \mathcal{O}_\pm$ is a submanifold of co-dimension* 1 *in M_2.*

PROOF. The normal space of $M_2 \cap \mathcal{O}_+$ is the span of e_2^2 and $\nabla\lambda_+ = e_+^2 - ce_2F^{-1}(\pi e_+^2)$ and this is genuinely two dimensional: if not, there is a dependence

$$c_1 e_2^2 + c_2[e_+^2 - ce_2F^{-1}(\pi e)] = 0 \quad \text{with } c_2 \neq 0\,,$$

and e_+ vanishes with e_2: Let $0 \leqslant a < b \leqslant 1$ be consecutive roots of $e_2 = 0$ with $e_2 > 0$ between. Then the lowest eigenvalue of F restricted to $a \leqslant x \leqslant b$ is 0. But $f_+ = f + ce_2$ exceeds f in that interval, so the restricted eigenvalues λ_1^* satisfy $\lambda_1^*(f_+) > \lambda_1^*(f) = 0$, contradicting the fact that $e_+(a) = e_+(b) = 0$ which implies $\lambda_1^*(f_+) \leqslant 0$.

STEP 5. *$M_2 \cap \mathcal{O}_\pm$ moves smoothly out from L_1 as c increases from $c = 0_+$; in particular, it is connected near f_0.*

PROOF. The eigenvalue $\lambda_+(f) = \lambda_2(f + ce_2)$ is smooth in c and vanishes in the patch $M_2 \cap \mathcal{O}$ at $c = 0$; moreover, its derivative at $c = 0$ is $\dot\lambda_+ = \int_0^1 e^3$ which vanishes along L_1. Choose local coordinates (x_1, f) on the patch: f along L_1 and x_1 corresponding to the direction $\nabla I_1 = -3e_2F^{-1}e_2^2$ normal to L_1 at f. Then $\lambda_+(f)$ is of the form ch with h a smooth function of c, x_1, and f, and for $c > 0$, $M_2 \cap \mathcal{O}_+$ is the locus $h = 0$. Now

$$\begin{aligned}
\frac{\partial h}{\partial x_1}(0, 0, f) &= \frac{\partial^2 \lambda_+}{\partial c\partial x_1} \quad \text{at } c = x_1 = 0\\
&= \frac{\partial}{\partial c}\int_0^1 [\nabla\lambda_+ \cdot \nabla I, \quad \text{at } x_1 = 0] \quad \text{at } c = 0\\
&= \frac{\partial}{\partial c}\int_0^1 [e_+^2 - ce_2F^{-1}\pi e_+^2]\,(-3e_2F^{-1}e_2^2) \quad \text{at } c = 0\\
&= \int_0^1 [2e_2F^{-1}(-e_2^2) - e_2F^{-1}(e_2^2)](-3e_2F^{-1}e_2^2)\\
&= 9\int_0^1 e_2^2[F^{-1}(e_2^2)]^2 > 0\,,
\end{aligned}$$

so the implicit function theorem permits you to solve $\lambda_+(f) = 0$ $[h = 0]$ by smooth choice of $x_1 = x_1(c, f)$. The proof is finished.

STEP 6 is to study just how $\mathcal{O}_\pm$ moves out from L_1. I_1 is < 0 to one side of L_1 and > 0 to the other as in fig. 7, in which the curve represents L_1 and the dot represents L_2; similarly, $I_2 = \int_0^1 e^2 F^{-1} e^2$ is > 0 to one side of the (connected) sublocus $L_2 \cap \mathcal{O}_+$ inside $L_1 \cap \mathcal{O}_+$ and < 0 to the other. The evaluations $\dot{\lambda}_+[= \partial\lambda_+/\partial c$ at $c = 0] = I_1$ and $\ddot{\lambda}_+[= \partial^2\lambda_+/\partial c^2$ at $c = 0] = -2I_2$ show that for small $c > 0$,

$$\begin{aligned} \lambda_+ &< 0 \quad \text{in } M_2 \cap (I_1 < 0) \\ \lambda_+ &> 0 \quad \text{in } M_2 \cap (I_1 > 0) \\ \lambda_+ &< 0 \quad \text{in } L_1 \cap (I_2 > 0) \\ \lambda_+ &> 0 \quad \text{in } L_1 \cap (I_2 < 0)\,. \end{aligned}$$

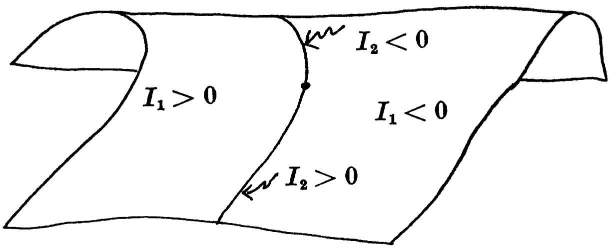


Figure 7

It follows from fig. 7 that $M_2 \cap \mathcal{O}_+$ moves out from L_1 *into* $M_2 \cap (I_2 > 0)$ *near* $L_1 \cap (I_2 > 0)$ and (oppositely) into $M_2 \cap (I_2 < 0)$ near $L_1 \cap (I_2 < 0)$, as in fig. 8. The same holds for $M_2 \cap \mathcal{O}_-$ except that $\lambda_- = -I_1$ has the opposite signature to λ_+, so that the motion is opposed: *from* $L_1 \cap (I_2 > 0)$ *into* $M_2 \cap (I_1 < 0)$ and *from* $L_1 \cap (I_2 < 0)$ *into* $M_2 \cap (I_1 > 0)$. The proof is finished by using the connectedness of $M_2 \cap \mathcal{O}_\pm$ to force them to intersect.

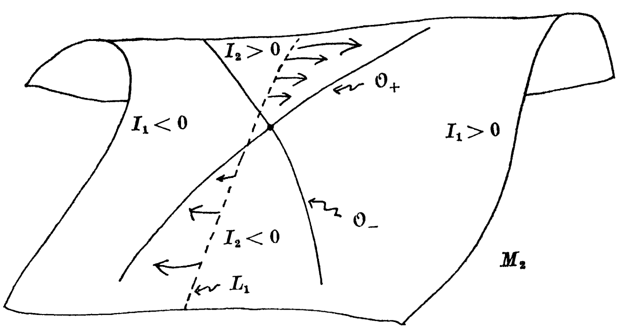


Figure 8

10. – An inequality.

The groundstate inequality:

$$\lambda_1(\tfrac{1}{2}f_1 + \tfrac{1}{2}f_2) > \tfrac{1}{2}\lambda_1(f_1) + \tfrac{1}{2}\lambda_1(f_2) \qquad \textit{unless } f_1 = f_2$$

has a simple extension to excited states:

$$\lambda_{i+j-1}(\tfrac{1}{2}f_1 + \tfrac{1}{2}f_2) > \tfrac{1}{2}\lambda_i(f_1) + \tfrac{1}{2}\lambda_j(f_2) \qquad \textit{unless } f_1 = f_2 .$$

The proof is postponed in favor of an application. It has just been proved that points of M_2 can be reflections of each other *across* M_2, *i.e.*, $f_\pm = f \pm ce_2$ with all three points f, f_+, f_- on M_2. In fact, *this is the only way for two points of M_2 to have the same image* (it is conceivable that they could be reflections of each other across M_3, M_4 etc.); indeed if $Af_- = Af_+$, then $f_0 = \frac{1}{2}(f_+ + f_-)$ is singular by basic geometry, and

$$\lambda_3(f_0) > \tfrac{1}{2}\lambda_2[f_-) + \tfrac{1}{2}\lambda_2(f_+) = 0 ,$$

so that f_0 lies above M_3: It cannot lie on M_1 because $e_0 = \frac{1}{2}(f_+ - f_-)$ is proportional to the singular direction at f_0 and f_0 M_1 would mean that e_0 was of one signature, forcing one of $f_\pm = f_0 \pm ce_0$ to lie above M_1 instead of on M_2.

PROOF OF THE INEQUALITY. This employs the max-min characterization of eigenvalues of Courant-Hilbert [4]. Let V be the span of any $n-1$ independent vectors in H^0. Then (4)

$$\lambda_n(f) = \max_{V} \min_{e \in V^0} Q[e]$$

with the understanding that $\int e^2 = 1$. Let V_1 be the span of $e_k(f_1)\colon k < i$ and V_2 the span of $e_k(f_2)\colon k < j$. Then with Q_1 for f_1 and Q_2 for f_2, general V of dimension $i + j - 1$, and special $V = V_3 \supset V_1 \oplus V_2$, you have

$$\begin{aligned}\lambda_{i+j-1}(\tfrac{1}{2}f_1 + \tfrac{1}{2}f_2) &= \max_{V} \min_{e \in V^0} \tfrac{1}{2}Q_1[e] + \tfrac{1}{2}Q_2[e] \\ &\geqslant \min_{e \in V_1^0} \tfrac{1}{2}Q_1[e] + \tfrac{1}{2}Q_2[e] \geqslant \tfrac{1}{2}\min_{e \in V_2^0} Q_1[e] + \tfrac{1}{2}\min_{e \in V_3^0} Q_2[e] = \tfrac{1}{2}\lambda_i(f_1) + \tfrac{1}{2}\lambda_j(f_2) .\end{aligned}$$

The equality is not possible unless $f_1 = f_2$: in fact, the equality of line 2 and line 3 requires $e_i(f_1) = \pm e_j(f_2)$, which is not the case if $f_1 \neq f_2$.

(4) V^0 is the annihilator of V. $Q[e] = \int_0^1 [(e')^2 + fe^2]$.

11. – Diagonal form.

Berger-Church [3] proposed the following definition: a map can be brought to *diagonal form* if it can be expressed in suitable coordinates on the range and the domain as a map of the form

$$(x_1, x_2, x_3, \ldots) \to (x_1' = x_1'(x_1),\, x_2' = x_2'(x_2),\, x_3' = x_3'(x_3), \ldots)\,.$$

This can be done for the map $f \to -D^2 f + K(f)$ if the derived function $K'(f)$ is increasing and crosses only the lowest eigenvalue π^2 of $-D^2$; see Berger-Church [3]. The present function $K'(f) = f$ crosses every eigenvalue $n^2\pi$ $(n = 1, 2, 3, \ldots)$ of $-D^2$ and *cannot be brought to diagonal form.* This follows from the fact that A is not 1:1 on M_2.

PROOF. A diagonal map is singular at a point $(x_1, x_2, x_3, \ldots)$ if and only if one or more of the functions $x_1'(x_1)$, $x_2'(x_2)$, *etc.* is *bad, i.e., fails to be strictly increasing.* It follows that either just one of them is bad and the singular set is a collection of non-intersecting planes, or else $\geqslant 2$ of them are bad and the singular planes associated to the first and the second intersect. The latter is certainly not the case for $Af = f + Gf^2/2$, the singular sheets being disjoint, nor is the former: if just x_1' is bad then x_2', x_3' etc. are good and the restriction of the map to any singular plane $x_1 = c$ is 1:1.

12. – Reaching from lower sheets.

It is known, and will be reproved in yet a third way, that every point of H^1 can be reached from M_1 by the field e_1. The present article is devoted to reaching from lower sheets, e.g., from M_2 by the field e_2. The principal fact is that *if f lies below M_n, then it can be reached from M_n by the field e_n and, indeed, in at least two different ways, unlike the simple reach from M_1.*

PROOF. Let $n = 2$ for definiteness so that $\lambda_2(f_-) < 0$. It is required to find $f \in M_2$ so as to have $f_- = f - ce_2$ with $-e_1'' + fe_2'' = 0$ and $e_2(0) = e_2(1) = 0$, e_2 having one interior root. The problem can be recast in a more convenient form: $f = f_- + ce_2$ so what is needed is a solution of the (non-linear) problem

(1) $$-e'' + f_- e = -ce^2 \quad \text{with } e(0) = e(1) = 0 \text{ and one interior root.}$$

The method of continuity is used. 1) is viewed as an initial value problem with $e(0) = 0$ and $e'(0) = 1$; for $c = 0$, it has a solution with $\geqslant 2$ interior roots in conformity with the fact that $\lambda_2(f_-) < 0$; for some positive value $c = c_1$ it has a solution $e = e_1$ with $e_1(1) = 0$ and no interior root, representing the known reach from M_1. The rest is plain: as c varies roots of $e = 0$ can be created or destroyed only at $x = 1$, so there must be an interior value $0 < c_2 < c_1$ for which $e = e_2$ vanishes at $x = 1$ and has one interior root, representing a reach to f_- from M_2. An immediate obstacle to this nice plan is the fact that you may not be able to solve 1) up to $x = 1$ for general values of c: for example, if $f = 0$, you cannot solve it if $\int_0^\infty (1 + \frac{2}{3} ce^3]^{-\frac{1}{2}} de$ is smaller than 1. Fortunately you can side-step this difficulty, replacing 1) by another differential equation 2) so that 2) has solutions up to $x = 1$ for any value of c, and 1) and 2) have *the same solutions with* $e(1) = 0$ *and one interior root.* The extension to lower sheets ($\geqslant 2$ interior roots) will be plain: for example, if $\lambda_4(f_-) < 0$, then you have $\geqslant 3$ interior roots for $c = 0$, and none for $c = c_1$, and so 3, 2, or 1 interior roots for intermediate $0 < c_4 < c_3 < c_2 < c_1$ representing reaches from M_4, M_3, and M_2.

STEP 1. If $0 \leqslant c \leqslant c_0$ is small then 1) has nice solutions up to $x = 1$ and 2) $\equiv$ 1) is perfectly adequate.

STEP 2. Fix a number $c_0 > 0$, take $c \geqslant c_0$, and suppose $e = e_2$ solves 1) with $e_2(0) = 0$, $e_2'(0) = 1$, $e_2(1) = 0$, and one interior root. Then $-e_2'' + f_- e_2 = -ce_2^2$ may be multiplied by e_2' and integrated up to $0 < x < 1$ to obtain

$$-\tfrac{1}{2}(e_2')^2 + \tfrac{1}{2} + \int_0^x f_- e_2 e_2' = -\tfrac{1}{3} e_2^3 .$$

Now, at the maximum or minimum of e_2,

$$\tfrac{2}{3} ce_2^3 = -1 - \int_0^x f(e_2^2)' = -1 - fe_2^2 + \int_0^x e_2^2 f_-' ,$$

so that

$$\tfrac{2}{3} c_0 |e_2|^2 \leqslant 1 + 2\|f_-\|_1 e_2^2 ,$$

which limits the size of $\|e_2\|_\infty$ independently of the values $c \geqslant c_0$. Let $K \geqslant \|e_2\|_\infty$ be the best bound of this type.

STEP 3. The modified problem for $c \geqslant c_0$ is now declared to be

$$(2) \qquad -e'' + f_- e = -ce^2 \quad \text{if } |e| \leqslant L$$
$$= -c|e|L \quad \text{if } |e| > L\,,$$

with an adjustable cut-off L somewhat larger than K; plainly, any solution of 1) with $e(0) = 0$, $e'(0) = 1$, and $e(1) = 0$ also solves 2). It is to be proved that L *can be chosen independently of* $c \geqslant 0$ *so that any solution of* 2) *with* $e(0) = 0$, $e'(0) = 1$, $e(1) = 0$, *and one interior root lies between* $-L$ *and* $+L$ *and so solves* 1). This is easy to see. Note first that at the maximum of e, $e'' \leqslant 0$ and if $e > 0$ exceeds L, then $f_- e \leqslant -e'' + fe = -ceL$ implies $L \leqslant -f_-/c$, so that this possibility can be avoided by choice of $L \geqslant \|f_-\|_1/c_0$. Then e is limited from above between $x = 0$ and its interior root $0 < r < 1$ and $e'(r) < 0$ is also limited. Now beyond $x = r$, $-e'' + f_- e < 0$ inviting comparison to the solution of $-e''_- + f_- e_- = 0$ with the same data at $x = r$: for a while, $e_- < 0$ and

$$0 \leqslant (-e'' + f_- e)e_- - (-e''_- + f_- e_-)e = e''_- e - e_- e''\,,$$

so that $e'_- e - e_- e'$ rises from its value $(=0)$ at $x = r$ and is $\geqslant 0$, so that e_-/e rises from *its* value $(=1)$ and $e_- \leqslant e$. This shows that e lies above e_- up to $x = 1$, *independently of* c *and of* L. The modified problem 2) may now be completed by taking

$$L \geqslant \max\left[\|f_-\|_1/c_0,\, -\min_{r \leqslant x \leqslant 1} e_-(x)\right].$$

FINISHING THE PROOF. The existence of a reach from M_2 to f_- is now assured; it remains only to prove the existence of a second distinct reach. The argument differs slightly for M_3. For M_2, you redo everything starting from $x = 1$ with $e(1) = 0$ and $e'(1) = -1$ instead of from $x = 0$ with $e(0) = 0$ and $e'(0) = 1$. Replacement of e by $-e$ produces a solution of $-e''_2 + f_- e_2 = +c_2 e_2^2$ with $c_2 > 0$, $e_2(0) = e_2(1) = 0$, $e'_2(0) > 0$, and one interior root. This solution cannot be proportional to the old, *the signature of the right hand side being reversed*, so a new reach is obtained.

M_2 is handled differently since the procedure for M_2 could produce the same reach (and will if f is symmetric about $x = \frac{1}{2}$). This time, you start at $x = 1$ with $e(1) = 0$ and $e'(1) = 1$. The a priori bounds upon which the validity of the modified problem depends continue to hold: only the order of the argument of step 3 is reversed. The second M_2 reflection, rescaled, represents a solution of the present problem with one interior root and parameter $c_2 > 0$. Now, for $c = 0$, $\lambda_3(f_-) < 0$ implies that the solution has $\geqslant 3$ interior roots, so there is an intermediate parameter $0 < c_3 < c_2$ for which $e = e_3$ has $e_3(0) = 1$ and 2 interior roots; naturally, $e_3'(0) < 0$. Replacement of e_3 by $e_3/e_3'(0)$ restores the original $e_3'(0) = 1$ and flips the signature of c_3 from $+$ to $-$ producing the original reach.

AMPLIFICATION 1. If f_- is symmetric about $x = \frac{1}{2}$, the second reach from M_2 or M_3 is produced from the first by replacement of $e = e_2$ or e_3 by $-e(1-x)$.

AMPLIFICATION 2. Let f_- lie below M_n. Then it can be reached from M_k in $\geqslant 2$ different ways for $k = 2, \ldots, n$ and in 1 way from M_1, as in fig. 9 [$n = 3$]. The reflected points such as f_+ situated at the ends of the dotted lines all map to the single point in which the solid lines indicate reaches, by basic geometry, and all these 6 solutions of $Af = g$ are distinct, the general result being that *$N(g) \geqslant 2n$ if the lowest point of the preimage lies below M_n.* It follows that there are points (such as f_+) above M_1 which can be reached from M_1 in $\geqslant 2n = 4, 6, 8, \ldots$ ways.

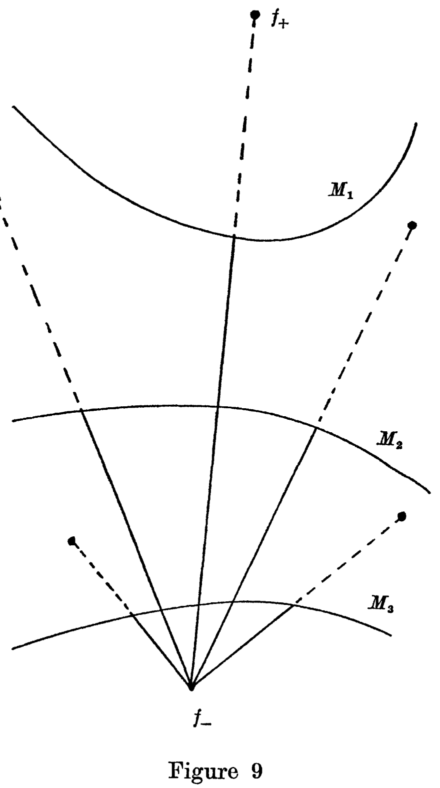


Figure 9

Lazer-Mc Kenna [6] prove a similar result for the equation $g = -D^2 f + K(f)$ with increasing $K'(f)$ crossing the first n eigenvalues of $-D^2$ *if g is sufficiently high up, then $N(g) \geqslant 2n$.*

AMPLIFICATION 3. The count $2n$ is found to be exact in the example of art. 14 but must be higher in general: for example, if f_- is just a little above a point $f_4 \in M_4$ where $\int_0^1 e_4^3 \neq 0$, then it can be reached from M_4 since the differential (5) $I - cF_4^{-1}\pi e_4$ of the map $f_4 \to f_4 + ce_4$ of $M_4 \times \mathbb{R}$ into H^1 is invertible. This raises the count to $7 = 2n + 1$, and it could go much higher; indeed, the situation seems very complicated in view of the presence of cusps and higher singularities.

AMPLIFICATION 4. The example of art. 14 has $g = c\boldsymbol{G}1 = (c/2)x(1-x)$. The preimage satisfies $-f'' + f^2/2 = c$ which may be integrated explicitly by means of simple elliptic functions. The lowest preimage should be approximately $f = -\sqrt{2c}$ and should cross the n-th sheet at $c = n^4\pi^4/2$ about, for a rough count of $N(g) = 2n = 2^{\frac{5}{4}} c^{\frac{1}{4}} \pi^{-1}$. This leads to the

CONJECTURE. *$N[g + c\boldsymbol{G}1] \sim$ universal constant $\times\, c^{\frac{1}{4}}$ as $c \uparrow \infty$.*

This would follow if the count were not too sensitive to the details of g as $c \uparrow \infty$; actually no upper bound to $N(g)$ seems to be known, so this is an attractive question.

HOW FAR CAN YOU REACH? The whole of H^1 can be reached from M_1. What about lower sheets M_2 *etc.* Let R_n be the points that can be reached from M_n and note that it is filled up by lines $f + \mathbb{R}e_n(f)$ with $f \in M_n$. It contains everything below M_n. The question is: *how far up can you reach?*

PROPOSITION 1. *R_n is closed.*

PROOF. Let $f_+ = f + ce_n$ with $f \in M_n$, $-\infty < c < \infty$, and $\int_0^1 e_n^2 = 1$, and let f_+ converge in H^1. Then Af_+ converges, and $f_- = f - ce_n$ is compact since $Af_- = Af_+$ and A is a proper map. The rest is routine.

PROPOSITION 2. *$R_n \subset R_{n-1}$ for every $n \geqslant 2$.*

Let $f_- = f - ce_n \in R_n$ be reached from $f \in M_n$ with $c > 0$. Then $-e_n'' + f_- e_n = -ce_n^2$ with $e_n(0) = e_n(1) = 0$ and $n - 1 \geqslant 1$ interior roots. Adjustment of c produces a solution with $n - 2$ interior roots, representing a reach from M_{n-1}. The proof is finished.

(5) π is the co-projection $I - e_4 \times e_4$.

PROPOSITION 3. *R_2 falls short of M_1.*

PROOF. Let f_+ above or on M_1 be reached from $f \in M_2$: $f_+ = f + ce_2$. The reflected point $f_- = f - ce_2$ lies below M_1 since $Af_- = Af_+$ and the map is 1:1 above M_1: But then f_- can be reached from M_1. The corresponding reflection of f_- *across* M_1 represents a new preimage of Af_+ *above* M_1 and that is not possible.

PROPOSITION 4. R_n contains an open neighborhood of M_n punctured along the cubic locus L_1 where $\int_0^1 e_n^3 = 0$.

PROOF. As in amplif. 3 above.

AMPLIFICATION 5. The precise upper boundary even of R_2 seems to be complicated. It could touch M_2 though only along the cubic locus; this is a moot point. Besides, ∂R_2 could have finite-dimensional corners (but not worse) at places that can be reached from M_2 in $\geqslant 2$ different ways: 2 different reaches could produce 2 transversal half-patches in ∂R_2. It is not even known if R_2 is *connected* or if it contains everything below ∂R_2 (whatever that means). It would be pleasant if ∂R_2 were convex, but that is not clear either. The points above M_1 which can be reached from M_1 in 1, 2, 3, ... distinct ways present a dual class of geometrical questions about which nothing much is known. The model of amplif. 8.2 might be helpful in these matters.

13. – Curvature of the singular sheets.

The principal curvatures of the sheets (and the associated directions) can be computed following McKean [7]: for M_1, *they are of one signature, so that the sectional curvatures of M_1 are all positive*, as was known already from its convexity; *for M_2* there enters one principle direction of opposite curvature; *for M_3, there enter two such*; *for M_4, three*; *and so forth.* M_1 and M_2 are seen in fig. 10. This is seen by computing the second fundamental form of the sheet in the ambient space H^1: it is a compact self adjoint operator in TM whose eigenvalues are the principal curvatures.

PROOF. It is simplest to use the H^0 format. Let $n(f) = -\, e_n^2(f)/\left(\int_0^1 e_n^4\right)^{-\frac{1}{2}}$ be the downward-pointing unit normal at $f \in M_n$ and let a and b be tangent vectors at f: $\int_0^1 ae_n^2 = \int_0^1 be_n^2 = 0$. The second fundamental form is $[a, b]$

$=\int_0^1 adnb$, in which dnb is the derivative of $n(f)$ in the direction b. This is easy to compute:

$$dnb = -2e_n\dot{e}_n\left(\int_0^1 e_n^4\right)^{-\frac{1}{2}} + \tfrac{1}{2}e_n^2\left(\int_0^1 e_n^2\right)^{-\frac{3}{2}} \times 4\int_0^1 e_n^3\dot{e}_n\,,$$
$$= 2e_n F^{-1}(e_n b)\left(\int_0^1 e_n^4\right)^{-\frac{1}{2}} + ce_n^2\,,$$

and $\int_0^1 ae_n^2 = 0$ permits the second piece to be ignored:

$$[a, b] = 2\left(\int_0^1 e_n^4\right)^{-\frac{1}{2}} \int_0^1 (e_n a)F^{-1}(e_n b);$$

the normalization is adopted below so that this integral drops out of the form. Now the (compact, self-adjoint) operator representing this form in TM_n is $\dot{f} \to 2\pi e_n F^{-1}(e_n\dot{f})$ in which π is the co-projection $I - e_n^2 \otimes e_n^2\left(\int_0^1 e_n^4\right)^{-1}$, *i.e.* the projection of the ambient space onto TM_n. Let $\dot{f} \in TM_n$ be an eigenvector with eigenvalue μ so that

$$2\pi e_n F^{-1}(e_n\dot{f}) = \mu\dot{f} \quad \text{with } \int_0^1 e_n^2\dot{f} = 0\,,$$

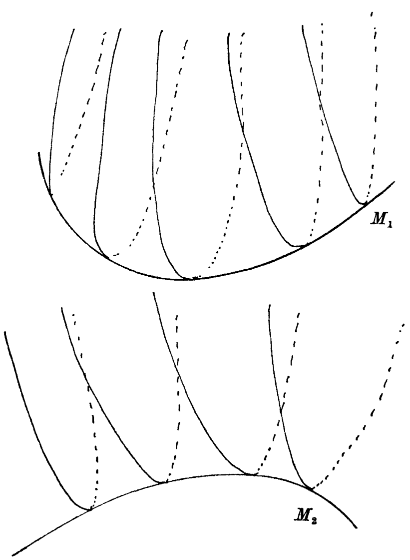


Figure 10

or what is the same as without the π, $2e_n F^{-1}(e_n \dot{f}) + ce_n^2 = \mu \dot{f}$. Let $e = 2F^{-1}(e_n \dot{f}) + ce_n$ so that $\dot{f} = ee_n/\mu$, noting that μ cannot vanishes unless $e = 0$ and $0 = Fe = 2e_n \dot{f}$ which is not the case. It follows that $2F^{-1}(ee_n^2/\mu) + ce_n = e$, whence $Fe = 2ee_n^2/\mu$, which is to say $-e'' + (f - 2e_n^2/\mu)e = 0$. This shows that the eigenvalues $\mu = \mu_m$ $(m \neq n)$ of the second fundamental form are determined by the rule:

$$\lambda_m(f - 2e_n^2/\mu) = 0 \qquad \text{for } m = 1, 2, 3, \ldots \neq n\,,$$

the corresponding principal directions being

$$\dot{f}_n = ee_n/\mu = \frac{1}{\mu}\, e_m \left(f - \frac{2e_n^2}{\mu}\right) e_n$$

with $e_n = e_n(f)$. The geometry is self-evident: $f - 2e_n^2/\mu$ lies above f if $\mu < 0$ and below if $\mu > 0$; plainly it crosses every singular sheet, M_n excepted; *in particular it has crossing at precissly* $n - 1$ *negative eigenvalues of* μ. The proof is finished.

AMPLIFICATION 1. $2/\mu_n$ is just the (signed) distance from $f \in M_n$ to the sheet M_m $(m \neq n)$ in the direction $-e_n^2$.

AMPLIFICATION 2. $2/\mu_m$ tends to 0 as $m \uparrow \infty$, so most of the principal curvatures are small; indeed, $\mu_m = O(m^{-2})$ by standard estimates and the scalar curvature $\varkappa = \sum_{i<j} \mu_i \mu_j$ is finite (6), the upshot being that the sheets are pretty flat.

14. – Example: The preimage of a line.

$G1 = \frac{1}{2}x(1-x)$, so the solutions of $Af = (a/2)x(1-x)$ satisfy $-f'' + f^2/2 = a$ with $f(0) = f(1) = 0$ and can be computed explicitly by means of simple elliptic functions: in short, the whole preimage of the line $\mathbb{R}x(1-x)$ can be found. It is convenient to rescale so as to have $-f'' + 6f^2 = a/2$. Then, $(f')^2 = 4f^3 - af + b$ with a constant of integration b, this being the Weierstrass $\wp$-function. It is required to adjust the value of b

(6) It is evaluated explicitly for M_1 in McKean [7].

so that $\wp$ is real on a horizontal segment of $\mathbf{C}$ of length 1 and vanishes at the ends; plainly, this cannot be done if a is so negative that $aG1$ falls below $A(M_1)$ in the original scale; contrariwise, it can be done in many ways if $a > 0$ is large. The details are outlined below. Let $2\omega_1$ and $2\omega_2$ be the primitive periods of $\wp$ and let $e_1 = \wp(\omega_1)$, $e_2 = \wp(\omega_2)$, $e_3 = \wp(\omega_1 + \omega_2)$ be the roots of the cubic: $4\wp^3 - a\wp + b = 4(\wp - e_1)(\wp - e_2)(\wp - e_3)$.

$$e_1 + e_2 + e_3 = 0\,, \tag{1}$$

$$e_1 e_2 + e_2 e_3 + e_3 e_1 = -\frac{a}{4}, \tag{2}$$

$$e_1 e_2 e_3 = -\frac{b}{4} \tag{3}$$

are noted for future use. Now if $\wp$ is real on a horizontal line passing through its fundamental cell, then its poles must be conjugate symmetric across that line. This permits only three possibilities, seen in fig. 11, in which the dark segment is where $\wp$ is to be considered: there $e_1 > e_3 > e_2$ and 1) forces $e_1 > 0$, so that on the dark segment $\wp \geqslant e_1$ has no roots at all. This leaves the second rectangular case and the rhombic case to be looked into.

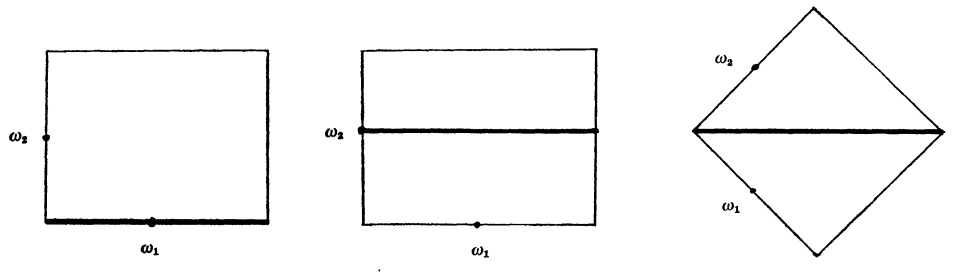


Figure 11

Rectangular case. $\wp$ is required to vanish on the dark segment and the distance between *some* two roots of $\wp = 0$ on the extended horizontal line is required to be 1.

Step 1. $e_2 < 0$ by 1) and $e_2 \leqslant \wp \leqslant e_3$ is required to vanish between ω_2 and ω_3, which makes $e_3 \geqslant 0$; $\wp$ also vanishes symmetrically to the other side of ω_3 and *no place else*, being of degree 2. The square of 1) compared with 2) yields $e_1^2 + e_2^2 + e_3^2 = a/2$ so $a > 0$; next, 1) implies $e_1 > |e_3| > - e_2$, so 2) yields $e_3^2 \leqslant a/12$; and similar such considerations lead to the final limitations on e_1, e_2, e_3.

$$-\sqrt{a/3} \leqslant e_2 \leqslant -\sqrt{a/4}\,, \quad 0 \leqslant e_3 \leqslant \sqrt{a/12} \leqslant e_1 \leqslant \sqrt{a/4}\,. \tag{4}$$

The choice of a and any one of e_1, e_2, e_3 between the indicated limits determines all the other parameters; especially, $b \geqslant 0$ by 3).

STEP 2. I_1 denotes $2\omega_1$, while I_2 and I_3 denote the distance between consecutive roots of $\mathcal{T} = 0$, with $\mathcal{T} < 0$ between for the first, and $\mathcal{T} > 0$ between for the second. $I_1 = I_2 + I_3$ with

$$I_2 = \int_{e_2}^{0} [(e_1 - p)(p - e_2)(e_3 - p)]^{-\frac{1}{2}}\, dp$$

and

$$I_3 = \int_{0}^{e_3} [(e_1 - p)(p - e_2)(e_3 - p)]^{-\frac{1}{2}}\, dp\,;$$

if any one of nI_1, $nI_1 + I_2$, or $nI_1 + I_3$ takes the value 1 for $n = 0, 1, 2, 3, \ldots$, then a solution of $-f'' + bf^2 = a/2$ with $f(0) = f(1) = 0$ is obtained; compare fig. 12.

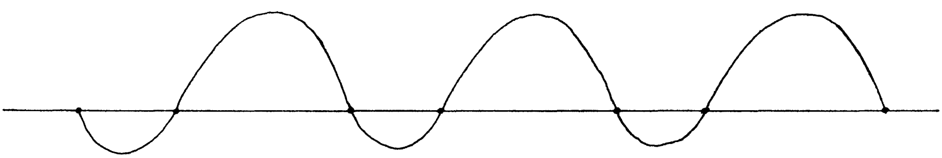


Figure 12

Note that two distinct solutions are obtained from $I_1 = 1$, $2I_1 = 1$, $3I_1 = 1$, *etc.*, by a self-evident reflection, but only 1 from any other combination.

STEP 3. Let $p = re_3$ in I_3 and eliminate e_1 and e_2 in favor of e_3 using 1) and 2):

$$I_3 = e_3^{-\frac{1}{2}} \int_0^1 [(1 - r)(a/4e_3^2 - r^2 - r - 1)]^{-\frac{1}{2}}\, dr.$$

I_1 and I_2 can be similarly expressed: in terms of the 3 new variables

$\varkappa = a/4e^2$ $(e = e_1, e_2, e_3)$, you find

$$I_1(\varkappa_1, a) = \sqrt{2}\, a^{-\frac{1}{4}} \varkappa_1^{\frac{1}{4}} \int_0^\infty [r(r^2 + 3r + 3 - \varkappa_1)]^{-\frac{1}{2}}\, dr \qquad (1 \leqslant \varkappa_1 \leqslant 3)$$

$$I_2(\varkappa_2, a) = \sqrt{2}\, a^{-\frac{1}{4}} \varkappa_2^{\frac{1}{4}} \int_0^1 [(1 - r)(r^2 + r + 1 - \varkappa_2)]^{-\frac{1}{2}} dr \quad (\tfrac{3}{4} \leqslant \varkappa_2 \leqslant 1)$$

$$I_3(\varkappa_3, a) = \sqrt{2}\, a^{-\frac{1}{4}} \varkappa_3^{\frac{1}{4}} \int_0^1 [(1 - r)(\varkappa_3 - r^2 - r - 1)]^{-\frac{1}{2}}\, dr \quad (3 \leqslant \varkappa_3 < \infty)\,.$$

The number $a > 0$ is viewed as a parameter and $1 \leqslant \varkappa_1 \leqslant 3$ is taken as the basic variable: $\varkappa_2$ and $\varkappa_3$ can be found from $2e_1 = \sqrt{a/\varkappa_1}$ and the associated values of e_2 and e_3. It is helpful for the next step to introduce the extreme values of I_1, I_2, I_3 corresponding to $\varkappa_1 = 1$ and $\varkappa_1 = 3$:

at $\varkappa_1 = 3$, $\varkappa_2 = \frac{3}{4}$, $\varkappa_3 = 3$,

$$I_1 = \infty\,,$$

$$I_2 = 3^{\frac{1}{4}} a^{-\frac{1}{4}} \int_0^1 (1 - r)^{-\frac{1}{2}} (r + \tfrac{1}{2})^{-1}\, dr \equiv a^{-\frac{1}{4}} J_1\,,$$

$$I_3 = \infty\,;$$

at $\varkappa_1 = 1$, $\varkappa_2 = 1$, $\varkappa_3 = \infty$,

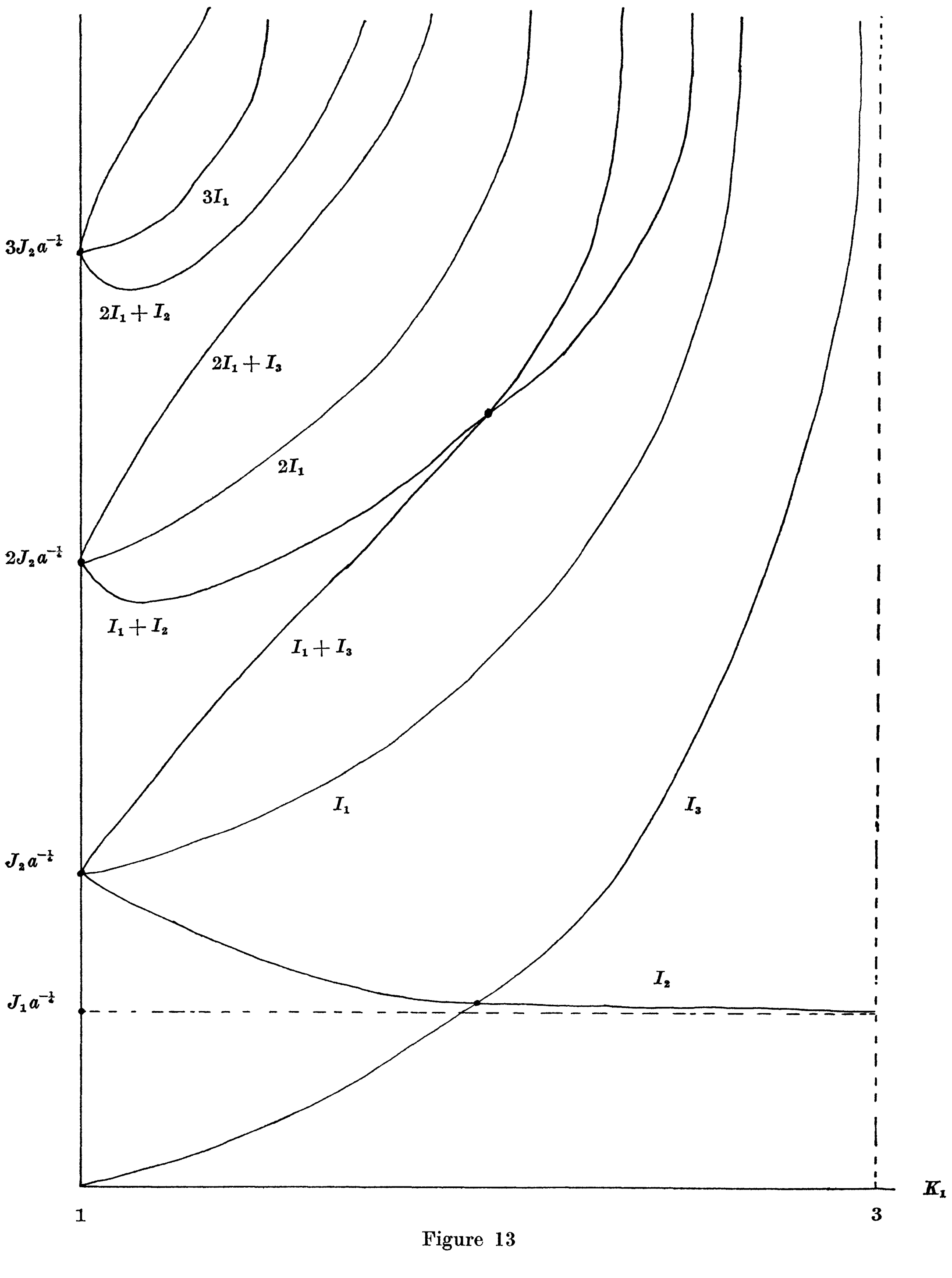
$$I_1 = 2^{\frac{1}{2}} a^{-\frac{1}{4}} \int_0^\infty [r(r + 1)(r + 2)]^{-\frac{1}{2}}\, dr \equiv a^{-\frac{1}{4}} J_2\,,$$

$$I_2 = 2^{\frac{1}{2}} a^{-\frac{1}{4}} \int_0^1 [(1 - r)\, r(r + 1)]^{-\frac{1}{2}}\, dr \equiv a^{-\frac{1}{4}} J_3\,,$$

$$I_3 = \infty\,.$$

$J_2 = J_3$ is seen by the substitution $r \to 1/r - 1$. J_1 is smaller.

STEP 4 is the justification of fig. 13 which displays the several combinations of I_1, I_2, I_3 as functions of $1 \leqslant \varkappa_1 \leqslant 3$ for fixed $a > 0$. The number of solutions of $-f'' + 6f^2 = a/2$ with $f(0) = f(1) = 0$ is found by counting the number of curves crossing the horizontal level 1, keeping in mind that $I_1 = 1$, $2I_1 = 1$, $3I_1 = 1$, *etc.* produce two solutions. The lengthy details are carried out in full in Scovel [8].



$3I_1$
$3J_2a^{-\frac{1}{2}}$
$2I_1+I_2$
$2I_1+I_3$
$2I_1$
$2J_2a^{-\frac{1}{2}}$
I_1+I_2
I_1+I_3
I_1
I_3
$J_2a^{-\frac{1}{2}}$
I_2
$J_1a^{-\frac{1}{2}}$
K_1
1
3

Figure 13

Rhombic case. e_1 and e_2 are complex conjugates, e_3 and $\omega_3 = \omega_1 + \omega_2$ are real, and $\wp$ is real on $\mathbb{R}$: it tends to ∞ at 0 and at $2\omega_3$ and has a minimum e_3 between; in particular, we need $e_3 < 0$, in which case $\wp = 0$ has two roots symmetrically placed about ω_3 and separated by a distance

$$I_2(\varkappa, a) = \sqrt{2}\, a^{-\frac{1}{4}} \varkappa^{\frac{1}{4}} \int_0^1 [(1-r)(r^2 + r + 1 - \varkappa)]^{-\frac{1}{2}} dr$$

with [7] $\varkappa = a/4e_3^2 \leqslant \frac{3}{4}$. Each root of $I_2(\varkappa, a) = 1$ produces one and only one solution to our problem.

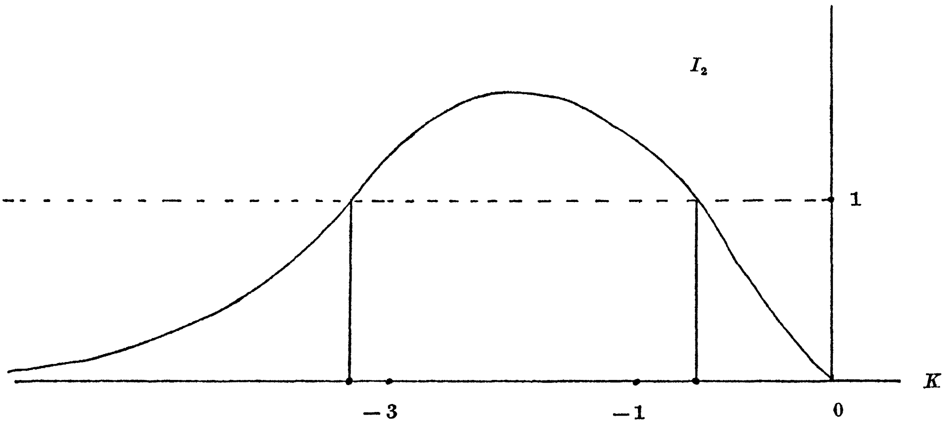


Figure 14

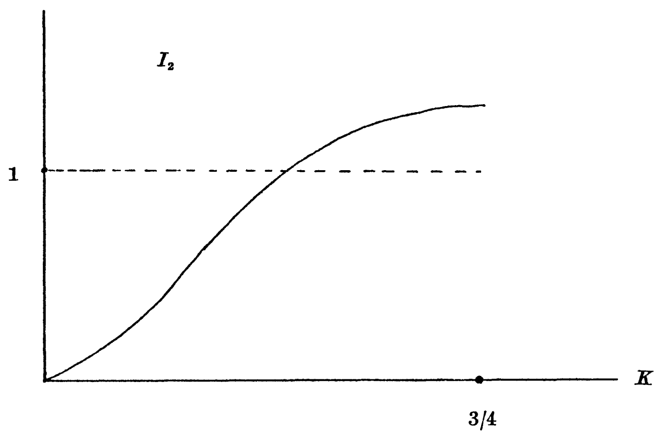


Figure 15

[7] The limitation $e_3^2 \geqslant \varkappa/3$ is inherent in the rhombic case.

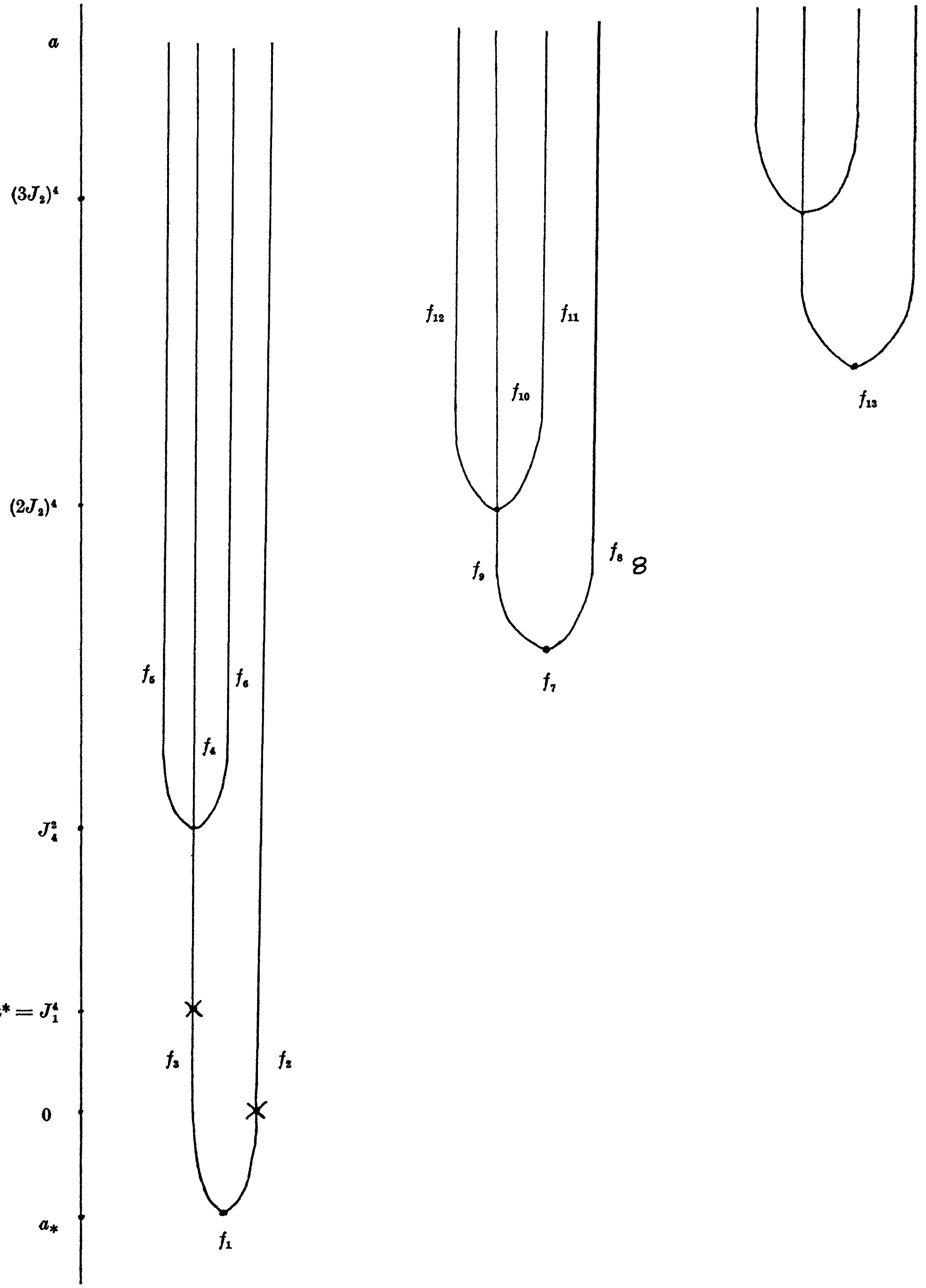


Figure 16

Figure 14 depicts I_2 for moderate values of $a < 0$: for large $a < 0$, the whole graph lies below the level 1 and no solutions are obtained; then it rises and produces at $a = a_*$, first one and later two solutions. The roots of $I_2 = 1$ tend to $-\infty$ and 0 as $a \uparrow 0$. The solution arising from the lefthand root tends to 0 with a and passes smoothly into the solution attached to $I_3 = 1$ in fig. 13. The picture is different for $a > 0$: I_2 increases from 0 to a maximum $3^{\frac{1}{4}} a^{-\frac{1}{4}} \int_0^1 (1-r)^{\frac{1}{2}} (r+\frac{1}{2})^{-1} dr$ at $\varkappa = \frac{3}{4}$ and for a not exceeding

$$a^* = 3\left[\int_0^1 (1-r)^{-\frac{1}{2}} (r+\tfrac{1}{2})^{-1} dr\right]^4 = J_1^4,$$

there is a unique root of $I_2 = 1$ representing the continuation *through* $\varkappa = 0$ of the right-hand root of fig. 14. The number $a^* = J_1^4$ is the asymptote of I_2 in fig. 13, and as a passes a^* the surviving rhombic case solution passes smoothly into the $I_2 = 1$ solution of the rectangular case.

SUMMARY. If $a < a_*$, there is no solution at all; at $a = a_*$, a rhombic solution appears, which splits in two as a increases further; one branch passes smoothly into the rectangular solution coming from $I_3 = 1$ in fig. 13 as a crosses 0 and persists from then on, while the other persists until $a = a^* = J_1^4$ when it changes into the rectangular solution coming from $I_2 = 1$. The latter persists until $a = J_2^4$ when it splits in three (8); these branches continue thereafter. The pattern repeats itself, starting a little below $a = (2J_2)^4$: at that moment, a new solution appears which immediately splits into two branches, one persisting from then on and the other splitting in three at $a = (2J_2)^4$; further solutions appear a little below $a = (3J_2)^4$, *etc.* Figure 16 tells the story more concisely; the little crosses indicate the transition from the rhombic case to the rectangular.

AMPLIFICATION. It is instructive to follow the motion of the preimages f_1, f_2, f_3 *etc.*, relative to the singular sheets; the discrepancy between the present scaled problem $-f'' + 6f^2 = a/2$ and the original $-f'' + f^2/2 = a$ is ignored. The bottom solution f_1 makes its debut at $a = a_*$; it lies on M_1 and immediately splits into f_2 and f_3 owing to the local folding along M_1; f_2 stays above M_1 forever, while f_3 moves down and splits into f_4, f_5, f_6 at $a = J_2^4$. This is when f_3 arrives at M_2. The split is 3-fold, so f_3 is not a fold point: instead it must be in the cubic locus $\int_0^1 e_2^3 = 0$; compare fig. 17, in which the lines joining f_4, f_5, f_6 represent reflections across M_2.

(8) $I_1 = 1$ produces two solutions, not one.

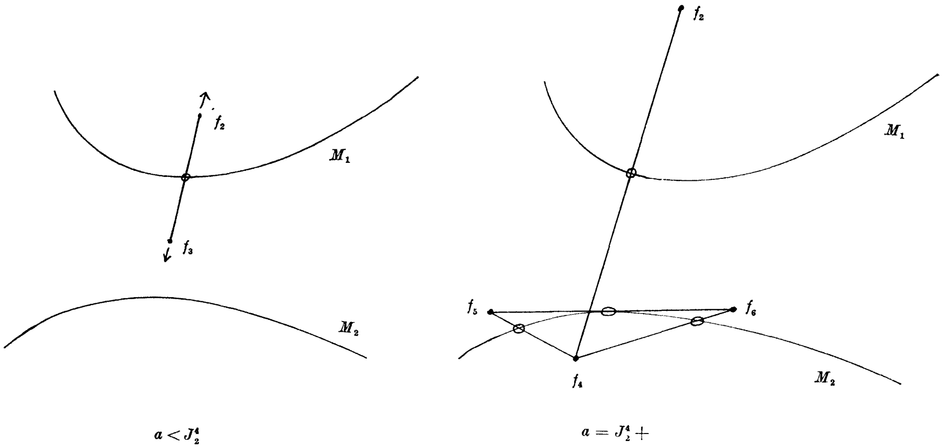


Figure 17

Let us confirm this picture, *to wit*: *f_4 lies below M_2 while f_5 and f_6 lie above.*

PROOF. f_4 comes from $I_1 + I_3 = 1$, which is to say that $0 \leqslant x \leqslant 1$ comprises a full period between three consecutive roots of $p = 0$ plus an interval where $p > 0$; see fig. 12. Let $0 < a < b < 1$ be the roots of $f_4' = 0$ closest to $x = 0$ and to $x = 1$. Then $-f_4'' + 6f_4^2 = a/2$ implies that $-e'' + 12 f_4 e = 0$ with $e = f_4'$, and since e has one interior root between $x = a$ and $x = b$, the restricted eigenvalue $\lambda_*^2(12f_4)$ vanishes. This proves that the unrestricted eigenvalue $\lambda_2(12f_4)$ is < 0, so that f_4 lies below M_2 (9). Now f_5, f_6 is the pair of solutions coming from $I_1 = 1$. They are mirror images of each other about $x = \frac{1}{2}$, so they lie on the same side of M_2. The map A is a compact perturbation of the identity to which the index (= local degree) of Leray-Schauder applies (see Berger [2]); in particular, the index (-1) of f_3 is the sum of the like indices of f_5 and f_6 and the inhex $(+1)$ of f_4: in short both f_5 and f_6 have index -1 and so lie above M_2.

What happens next? f_4, f_5, f_6 do not merge or split after $a = J_2^4$ so they do not touch any singular sheets but remain: f_4 between M_2 and M_3, and f_5, f_6 between M_1 and M_2, as in fig. 18. The next event is that f_7 suddenly appears at a fold point of M_3: $\lambda_2(12f_7) < 0$ is proved as for f_4 and $\lambda_4(12f_7) > 0$ similarly by considering the roots of f_7' next below $x = 0$ and next above $x = 1$. Then f_7 splits into f_8 above M_3 and f_9 below, of which the former stays between M_2 and M_3, while the latter hits M_4 at a point of its cubic locus and splits in 3. The pattern repeats itself from then on; in particular, *if*

(9) The 12 is to be ignored: it is an artifact of the scaling.

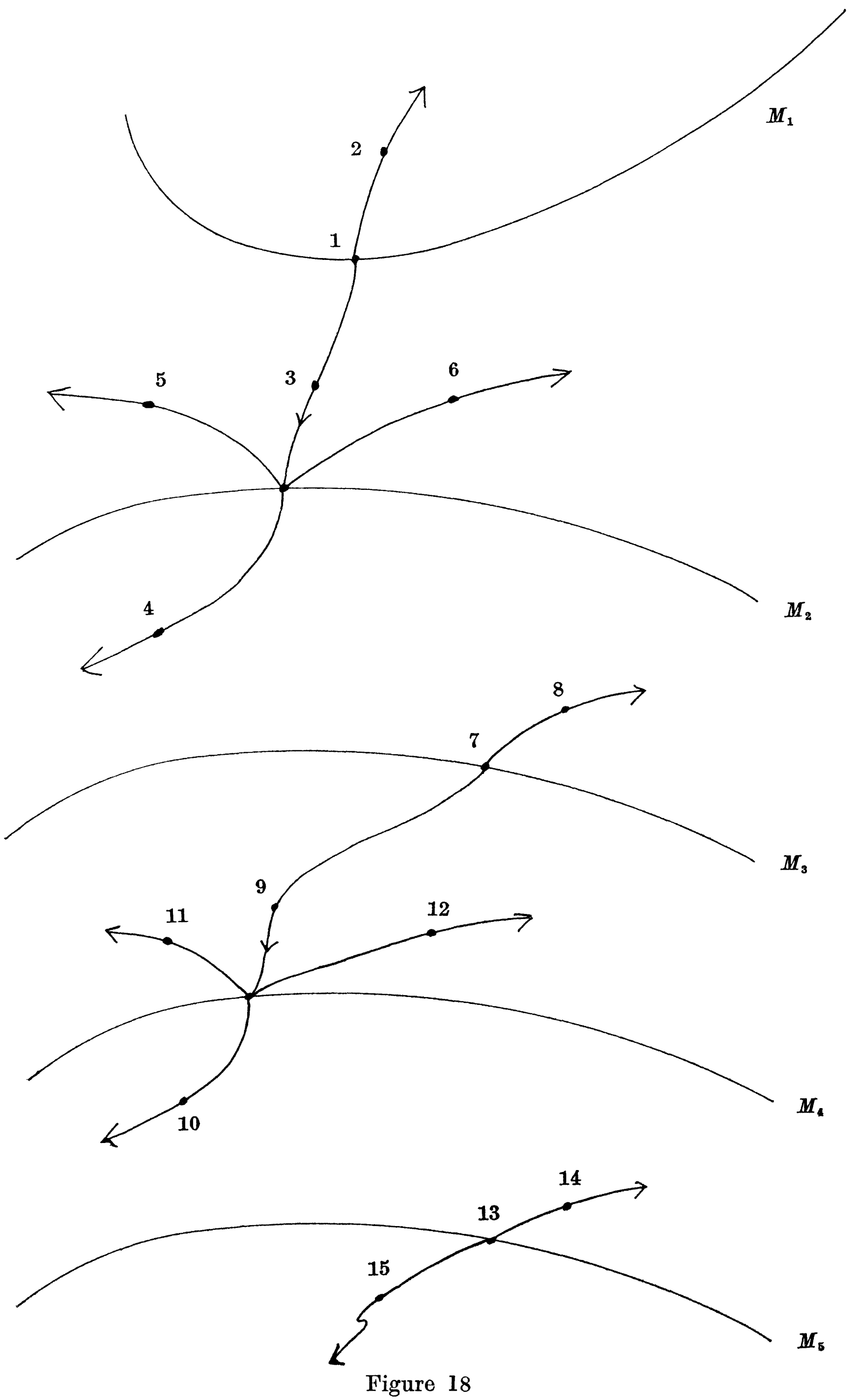


Figure 18

the lowest f lies between M_n and M_{n+1}, then there are exactly two more f's between M_i and M_{i+1} for $1 \leqslant i < n$ and still another above M_i for a total count of $2n$, as was announced before.

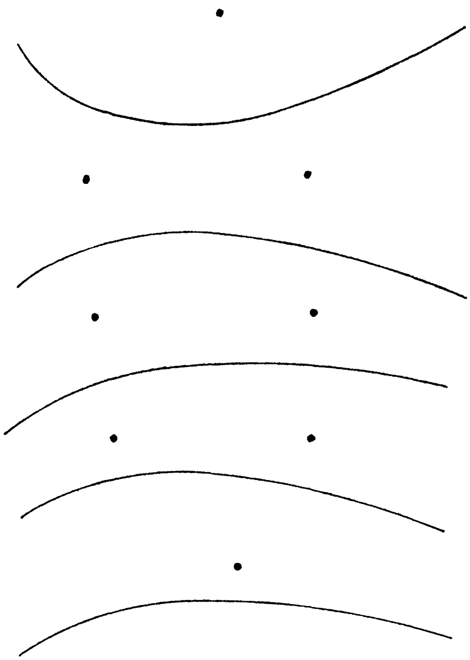


Figure 19

BIBLIOGRAPHY

[1] A. Ambrosetti - G. Prodi, *On the inversion of some differentiable maps between Banach spaces with singularities*, Ann. Mat., **93** (1972), 231-246.

[2] M. S. Berger, *Non-linearity and Functional Analysis*, Academic Press, N.Y.C., 1977.

[3] M. S. Berger - P. T. Church, *Complete integrability and perturbation of a nonlinear Dirichlet problem (I)*, Indiana Univ. Math. J., **28** (1979) 935-952.

[4] R. Courant - D. Hilbert, *Methods of Mathematical Physics*, vol. **1**, Interscience Publishers, New York, 1953.

[5] M. Golubitski - V. Guillemin, *Stable Mappings and their Singularities*, Springer-Verlag, N.Y.C., 1973.

[6] A. C. Lazer - P. J. McKenna, *On a conjecture related to the number of solutions of a nonlinear Dirichlet problem*, Proc. Roy. Soc. Edinburgh, to appear.

[7] H. P. McKean, *Curvature of an ∞-dimensional manifold related to Hill's equation*, J. Differential Geom., **17** (1982), 523-529.

[8] J. C. SCOVEL, *Geometry of some nonliner differential operators,* Ph.D. Thesis, New York University, 1983.
[9] J. C. SCOVEL, *Geometry of some simple nonlinear differential operators,* part II, sub. to J.D.E. Mar 1985
[10] H. WHITNEY, *On singularities of mappings,* Ann. of Math. (1955), 374-416.

Courant Institute
New York, New York,
Brandeis University
Waltham, Massachusetts