ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 24, Number 4, Fall 1994

THE RIEMANN-HILBERT PROBLEM FOR SINGULAR POSITIVE LOOPS

EDWIN J. BEGGS

0. Introduction. Let $g : S^1 \to M_m(\mathbf{C})$ be a matrix valued function on the unit circle in $\mathbf{C}$. The Riemann-Hilbert problem is concerned with factorizing g into the product of two matrix valued loops, one analytic inside the circle, and the other analytic outside. This problem arises naturally in the inverse scattering procedure for solving some nonlinear partial differential equations. In this method the Cauchy data for the system is converted into the scattering matrix $S(\lambda)$, a matrix valued function of a parameter λ taking values in $\mathbf{R}_\infty \cong S^1$. From this we can calculate the reflection $g(\lambda, x, t)$, which is a loop valued function on space-time, and the corresponding analytic factors are solutions of the linear system of an integrable classical field equation [**5, 6**]. However, in the case of the chiral equation there is a complication; the loop g may not exist for all scattering data. To be more specific, the inverse scattering procedure supplies a formula for the reflection at $x = t = 0$ of the form $g = qp^{-1}$, where q and p are loops calculated by upper and lower triangular factorization of the scattering. But there is no reason why p should be invertible on the circle. This problem has an interesting physical interpretation; these singular reflection loops arise in the transitional cases between solutions with different numbers of solutions. To find such a state all we have to do is to continuously deform the initial data for a solution with one number of solutions to initial data for another number of solutions. To examine the topology of the space of solutions to the chiral equation we should therefore pay some attention to these degenerate cases. In the chiral equation the situation can be partly remedied by defining a positive modified reflection $\tilde{g} = q\bar{q}$, but this may still fail to be invertible on the circle. This will be explained further in a future paper [**3**]. This brings us to the subject of the present paper, which is to solve the Riemann-Hilbert problem for positive loops which may fail to be invertible. The answer we shall obtain is that a continuous positive loop can be factored if its

Received by the editors on October 21, 1992.

log (in the sense of the calculus of normal operators) is integrable on the circle.

In the references [**7**] and [**11**] the problem is approached from the point of view of stochastic processes. In [**1**] the problem of L^∞ continuity of the factorization is considered in the nonsingular case by putting L^2 bounds on the derivatives of the positive loops being factorized.

The emphasis of the present paper is on an elementary method of solving the nonsingular case, in which the continuity and smoothness of the factorization are obvious. In addition, the inverse scattering method requires the factors to be evaluated at particular points on the circle. Some care is required in this, as there are continuous loops which factor into discontinuous ones [**8**]. The problem of loops which can be evaluated at points is also treated by elementary functional analysis. The somewhat strange class of loops which can be evaluated at a point used in the paper is justified on two grounds. Firstly, it is the natural class for the functional analytic proof. Secondly, the fact that it is more general than the usual condition of differentiability at a point considerably eases the analytical difficulties in the inverse scattering procedure. The points where the evaluation must take place are precisely the points where the scattering is most likely to be badly behaved. Again there is a simple corollary that the factorization process is smooth.

The singular case is treated as a limit of the nonsingular case. In this case, the space of loops is not a smooth manifold, but the continuity result still holds if the correct topology is used.

1. Definitions of various operators on $L^2(S^1, M_m(\mathbf{C}))$. We take the space L^2 to consist of the Lebesgue square integrable functions from the circle to the $m \times m$ complex matrices, with the usual inner product

$$\langle f, k\rangle = \oint_{S^1} \mathrm{Trace}\,(f(z)k(z)^*). \tag{1.1}$$

(Conventionally, we take the integral with respect to angular measure on the circle giving it length 2π.) The norm $||\cdot||_2$ on L^2 is the usual L^2 norm of the pointwise matrix operator norm on the circle. This is equivalent to the norm derived from the inner product. Similarly L^1 is defined using the L^1 norm of the matrix operator norm.

The space $L^\infty(S^1, M_m(\mathbf{C}))$ is defined to be the space of essentially bounded measurable matrix valued functions, with norm $||\cdot||_\infty$ the essential supremum of the matrix operator norm.

Then L^+ can be defined to be the closed subspace of L^2 consisting of those functions which have only got positive powers of z (including z^0) in their Laurent series, and L^- as the functions with only negative powers (including z^0).

Definition 1.2. Define the following maps to the set of holomorphic functions on the unit disk $\Delta_+ = \{z \in \mathbf{C} : |z| < 1\}$ and $\Delta_- = \mathbf{C}_\infty - \overline{\Delta_+}$. (Here $\mathbf{C}_\infty$ is the Riemann sphere.)

$$\pi_+ : L^1(S^1, M_m) \to \mathrm{Hol}\,(\Delta_+, M_m)$$

and

$$\pi_- : L^1(S^1, M_m) \to \mathrm{Hol}\,(\Delta_+, M_m).$$

$$\pi_+(f)(z) = \frac{1}{2\pi}\oint \frac{\zeta}{\zeta - z} f(\zeta) \centerdot d\mu_\zeta \qquad z \in \Delta_+ \tag{1.2a}$$

$$\pi_-(f)(z) = \frac{-1}{2\pi}\oint \frac{z}{\zeta - z} f(\zeta) \centerdot d\mu_\zeta \qquad z \in \Delta_- \tag{1.2b}$$

using angular measure μ to integrate over $\zeta \in S^1$. These functions have a particularly simple form when expressed in terms of the Laurent series (1.2c)

$$f(z) = \sum_{n\in\mathbf{Z}} f_n z^n, \qquad \pi_+ f(z) = \sum_{n\geq 0} f_n z^n, \qquad \pi_- f(z) = \sum_{n\leq 0} f_n z^n.$$

From these expressions it is automatic that $\pi_+ : L^2 \to L^+$ and $\pi_- : L^2 \to L^-$ are continuous.

Definition 1.3. If f is a matrix valued function analytic on Δ_+ (respectively, Δ_-) define $\bar{f}$, a function analytic on Δ_- (respectively, Δ_+), by the formula $\bar{f}(z) = (f(1/\bar{z}))^*$. Note that this bar construction gives an antilinear isomorphism between L^+ and L^-.

Proposition 1.4. *The operations of left and right multiplication by L^∞ loops on L^2 are continuous, and the corresponding left (L) or right (R) "multiply by" maps are isometries from L^∞ to $B(L^2, L^2)$, the bounded operators on L^2.*

Proof. We consider only the left multiplication case. Suppose that $g \in L^\infty$ and that $x \in L^2$. Then in the L^2 norm $||\cdot||_2$

$$||L_g x||_2^2 = ||gx||_2^2 = \oint |g(z)x(z)|^2 \le ||g||_\infty^2 \cdot ||x||_2^2.$$

Note that the operator norm of L_g is actually $||g||_\infty$ since we could have chosen x to be the characteristic function of a set of strictly positive measure on which $|g(z)|$ almost reaches its essential supremum. □

Now it is not too difficult to see how to factor a large class of L^∞ loops. If for $g \in L^\infty$, the map $\pi_+ L_g : L^+ \to L^+$ is one-to-one and onto, then define v to be $(\pi_+ L_g)^{-1}(1)$. Now gv is in L^-, and its zero$^{\text{th}}$ Laurent coefficient is $(gv)_0 = 1$. The loop g can now be expressed in the form $gv = w$, where $w \in L^-$ and $v \in L^+$, and v and w are uniquely defined by the property that $w_0 = 1$. The set of g for which this can be done is open in L^∞ and contains 1. Moreover, if we restrict attention to this open set, then v and w are smooth functions of g, since inversion of one-to-one and onto operators is smooth in a Hilbert space. We are left with two problems: for which g is the map one-to-one and onto, and when can we invert v to express g as a product of elements of L^+ and L^-?

2. How to factor positive invertible loops. Here we show that if $g \in L^\infty$ is positive and invertible, $g^{-1} \in L^\infty$, then the map $\pi_+ L_g$ is one-to-one and onto, and that the corresponding $v = (\pi_+ L_g)^{-1}(1)$ is invertible in L^+.

Proposition 2.1. *If $g \in L^\infty(S^1, M_m)$ takes strictly positive values for almost all points in S^1, then the map $\pi_+ L_g : L^+ \to L^+$ is one-to-one.*

Proof. If not, there is a $v \in L^+$ so that $\pi_+ gv = 0$. Thus, $gv = w$ for some $w \in L_-$ with $w_0 = 0$. Then the function

$$\bar{v}gv = \bar{v}w = \bar{w}v$$

is in L^1 and has all its nonconstant Laurent terms equal to zero since it is both the product of two elements of L^+ and of two elements of L^-. This means that the function is constant almost everywhere, and that constant is $\bar{v}_0 w_0$, which is zero. Then since $\bar{v}gv$ is zero almost everywhere on S^1, we have

$$0 = \oint \langle \bar{v}gv, 1\rangle = \oint \langle gv, v\rangle.$$

Since g is strictly positive almost everywhere on S^1, this means that $v = 0$ almost everywhere. □

Proposition 2.2. *If $g \in L^\infty(S^1 M_m)$ takes strictly positive values for almost all points in S^1, then the map $\pi_+ L_g : L^+ \to L^+$ has dense image.*

Proof. If $u \in L^+$ is perpendicular to the image of $\pi_+ L_g$, then for all $v \in L^+$,

$$\langle \pi_+ gv, u\rangle = \langle gv, u\rangle = \langle v, gu\rangle = \langle v, \pi_+ gu\rangle = 0.$$

Thus $\pi_+ gu$ is zero, and hence u is zero by the previous proposition. □

Now we are left with the problem of showing that $\pi_+ L_g$ has closed image.

Proposition 2.3. *There is a strictly positive constant C depending only on m, so that for all positive $L^\infty(S^1, M_m(\mathbf{C}))$ loops g with L^∞ pointwise inverse g^{-1}, the operator norm of*

$$(\pi_+ L_g)^{-1} : \text{ image } \pi_+ L_g(L^+) \to L^+$$

is less than or equal to $||g^{-1}||_\infty / C$.

Proof. First suppose that a is a strictly positive matrix with eigenvalues $\lambda_1, \dots, \lambda_m$. Then for $x \in M_m$ we can change to an orthonormal basis of eigenvectors of a by a unitary transformation, and in this basis

$$\langle ax, x\rangle = \sum_{i,j} \lambda_i x_{ij} \bar{x}_{ij} \geq \min\{\lambda_i\} \cdot \langle x, x\rangle.$$

Since $\min\{\lambda_i\} = (\max\{\lambda_i^{-1}\})^{-1}$, and since the operator norm of a positive matrix is its maximum eigenvalue,

$$\langle ax, x\rangle \geq \frac{1}{|a^{-1}|_{\mathrm{op}}} \cdot \langle x, x\rangle.$$

Applying this inequality pointwise to the loop g and $v \in L^+$,

$$\langle gv(z), v(z)\rangle \geq \frac{1}{|g^{-1}(z)|_{\mathrm{op}}} \cdot \langle v(z), v(z)\rangle \geq \frac{1}{||g^{-1}||_\infty} \cdot \langle v(z), v(z)\rangle,$$

and integrating round S^1,

$$\oint_{S^1} \langle gv, v\rangle = \oint_{S^1} \langle \pi_+ gv, v\rangle \geq \frac{1}{||g^{-1}||_\infty} \cdot \oint_{S^1} \langle v, v\rangle.$$

Using the equivalence of the operator and inner product norms on $m \times m$ matrices, there is a constant $C = C(m) > 0$ so that

$$||\pi_+ gv||_2 ||v||_2 \geq \frac{C}{||g^{-1}||_\infty} \cdot ||v||_2^2,$$

and dividing by $||v||_2$ gives the required result. □

Corollary 2.4. *Under the conditions of Proposition* 2.3*, the image of the map* $\pi_+ L_g : L^+ \to L^+$ *is complete and therefore closed.*

Proof. This is immediate from mapping back a Cauchy sequence in the image to L^+ by Proposition 2.3. □

Now we know enough to obtain the formula $gv = w$ according to the last paragraph of Section 1, but we still have to find if v and w are invertible:

Proposition 2.5. *If g is a positive L^∞ loop which is invertible almost everywhere on S^1, and if $gv = w$ for some $v \in L^+$ and $w \in L^-$ with $w_0 = 1$, then $v^{-1} \in L^+$ and $w^{-1} \in L^-$.*

Proof. As in Proposition 2.1 the expression $\bar{v}gv = \bar{w}v = \bar{v}w$ is a constant, $\bar{v}_0$. Now if $\bar{v}_0$ were not invertible, there would be a nonzero vector $a \in \mathbf{C}^m$ so that $\bar{v}_0 a = 0$. Then

$$0 = \oint \langle \bar{v}gva, a\rangle = \oint \langle gva, va\rangle,$$

and since g is strictly positive almost everywhere, this means that $va = 0$ almost everywhere. Thus, $gva = wa = 0$ almost everywhere on S^1, so wa is zero on Δ_-. But then $w(z = \infty)a = w_0 a = a = 0$, a contradiction. We deduce that $\bar{v}_0$ is invertible, and then

$$(\bar{v}_0^{-1}\bar{v})w = 1 \quad \text{and} \quad (\bar{v}_0^{-1}\bar{w})v = 1. \qquad \square$$

Now the results can be combined in the following theorem:

Theorem 2.6. *If g is a positive L^∞ loop so that g^{-1} is also in $L^\infty(S^1, M_m)$, then there are functions $v \in L^+$ and $w \in L^-$ with $w_0 = 1$ so that $gv = w$. Further, $v^{-1} \in L^+$ and $w^{-1} \in L^-$, and v and w are smooth functions from the set of loops g with these properties (given the L^∞ norm) to L^2. Alternatively, we can write $g = y\bar{y}$, where $y \in L^-$ also depends smoothly on g.*

Proof. The existence and smoothness of v and w so that $gv = w$ is shown in Propositions 2.1 to 2.5 and the discussion following Proposition 1.4. Now from the proof of Proposition 2.5, we find $\bar{w}v = \bar{v}_0$, an invertible positive matrix, thus

$$g = wv^{-1} = w\bar{v}_0^{-1}\bar{w} = (w\bar{v}_0^{-1/2})(\overline{w\bar{v}_0^{-1/2}}),$$

so define $y = w\bar{v}_0^{-1/2}$. Taking the square root is a smooth operation by 8.3. $\square$

3. The problem with noninvertible loops. In the previous section we assumed that the loop that we wanted to factor had an L^∞ inverse. It might be hoped that the same procedure could work for more general loops, but we shall now show that the map $\pi_+ L_g$ cannot be onto even under relatively mild noninvertibility conditions. This result is adapted from a proof by Carathéodory [**4**]. First we prove a lemma:

Lemma 3.1. *If $g \in L^\infty(S^1, M_m)$ is invertible almost everywhere, but $g^{-1} \notin L^\infty$, then given any $\varepsilon > 0$, there is a unit vector $a \in \mathbf{C}^m$, and a set N of strictly positive measure on the circle, so that $|g(z)a| < \varepsilon$ for all $z \in N$.*

Proof. If not, then for all unit vectors a,

$$\mu\{z : |g(z)a| < \varepsilon\} = 0,$$

where μ is angular measure on the circle. Since there is a countable dense subset of the set of unit vectors, this proves that

$$\mu\{z : |g(z)a| \geq \varepsilon |a| \quad \forall\, a \in \mathbf{C}^m\} = 2\pi.$$

Thus $|g^{-1}(z)|_{\mathrm{op}}| \leq 1/\varepsilon$ almost everywhere, a contradiction. □

Proposition 3.2. *Suppose that g is an L^∞ loop which is strictly positive almost everywhere, but that $g^{-1} \notin L^\infty$. Then the map $\pi_+ L_g : L^+ \to L^+$ is not onto.*

Proof. If it were onto, then the open mapping theorem would imply that it was an isomorphism, since we already know that it is one-to-one by Proposition 2.1. But, given any $\varepsilon > 0$, we will find an $x \in L^+$ so that $||\pi_+ g x||_2 \leq \varepsilon ||x||_2$, and so show that the map cannot be an isomorphism.

First we use the result of the last lemma to show that there is a matrix a of unit norm and a subset N of S^1 of strictly positive measure so that

$$|g(z)a| < \varepsilon/2 \qquad \forall\, z \in N.$$

Then we define a step function $u : S^1 \to \mathbf{R}$ by

$$u(z) = \begin{cases} C/\mu(N) & z \in N \\ -C/(2\pi - \mu(N)) & z \in S^1 - N, \end{cases}$$

where μ is the usual measure giving the circle length 2π, and C is a constant which we shall determine later. Now extend u to a harmonic function on Δ_+ by using the Poisson kernel, and let y be a harmonic conjugate to u on Δ_+ normalized so that $y(0) = 0$. Then the function $f = e^{u+iy}$ is analytic on Δ_+, and since $|f| = e^u$, it follows that f is in L^+. The L^2 norm of f is

$$||f||_2^2 = \mu(N) \cdot e^{2C/\mu(N)} + (2\pi - \mu(N)) \cdot e^{-2C/(2\pi - \mu(N))}.$$

If we define $x = af$, then $||x||_2 = ||f||_2$ and

$$||gx||_2^2 \leq \left(\frac{\varepsilon}{2}\right)^2 \mu(N) \cdot e^{2C/\mu(N)} + ||g||_\infty^2 (2\pi - \mu(N)) \cdot e^{-2C/(2\pi - \mu(N))}.$$

Thus we can choose C sufficiently large so that $||gx||_2 < \varepsilon ||x||_2$. □

This result is rather worrying; however, the fact that we can factor some noninvertible loops is easily seen in the commutative case ($m = 1$). Before continuing it will be convenient to define new projection operators

$$\mathring{\pi}_+ \left(\sum_{n \in \mathbf{Z}} a_n z^n \right) = \frac{a_0}{2} + \sum_{n > 0} a_n z^n$$

and

$$\mathring{\pi}_- \left(\sum_{n \in \mathbf{Z}} a_n z^n \right) = \frac{a_0}{2} + \sum_{n < 0} a_n z^n. \tag{3.3}$$

This definition of the projections has the advantage that the following statements hold:

$$g = \mathring{\pi}_+ g + \mathring{\pi}_- g \quad \text{and} \quad \overline{\mathring{\pi}_+ g} = \mathring{\pi}_- \bar{g}. \tag{3.4}$$

Now look at the case where g is a one-dimensional positive loop, i.e., $g : S^1 \to \mathbf{R}^+$. If we assume that $\log(g)$ is in L^2, then we can write

$$\log(g) = \mathring{\pi}_+ \log(g) + \mathring{\pi}_- \log(g),$$

so g can be factorized as

$$g = w\bar{w} \quad \text{with} \quad w = \exp(\mathring{\pi}_- \log(g)).$$

Note that $|w| = \sqrt{g}$, so the corresponding $v = \exp(-\mathring{\pi}_+ \log(g))$ may not be in L^2.

4. A retraction from continuous positive loops to strictly positive loops. The strategy to factorize noninvertible loops is the following: if g is a positive Hermitian loop with some singular values, we try to factorize g by deforming it into an invertible loop, then factor the resulting invertible loop and take the limit as we tend to the original noninvertible loop. To construct the retraction that we require for this procedure, we use the spectral theory of normal operators [**9**].

Let $b : \mathbf{R} \to \mathbf{R}$ be a smooth positive decreasing function such that $b(x) = 1$ for all $x \leq 1$ and $b(x) = 0$ for all $x \geq 3/2$. Then define a continuous function $f : [0,1] \times (0,\infty) \to (0,\infty)$ by

$$f(t,r) = 1 - b\left(\left(r + \frac{1}{r}\right)(1-t)\right) + r \,.\, b\left(\left(r + \frac{1}{r}\right)(1-t)\right). \tag{4.1}$$

Conventionally we add $f(t,0) = 1$ for all t. Also note that $f(t,r) = 1$ for all $t < 1/4$ and $f(1,r) = r$ for all $r > 0$. Now the induced function $f : [0,1) \to C^0(\mathbf{R}^+, \mathbf{R}^+)$ is continuous and smooth on $(0,1)$, where the uniform norm is used for the space of continuous functions. The differential

$$f_t(t,r) = (1-r)\left(r + \frac{1}{r}\right) \,.\, b'\left(\left(r + \frac{1}{r}\right)(1-t)\right) \tag{4.2}$$

is negative for $r \leq 1$ and positive for $r \geq 1$. Further, on any compact subset of $\mathbf{R}^{*+}$, $f(t,r)$ is eventually constant, i.e., there is a $T < 1$ so that $f(t,r) = r$ if $t \geq T$ and r is in the compact set.

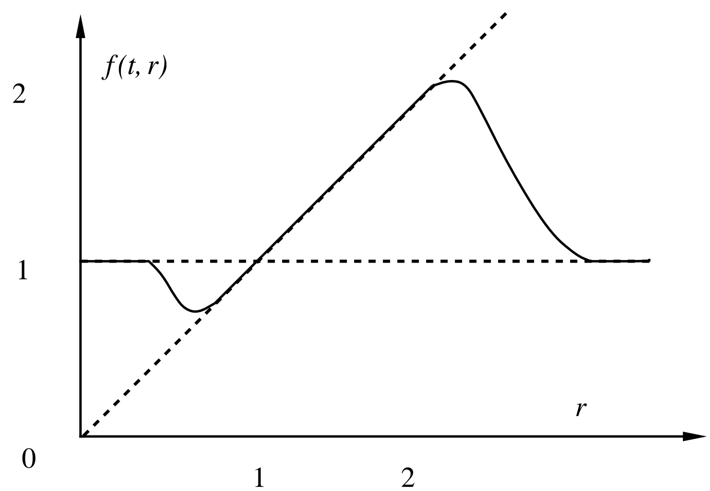

FIGURE 1.

Now use the spectral theory of normal operators to define

$$g(t) = f(t, g) \qquad t \in [0, 1] \tag{4.3}$$

for g a continuous positive loop. Then $g(0) = 1$, $g(1) = g$ where g is invertible on S^1, and the function

$$g : [0, 1) \to C^0(S^1, M_m)$$

is continuous and smooth on $(0, 1)$. Note that $g(t) \to g$ uniformly as $t \to 1$ in any compact subset K of the region in S^1 where g is invertible. In fact, $g(t)$ is eventually constant on such K. In addition $g(t)$ is positive and invertible everywhere on S^1 for $t < 1$.

5. Factoring positive continuous noninvertible loops. Combining the retraction of Section 4 with Proposition 2.6, we have demonstrated the existence of a continuous function

$$y : C^0(S^1, M_m^{\text{pos}}) \times [0, 1) \to L^- \tag{5.1}$$

which is smooth with respect to the interval valued parameter on $C^0(S^1, GL_m^{\rm pos}) \times (0,1)$ constant for $t \in [0,1/4)$ and has the property that

$$g(t) = y(g,t)\overline{y(g,t)} \qquad \forall\,(g,t) \in C^0(S^1, M_m^{\rm pos}) \times [0,1). \tag{5.2}$$

(Here $M_m^{\rm pos}$ denotes the positive matrices). Now it might be thought that all we have to do is to take the limit of y as $t \to 1$; however, we shall have to modify y and impose extra conditions on g before this is possible. Start by differentiating the equation (5.2): (we use $\dot g$ for dg/dt, etc.)

$$\dot g(t) = \dot y \bar y + y \dot{\bar y}, \qquad t \in (0,1). \tag{5.3}$$

This implies that

$$y^{-1}\dot y = \mathring{\pi}_-(y^{-1}\dot g \bar y^{-1}) + a(g,t), \tag{5.4}$$

where $a(g,t)$ is a constant anti-Hermitian loop depending continuously on $g, t \in C^0(S^1, M_m^{\rm pos}) \times [0,1)$. To simplify the situation enough to take the limit as $t \to 1$, we need to remove $a(g,t)$ and estimate the size of the remainder. To get rid of $a(g,t)$, take a unitary matrix value function $u(g,t)$ and observe that

$$(yu)^{-1}\frac{d}{dt}(yu) = u^{-1}\mathring{\pi}_-(y^{-1}\dot g \bar y^{-1})u + u^{-1}au + u^{-1}\dot u. \tag{5.5}$$

Let $u(g,t)$ satisfy the condition $\dot u u^{-1} = -a$, (which can be solved by right invariant integration from the starting point $u(g,0) = 1$), and observe that u is a continuous function

$$u : C^0(S^1, M_m^{\rm pos}) \times [0,1) \to U_m(\mathbf{C})$$

which is smooth with respect to the interval valued parameter on $C^0(S^1, GL_m^{\rm pos}) \times (0,1)$ and constant for $t < 1/4$. Now define a new loop $x = yu$, and then (5.5) becomes

$$x^{-1}\dot x = \mathring{\pi}_-(x^{-1}\dot g \bar x^{-1}) \qquad t \in (0,1). \tag{5.6}$$

Again $x(g,t)$ is continuous on $C^0(S^1, M_m^{\rm pos}) \times [0,1)$, smooth with respect to the interval valued parameter on $C^0(S^1, GL_m^{\rm pos}) \times (0,1)$, and constant (equal to 1) for $t < 1/4$.

Now we can see what happens if we try to take the limit of $x(g,t)$ as $t \to 1$. The simplest way to look at this problem is to substitute

$$x(g,t) = \sqrt{g(t)} \cdot c(g,t) \tag{5.7}$$

where c is a unitary loop. Putting this expression into (5.6) gives a differential equation for c:

$$c^{-1}\dot{c} + c^{-1}\sqrt{g^{-1}}\left(\frac{d}{dt}\sqrt{g}\right)c = \mathring{\pi}_-(c^{-1}\sqrt{g^{-1}}\dot{g}\sqrt{g^{-1}}c). \tag{5.8}$$

By definition of the retraction, $\sqrt{g}$ commutes with $\dot{g}$ (since $\sqrt{g(t)} = \sqrt{f(t,g)}$ and $\dot{g} = f_t(t,g)$), and $(d/dt)\sqrt{g} = (1/2)g/\sqrt{g}$, so (5.8) can be rewritten as

$$c^{-1}\dot{c} + \frac{1}{2}c^{-1}\left(\frac{d}{dt}\log(g)\right)c = \mathring{\pi}_-\left(c^{-1}\left(\frac{d}{dt}\log(g)\right)c\right). \tag{5.9}$$

Now use the formula for $\mathring{\pi}_-$,

$$\mathring{\pi}_-(k)(z) = -\frac{1}{2\pi} \cdot \oint \frac{z}{\zeta - z} \cdot k(\zeta) - \frac{1}{2} \cdot k_0. \tag{5.10}$$

In our case $k = c^{-1}((d/dt)\log(g))c$ is Hermitian, and we only want the anti-Hermitian pat of $\mathring{\pi}_-(k)$ since c is unitary, so

$$c^{-1}\dot{c}(z) = -\frac{1}{2\pi}\oint \operatorname{imag}\left(\frac{z}{\zeta - z}\right)\left(c^{-1}\left(\frac{d}{dt}\log(g)\right)c\right)(\zeta), \tag{5.11}$$

where the integration is over $\zeta \in S^1$ with respect to angular measure.

We should now look at $c^{-1}\dot{c}$ to see if c tends to a limit as $t \to 1$. At first sight there are two problems, firstly (5.11) is only valid for $|z| < 1$, and secondly even if we put $|z| = 1$, the integrand in (5.11) has a singularity. But the choice of retraction was made to make this process simpler. If $z_0 \in S^1$ is in the interior of the complement of the singular points of g, then there is also a radius 4δ open disk about z_0 in the interior. Then $\dot{g}(t)$ vanishes in the radius 2δ disk for all t greater than some $T(\delta, z_0) < 1$. Now the integral (5.11) is continuous as a function of z in the radius 2δ disk for $t > T$, so we can consider z to lie on the circle in this disk. Further, if z is in the radius δ disk about z_0, then $|z/(\zeta - z)| \leq 1/\delta$ for any ζ for which $c^{-1}((d/dt)\log(g))c$ is nonzero if $t > T(\delta, z_0)$. Then if $1 > t > T(\delta, z_0)$ and $|z - z_0| < \delta$, the following inequality holds in operator norm

$$|c^{-1}\dot{c}(t,z)| \leq \frac{1}{2\pi\delta} \,.\, \oint \left| \frac{d}{dt} \log(g(t,\zeta)) \right|. \tag{5.12}$$

We would like to use this equation to show that $c(t)$ converges uniformly in the radius δ disk $B_\delta(z_0)$ as $t \to 1$. Fortunately, this can be done by observing that since g is bounded, $(d/dt)\log(g)$ is negative definite for large enough t, say for $1 > t > T'(\delta, z_0) > T(\delta, z_0)$. Now the following result can be used to make sense of the integral:

Proposition 5.13. *There is a linear function $\alpha : M_m(\mathbf{C}) \to \mathbf{C}$ which is real on the Hermitian matrices, and which has the property that*

$|F|_{\text{op}} \leq \alpha(f)$ *for F a positive linear operator, and*

$|F|_{\text{op}} \leq -\alpha(f)$ *for F a negative linear operator.*

Proof. Given an orthogonal basis $\{e_i\}$ for $\mathbf{C}^m$, define

$$\alpha(F) = 6m^2\Big(\sum_k \langle Fe_k, e_k\rangle + \sum_{k,j} \langle F(e_k + e_j), e_k + e_j\rangle + \sum_{k,j} \langle F(e_k + ie_j), e_k + ie_j\rangle \Big). \qquad \square$$

Continuing from (5.12), if $T'' > T'$ and $z \in B_\delta(z_0)$, then

$$\begin{aligned} |c(T'',z) - c(T',z)| &= \left| \int_{T'}^{T''} \dot{c}(t) \,.\, dt \right| \\ &\leq \int_{T'}^{T''} |c^{-1}\dot{c}(t)| \,.\, dt \\ &\leq \frac{1}{2\pi\delta} \,.\, \int_{T'}^{T''} \oint \left| \frac{d}{dt} \log(g(t)) \right| \,.\, d\phi \,.\, dt \end{aligned} \tag{5.14}$$

(ϕ is the angular variable on S^1). Since the integrand is positive we can reverse the order:

$$|c(T'',z) - c(T',z)| \leq \frac{1}{2\pi\delta} \cdot \oint \int_{T'}^{T''} \left| \frac{d}{dt} \log(g(t)) \right| \cdot dt \cdot d\phi \tag{5.15}$$

As $(d/dt)\log(g(t))$ is negative definite, the inequality can be rewritten using α,

$$|c(T'',z) - c(T',z)| \leq -\frac{1}{2\pi\delta} \cdot \oint \int_{T'}^{T''} \alpha\left(\frac{d}{dt} \log(g(t)) \right) \cdot dt \cdot d\phi, \tag{5.16}$$

and on taking α outside the integral,

(5.17)

$$\begin{aligned} |c(T'',z) - c(T',z)| &\leq -\frac{1}{2\pi\delta} \cdot \oint \alpha\left(\int_{T'}^{T''} \frac{d}{dt} \log(g(t)) \cdot dt \right) \cdot d\phi \\ &\leq -\frac{1}{2\pi\delta} \cdot \oint \alpha(\log(g(T'')) - \log(g(T'))) \cdot d\phi \\ &\leq \frac{|\alpha|}{2\pi\delta} \cdot \oint |\log(g(T'')) - \log(g(T'))| \cdot d\phi. \end{aligned}$$

For $t < 1$, $|\log(g(t))| \leq |\log(g(1))|$. If we suppose that $\log(g(1))$ is in L^1, the dominated convergence theorem says that $\log(g(t)) \to \log(g(1))$ in L^1 as $t \to 1$. The inequality (5.17) then shows that $c(t)$ is uniformly Cauchy on $B_\delta(z_0) \cap S^1$ as $t \to 1$, and this implies that $x(t)$ is uniformly Cauchy on $B_\delta(z_0) \cap S^1$ as $t \to 1$. The results so far can be summarized in

Lemma 5.18. *If $g(1)$ is a continuous positive loop such that $\log(g(1))$ is in L^1, then for sufficiently large t, $c(t)$, and hence $x(t)$, converges uniformly on compact subsets of the complement of the singular points of $g(1)$ in S^1. Since $x(t)$ is also bounded, this implies that $x(t)$ converges to a limit $x(1)$ in $L^2(S^1)$. This in turn implies that $g(1) = x(1)\overline{x(1)}$.*

It should be noted that, with rather more care, the limits could also be taken for unbounded g, provided its log was still in L^1.

Here we aim to complete the proof of the following result:

Theorem 6.1. *If g is a positive continuous loop such that* $\log(g) \in L^1$*, then there is an* $x(g,1) \in L^-$ *so that* $g = x\bar{x}$*. Further, if* $\{g_1, g_2, \dots\}$ *is a sequence of positive continuous loops so that* $g_n \to g$ *in* L^∞ *and* $\log(g_n) \to \log(g)$ *in* L^1*, then* $x(g_n, 1) \to x(g,1)$ *in* L^2*. (We shall call the metric topology resulting from this the* $L^\infty \cap L^1(\log)$ *topology).*

Proof. Note that it is enough to show L^2 convergence on compact subsets of the complement of the singular set of g in S^1. Let $U(\delta)$ be the points of S^1 less than δ away from the singular points of g, and look at the compact set $D = S^1 - U(\delta_0)$ for some fixed $\delta_0 > 0$. To simplify various statements we will write x_n for $x(g_n)$, x for $x(g)$, and also set $g_\infty = g$ and $x_\infty = x$. Since the sequence $\{g_n\}$ is bounded in L^∞, there is a $T' < 1$ such that $(d/dt)\log(g_n(t))$ is negative definite for $t \geq T'$ and $\infty \geq n \geq 1$. Now consider the following results, whose proofs are elementary:

Lemma 6.2. *Given* $\delta > 0$ *there is an* $N(\delta) \geq 1$ *and a* $T(\delta)$ *with* $T' < T(\delta) < 1$ *so that*

$$g_n(t,z) = g_n(1,z)$$

and

$$g(t,z) = g(1,z) \qquad \forall\, z \in S^1 - U(\delta) \quad \forall\, t \in [T(\delta), 1] \quad \forall\, n \geq N(\delta).$$

Lemma 6.3. *If* $f_n \to f$ *in* L^1*, then in the* L^1 *norm over* $U(\delta)$*,*

$$||f_n||_{U(\delta)} \to 0 \quad \textit{and} \quad ||f||_{U(\delta)} \to 0$$

uniformly over n *as* $\delta \to 0$*.*

Proof of Theorem 6.1 (continued). Choose $\varepsilon > 0$. For all $t \in [0,1)$ and $z \in S^1$, there is a constant G depending on the sequence $\{g_n\}$ so that

(6.4)

$$\begin{aligned}|x_n(1,z)-x(1,z)| &\le |x_n(1,z)-x_n(t,z)|+|x_n(t,z)-x(t,z)|\\ &\quad +|x(t,z)-x(1,z)|\\ &\le G\cdot(|c_n(1,z)-c_n(t,z)|+|c(t,z)-c(1,z)|)\\ &\quad +|x_n(t,z)-x(t,z)|\\ &\quad +|\sqrt{g_n(1,z)}-\sqrt{g_n(t,z)}|+|\sqrt{g(1,z)}-\sqrt{g(t,z)}|.\end{aligned}$$

Now if $t \ge T(\delta_0)$, $n \ge N(\delta_0)$, and $z \in D$, this equation can be simplified to

(6.5)

$$\begin{aligned}|x_n(1,z)-x(1,z)| &\le G\cdot(|c_n(1,z)-c_n(t,z)|+|c(t,z)-c(1,z)|)\\ &\quad +|x_n(z,t)-x(z,t)|.\end{aligned}$$

Choose a strictly positive $\delta_1 < \delta_0$, and observe that if $\infty \ge n \ge N(\delta_1)$ and $t \ge T(\delta_1)$ then $(d/dt)\log(g_n(t)) = 0$ in $S^1 - U(\delta_1)$. By (5.11) there is a constant $M > 0$ depending on δ_0 and δ_1 so that

(6.6)

$$|\dot{c}_n(z,t)| \le M\cdot\oint\left|\frac{d}{dt}\log(g_n(t))\right| \qquad \forall\, z\in D,\ t\ge T(\delta_1),\ \infty\ge n\ge N(\delta_1).$$

Then a calculation in the same manner as (5.17) shows that

$$\begin{aligned}\int_t^1 |\dot{c}_n(z,t)|\,dt &\le M|\alpha|\oint|\log(g_n(\zeta,1))-\log(g_n(\zeta,t))|\\ &\qquad \forall\, z\in D,\ t\ge T(\delta_1),\ \infty\ge n\ge N(\delta_1).\end{aligned} \tag{6.7}$$

Since the operator norm of an Hermitian matrix is the maximum absolute value of its eigenvalues, $|\log(g_n(\zeta,t))| \le |\log(g_n(\zeta,1))|$, so for a strictly positive $\delta < \delta_1$,

$$\begin{aligned}\int_t^1 |\dot{c}_n(z,t)| &\le 2M|\alpha|\cdot\int_{U(\delta)}|\log(g_n(\zeta,1))|\\ &\qquad \forall\, z\in D,\ t\ge T(\delta),\ \infty\ge n\ge N(\delta).\end{aligned} \tag{6.8}$$

This can be substituted into (6.5) to give

(6.9)

$$\begin{aligned}|x_n(1,z)-x(1,z)| &\le 2MG|\alpha|\cdot\int_{U(\delta)}(|\log(g_n(\zeta,1))|+|\log(g(\zeta,1))|)\\ &\qquad +|x_n(z,t)-x(z,t)|\end{aligned}$$

for all $z \in D$, $t \geq T(\delta)$, $\infty > n \geq N(\delta)$. By Lemma 6.3 we can choose δ so that $0 < \delta < \delta_1$ and

$$2MG|\alpha| \cdot \int_{U(\delta)} |\log(g_n(\zeta, 1))| < \frac{\varepsilon}{4\sqrt{2\pi}} \tag{6.10}$$

uniformly in all n including ∞. This is now substituted into (6.9), and L^2 norms are taken over D to give

$$\begin{aligned} ||x_n(1,z) - x(1,z)||_D &\leq \frac{\varepsilon}{2} + ||x_n(t,z) - x(t,z)||_D \\ &\leq \frac{\varepsilon}{2} + ||x_n(t,z) - x(t,z)||_{S^1} \end{aligned} \tag{6.11}$$

for $t = \max\{T', T(\delta)\}$ and all $n \geq \max\{N(\delta_1), N(\delta)\}$. But now we have fixed $t < 1$, $g(t)$ is invertible and $g_n(t) \to g(t)$ uniformly on S^1. By the continuity results for invertible loops, $||x_n(t,z) - x(t,z)||_{S^1} \to 0$, so there is an $N' > 0$ so that $||x_n(1,z) - x(1,z)||_D < \varepsilon$ for all $n \geq N'$. This completes the proof of Theorem (6.1). □

7. The problem of evaluating a loop at a point. So far we have only been dealing with L^2 factorizations of loops, and evaluation at a point does not make much sense for an L^2 function. But in applications of the Riemann Hilbert problem it is frequently important to be able to take a value at a single point. To be able to do this, we slightly modify the preceding procedure.

Definition 7.1. For some $\omega \in S^1$, define the vector spaces $E(\omega)$ and $E^\infty(\omega)$ to be

$$E(\omega) = \left\{ (f,b) \in L^2(S^1, M_m) \times M_m \mid \text{ the map } z \mapsto \frac{f(z)-b}{z-\omega} \text{ is in } L^2(S^1, M_m) \right\},$$

$$E^\infty(\omega) = \left\{ (f,b) \in L^\infty(S^1, M_m) \times M_m \mid \text{ the map } z \mapsto \frac{f(z)-b}{z-\omega} \text{ is in } L^2(S^1, M_m) \right\}.$$

$E(\omega)$ and $E^\infty(\omega)$ have norms given by $||f||_{L^2}+||(f-b)/(z-\omega)||_{L^2}+|b|_{\mathrm{op}}$ and $||f||_{L^\infty}+||(f-b)/(z-\omega)||_{L^2}+|b|_{\mathrm{op}}$, respectively.

It is easy to see that b is uniquely defined by $(f,b)\in E(\omega)$, and for convenience we will say that $f\in E(\omega)$ and refer to b as $f(\omega)$. In effect, $E(\omega)$ is a modification of L^2 in which evaluation at ω makes sense and is continuous.

Also define $E^+(\omega)$ to be $E(\omega)\cap L^+$, and $E^-(\omega)$ to be $E(\omega)\cap L^-$.

The following result is an exercise in Cauchy sequences:

Proposition 7.2. *$E(\omega)$ is complete.*

Proposition 7.3. *The operation of left multiplication by $E^\infty(\omega)$ loops on $E(\omega)$ defined by $L_{(g,c)}(f,b)=(gf,cb)$ is continuous. Also the corresponding left "multiply by" map (L) is continuous from $E^\infty(\omega)$ to $B(E(\omega),E(\omega))$, the bounded operators on $E(\omega)$.*

Proof. Suppose that $(g,c)\in E^\infty(\omega)$ and that $(f,b)\in E(\omega)$. Then in the pointwise operator norm,

$$\left|\frac{gf-cb}{z-\omega}\right| \le \left|\frac{gf-gb}{z-\omega}\right| + \left|\frac{gb-cb}{z-\omega}\right| \le ||g||_\infty\left|\frac{f-b}{z-\omega}\right| + |b|\left|\frac{g-c}{z-\omega}\right|$$

and on applying the L^2 norm,

$$\left\|\frac{gf-cb}{z-\omega}\right\|_2 \le ||g||_\infty\left\|\frac{f-b}{z-\omega}\right\|_2 + |b|\left\|\frac{g-c}{z-\omega}\right\|_2. \qquad \square$$

Proposition 7.4. *The projections $\pi_+ : E(\omega)\to E^+(\omega)$ and $\pi_- : E(\omega)\to E^-(\omega)$ defined by*

$$\pi_+(f,x)=\left(\pi_+(f), x+\frac{1}{2\pi}\oint\frac{\zeta}{\zeta-\omega}(f(\zeta)-x)\right),$$
$$\pi_-(f,x)=\left(\pi_-(f), x-\frac{1}{2\pi}\oint\frac{\omega}{\zeta-\omega}(f(\zeta)-x)\right),$$

are continuous.

Proof. We consider only π_+. By Definition 1.2, for $z \in \Delta_+$,

$$\frac{\pi_+ f(z) - \pi_+ f(\omega)}{z - \omega} = \frac{1}{2\pi} \oint \frac{\zeta}{\zeta - z} \frac{f(\zeta) - x}{\zeta - \omega} = \pi_+ \left(\frac{f(\zeta) - x}{\zeta - \omega} \right)(z).$$

The result follows by noting that $(f(z) - x)/(z - \omega)$ is in L^2 and using the previously defined properties of π_+. $\square$

We can now develop a method for factorization with these new spaces. All that has to be done is to follow the previous method: If $g \in E^\infty(\omega)$ and $\pi_+ L_g : E^+(\omega) \to E^+(\omega)$ is one-to-one and onto, then there is a $v \in E^+(\omega)$ so that $gv = \omega \in E^-(\omega)$ with $w_0 = 1$. Further, this v varies smoothly with g in the open set of $E^\infty(\omega)$ in which $\pi_+ L_g$ is one-to-one and onto. But for which g is this condition satisfied? The one-to-one result is easy:

Proposition 7.5. *If $g \in E^\infty(\omega)$ is such that $\pi_+ L_g : L^+ \to L^+$ is one-to-one, then $\pi_+ L_g : E^+(\omega) \to E^+(\omega)$ is also one-to-one.*

Proof. If $\pi_+ gx = 0$ for some $x \in E^+(\omega)$, note that $x \in L^+$ is also. $\square$

The onto result is rather more difficult and requires an extra condition on g.

Proposition 7.6. *If $g \in E^\infty(\omega)$ is such that $\pi_+ L_g : L^+ \to L^+$ is onto, $g(\omega)$ is invertible, and the map $z \mapsto (g(z) - g(\omega))/(z - \omega)$ is in L^∞, then $\pi_+ L_g : E^+(\omega) \to E^+(\omega)$ is also onto.*

Proof. Given $y \in E^+(\omega)$, by the hypothesis there is an $x \in L^+$ so that $\pi_+ gx = y$. Then

$$\pi_+((g - g(\omega))x) + g(\omega) \centerdot \pi_+ x = y.$$

Now observe that the map $z \mapsto (g(z) - g(\omega))/(z - \omega) \centerdot x(z)$ is in L^2, so $(g - g(\omega))x$ is in $E(\omega)$, as is $\pi_+((g - g(\omega))x)$. But then $g(\omega)\pi_+ x$ is in

$E^+(\omega)$, so $\pi_+ x$ is also, and as x is in L^+, we see that $x = \pi_+ x \in E^+(\omega)$. □

Theorem 7.7. *Suppose that g is a positive Hermitian loop so that both g and g^{-1} are in L^∞. Additionally suppose that for some $\omega \in S^1$, the map $z \mapsto (g(z) - g(\omega))/(z - \omega)$ is in L^∞. Then g can be uniquely factored as $gv = w$, where $v \in E^+(\omega)$, $w \in E^-(\omega)$, and $w_0 = 1$. Further, v and w are continuous* (*in fact smooth*) *functions of g if $\{g\}$ is given the $E^\infty(\omega)$ topology. The matrices $v(\omega)$ and $w(\omega)$ are invertible for such g.*

Proof. All except the last statement follows directly from the preceding discussion. For the last part, note that as in Proposition 2.5, the function $\bar{v}gv = \bar{v}w = \bar{w}v$ is an invertible constant. But that constant is also $\bar{v}(\omega)w(\omega)$. □

It is now natural to ask if any sense can be made of evaluation at a point in the singular case. If the invertible continuous loop g obeys the conditions of Theorem 7.7, then we can carry out the retraction as usual and find the function $x(g, t)$. Since $g(\omega)$ is invertible, we note that $g(t) = g$ on a neighborhood of ω for t sufficiently large. Thus, for t sufficiently large (including 1), $x(g, t)$ is in $E^-(\omega)$, so we can evaluate x at ω. The problem arises when g is not invertible, but in this case if $g(\omega)$ is invertible, we can still say that $x(g, t) \in E^-(\omega)$ for sufficiently large $t < 1$.

Proposition 7.8. *Given an $\omega \in S^1$, suppose that g is a continuous positive Hermitian loop so that $g(\omega)$ is invertible. Additionally suppose that the map $z \mapsto (g(z) - g(\omega))/(z - \omega)$ is in L^∞, and that* $\log(g)$ *is in L^1. Then $x(g, t) \in E^-(\omega)$ for sufficiently large $t < 1$, and* $\lim_{t \to 1} x(g, t)(\omega)$ *exists and coincides with $x(g, 1)(\omega)$ in the case where g is invertible.*

Proof. This is mainly covered by the discussion above, with the exception of the limit, which comes from (5.17) and the discussion after it if we set $z_0 = \omega$. □

We shall call the limit $x(g,1)(\omega)$, and it is to be understood that this does not imply that $x(g,1) \in E(\omega)$. There is now the matter of whether the evaluation at ω is continuous in the noninvertible case.

Theorem 7.9. *Fix an $\omega \in S^1$. Consider the set of all positive continuous loops $\{g\}$ with the properties that $g(\omega)$ is invertible, $\log(g) \in L^1$, and such that the map $z \mapsto (g(z) - g(\omega))/(z - \omega)$ is in L^∞. This set is topologized by the $E^\infty(\omega)$ and $L^1(\log)$ norms. Then the function $x(g,1)(\omega)$ (defined by the limit above) is a continuous matrix valued function on the set of such g.*

Proof. The proof follows the same lines as that of Theorem 6.1. Let g and g_n $(n = 1, 2, \dots)$ be in the set above, with g_n being a sequence with limit g in the topology above. Note that there is a neighborhood U of ω in S^1, an $N \geq 1$ and a $T < 1$ so that $g_n(t) = g_n$ on U for all $n \geq N$ and $t > T$. Fix $\varepsilon > 0$, and refer to the proof of Theorem 6.1 with $\omega \in D$. By use of Theorems 6.9 and 6.10, with $z = \omega$,

$$|x_n(1,\omega) - x(1,\omega)| \leq \frac{\varepsilon}{2} + |x_n(t,\omega) - x(t,\omega)|$$

for some fixed $t < 1$ and all n larger than some constant. But since $g_n(t) \to g(t)$ in $E^\infty(\omega)$, for sufficiently large n, $|x_n(t,\omega) - x(t,\omega)| < \varepsilon/2$. □

Appendix

The continuity of the operator calculus. In this section we prove some relatively simple results on the continuity and differentiability of the calculus of normal operators on Hilbert spaces which are assumed in the paper. Denote the set of positive operators by P, and the normal operators by N. The notation $\sigma(A)$ will be used for the spectrum of the operator A.

Proposition 8.1. *For a fixed positive operator A, the map: $C^0(\mathbf{R},\mathbf{R}) \to N$ defined by $f \mapsto f(A)$ is continuous and linear. If A is positive, then the map is continuous on $C^0(\mathbf{R}^+,\mathbf{R})$. Here the mapping spaces are given the compact open topology.*

Proof. Linearity is obvious from the usual definition. Continuity follows from the inequality

$$|f(A)| \leq \sup_{\sigma(A)} ||f||. \qquad \square$$

Proposition 8.2. *The map $P \times C^0(\mathbf{R}, \mathbf{R}) \to N$ defined by $(A, f) \mapsto f(A)$ is continuous.*

Proof. Fix $A \in P$, $f \in C^0(\mathbf{R}, \mathbf{R})$ and $\varepsilon > 0$. Choose a bounded interval J on $\mathbf{R}$ whose interior contains $\sigma(f)$. By the Stone-Weierstrass theorem there is a polynomial q which is within $\varepsilon/5$ of f on the interval J. Additionally we know that there is a neighborhood of A in the operator norm such that every element of the neighborhood has spectrum contained in J. Now, given $B \in P$ with $\sigma(B) \subseteq J$ and $g \in C^0(\mathbf{R}, \mathbf{R})$ we can write the inequality

$$\begin{aligned} |f(A) - g(B)| &\leq |f(A) - f(B)| + |f(B) - g(B)| \\ &\leq |f(A) - q(A)| + |q(A) - q(B)| \\ &\quad + |f(B) - q(B)| + \sup_J ||f - g|| \\ &\leq \frac{2\varepsilon}{5} + |q(A) - q(B)| + \sup_J ||f - g||. \end{aligned}$$

Now all we have to do is to note that $\sup_J ||f - g||$ is less than $\varepsilon/5$ for g sufficiently close to f, and that $q(B)$ tends to $q(A)$ as $B \to A$ in the operator norm, since each is just a polynomial. $\square$

Proposition 8.3. *On a finite dimensional Hilbert space, the operator of taking the positive square root of a strictly positive operator is smooth.*

Proof. First note that the strictly positive operators form an open set in the subspace of Hermitian operators. The map f on this open set defined by $A \mapsto A^2$ is known to be smooth and a homeomorphism. The fact that its derivative $f'(A; B) = AB + BA$ is an isomorphism can be most clearly seen if a basis is chosen so that A is diagonal. Then the inverse function theorem completes the proof. $\square$

Acknowledgments. I would like to express my gratitude to Graeme Segal for suggesting the problem and for much helpful advice. Also I would like to thank the Department of Education (Northern Ireland), and St. Catherine's and Christ Church Colleges Oxford for their support. This work was written up at Swansea, while I was supported by the Science and Engineering Research Council.

REFERENCES

1. B.D.O. Anderson, *Continuity of the spectral factorization operation*, Mat. Apl. Comput. **4** (1985), 139–156.

2. E.J. Beggs, *Inverse scattering and the Chiral equation*, Thesis, Oxford, 1988.

3. ———, *Dispersive solutions of the Chiral equation*, in preparation.

4. C. Carathéodory, *Theory of functions*, Vol. 1, Chelsea, New York, 1983.

5. L.D. Faddeev and L.A. Takhtajan, *Hamiltonian methods in the theory of Solitons*, Springer-Verlag, 1987.

6. S. Novikov, S.V. Manakov, L.P. Pitaevskii and V.E. Zakharov, *Theory of Solitons*; Contemporary Soviet Mathematics, Consultants Bureau, 1984.

7. S.C. Power, *Spectral characterization of the Wold-Zasuhin decomposition and prediction-error operator*, preprint, Univ. of Lancaster.

8. A. Pressley and G.B. Segal, *Loop groups*, O.U.P., 1988.

9. W. Rudin, *Functional analysis*, McGraw-Hill, New York, 1973.

10. G.B. Segal and G. Wilson, *Loop groups and equations of KdV type*, Pub. Math. I.H.E.S. **61** (1985), 5–65.

11. N. Weiner and E.J. Akutowicz, *A factorisation of positive Hermitian matrices*, J. Math. Mech. **8** (1959), 111–120.

12. A. Zygmund, *Trigonometric series*, Vol. 1, C.U.P., 1968.

University College, Swansea, Wales SA2 8PP