OBSERVATION AND PREDICTION FOR THE HEAT EQUATION

by

V. J. Mizel and T. I. Seidman

Report 68-32

September, 1968

OBSERVATION AND PREDICTION FOR THE HEAT EQUATION

V. J. Mizel and T. I. Seidman

§1. Consider an insulated uniform rod with unknown temperature distribution. With proper normalization the temperature $u = u(t,x)$ satisfies

$$(1) \qquad \begin{aligned} &u_t = u_{xx} \qquad (t > 0,\ 0 < x < 1),\\ &u_x(t,0) = u_x(t,1) = 0 \end{aligned}$$

Suppose it possible to observe $f(t) = u(t,0)$ for $0 < t < T$.

We ask: Is it possible, given f, to determine $w(x) = u(T,x)$ for $0 < x < 1$? Assuming the answer to this first question is 'yes', is this a well-posed problem?

We may formulate the problem more precisely as follows. The boundary data f must satisfy certain consistency conditions, i.e., must lie in a certain manifold to. We are then asking whether the operator $A: f \mapsto w$ is well-defined from to to $L^2(0,1)$ and whether it is continuous, topologizing to by the L^2 norm on $(0,T)$:

$$\|f\|^2 = \int_0^T |f(t)|^2 dt.$$

As is well known, we may write the solution of (1) as

$$(2) \qquad u(t,x) = \sum_0^\infty c_n e^{-n^2\pi^2 t} \cos n\pi x$$

with appropriate coefficients $\{c_n : n = 0,1,\ldots\}$ such that

$$\Sigma_n |c_n|^2 e^{-n^2\pi^2 t} < \infty$$

for $t > 0$. With the substitution $s = e^{-\pi^2 t}$, we have

2

f(t)

for $0 < t < T$ or $a = e^{-7r^2 T} < s < 1$. Then

$$Af = w(x) = E_n\ c_n a^{n^2} \cos mrx$$

$$= £_n K>_n(<p)\ a^{n^2} \cos mrx$$

where each I_N (N = 0,1,...) is the linear functional - assuming it is well-defined - such that

$$V^L n^c n^{s^{n^2}} > - <V$$

Letting

$$[A] = sp\{s^{n^2} : n = 0,1,...\},$$

we have, assuming each $\wedge_N$ is continuous on $M = \overline{[A]} £ L^Q(a,l)$,

(3) $$l|A||^2_- < ir^2 L_{na}^{2n^2} ||t_n||^2.$$

What is needed, therefore, is a suitable estimate for IKMII •
The Hahn-Banach Theorem assures us that if

(4) $$s^{N^2}\ jL\ (K£\backslash = iF\{s^{n^2} : n = 0,1,...,n \neq N\}$$

then $\wedge_N$ exists as a continuous linear functional on [M] with

(5) $$IM\sim^{1} = H^{sN2} - [\Lambda_N]\|_1$$

where $||\cdot|_U$ is the L^2 norm on (a,l), so that

(3') $$||A||^2 \le 7r^2 L_n\ a^{2n^2} ||s^{N^2} - [\Lambda_N]||_1^{-2}.$$

The problem is thus reduced to showing that (4) holds, so that

each t_N is well-defined, and obtaining a lower bound for (5).

If we had $a = 0$ then Muntz' Theorem (for which see, e.g.. [2]) assures (4) since $\Sigma n^{-2} < \infty$. For $0 < a < 1$, a result due to Clarkson and Erdös [1] shows that $[\Lambda_N]$ is not dense in $L^2(a,1)$ but does not give (4), much less a lower bound for (5). We shall obtain a sharpened form of the Clarkson-Erdös result, for a certain class of sequences including $\{n^2 : n = 0,1,...\}$, showing that (4) holds if a is sufficiently close to zero (i.e., for large enough T) and that then $\|\ell_N\| = \mathcal{O}(N \log N)$ so that the right hand side of (3) converges and A is continuous.

§2: Let $\Lambda = (\lambda_0, \lambda_1, ...)$ be a sequence of non-negative reals with $0 \le \lambda_0 < \lambda_1 < \ldots$; let $\Lambda_N = \Lambda \setminus \{\lambda_N\}$. Let $[\Lambda] = \mathrm{sp}\{s^{\lambda_n}: \lambda_n \in \Lambda,\ n = 0,1,...\}$. We assume that

(6) $$\Sigma_\Lambda \lambda_n^{-1} < \infty$$

and set

$$T_N = \Sigma_{n>N} \lambda_n^{-1}.$$

It is convenient to introduce

$$r_n^N = \begin{cases} \lambda_n/\lambda_N & n < N \\ \lambda_N/\lambda_n & n > N \end{cases}, \quad r_N = \max_n\{r_n^N : n \ne N\} < 1$$

and

$$\varphi(r) = -\log[(1-r)/(1+(1+\lambda_1^{-1})r)].$$

If we set, for n,N = 0,1,... with n ^ N,

$$\varphi_n^N = \begin{cases} -\log[(1-r_n^N)/(1+r_n^N+x_N^{-1})] & n < N, \\ \text{Ilogtd}\wedge/\text{d+r}\wedge;^1)] & n > N, \end{cases}$$

then

(7) $0 < \text{<}p_n^* \leq \text{<}p(r_n^*)$.

It is known (see, e.g., [2]) that, for any sequence A,

(8, $||s\text{V}^{[\wedge_N]}||^2 = \frac{1}{1+2\lambda_N}$ (^ j ^{2}

where ||*|| denotes the L˜ norm on (0,1) . Thus, by (7),

(8') $||s^{\lambda_N} - [\wedge J|| = (1 + 2\lambda_N)^{-1/2} \exp[-\Sigma_n \varphi_n^N]$

$\geq (1+2X_N)^{-1/2} \exp[-\Sigma_n \varphi(r_n^N)]$.

<u>Lemma 1</u>: Let A, etc., be as above. Then

(9) $||B^{\lambda N}$- [A^IH $\geq$ $(1+2A_N)^{-1/2}$ exp[-N<p(r_N)- V ^] > 0

where a^ = max{<p' (0) , <p(r_N)/r_N).

<u>Proof;</u> We break up the sum on the right of (8') into $\Sigma_{n<M}$ and $\Sigma_{n>N}$. Then J as <P is an increasing function of r,

$$\Sigma_{n<N} \varphi(r_n^N) \leq N\varphi(r_N) .$$

From the form of the function <p we have

cp(r) $\leq$ ^ r ($0 < r \leq r_N$)

so that

$$\Sigma_{n>N} * \wedge \wedge^{E} n>N \ V n \overset{N}{=} V \ N \ V$$

Combining these inequalities with (8') gives (9) . ||

Lemma 2; Let Pe [A^] so that $(s^{\lambda_N} - P) = S \wedge s^{\lambda_n}$ with all

but finitely many b^n zero and with $b_N = 1$. Then

(10) $|b_n| \le ||s^{\lambda} V P|| \ / \ ||s^{\lambda_n} - [A]||$ (n = 0,1,...).

Proof; Observe that (10) is immediate if n = N or if $b_n = 0$. Otherwise (n ^ N, b^ 0) , note that

$$||s^{\lambda_N} - P|| = ||\Sigma_\nu b_\nu s^{\lambda_\nu}|| = |b_n| \, ||s^{\lambda_n} - \Sigma_{\nu \neq n} (b_\nu / b_n) s^{\lambda_\nu}||$$
$$= |b_n| \, ||s^{\lambda_n} - P_o||$$

where P_Q = ^ (^ A ^ $s^{\lambda_\nu} e$ $[A_n]$. Thus,

$$||s^{\lambda_N} - P|| \geq |b_n| \, ||s^{\lambda_n} - [\Lambda_n]||$$

which is just (10). ||

Note that by (6) the sequence A satisfies:

(11) $T_n a^{\lambda_n} < \infty$ for $0 < a < 1$.

We add the assumption that A satisfies the additional condition:

(12) E V N + $(N/A_N)(p(r_N)] \to 0$ as $N \to \infty$.

Lemma 3: Let A, etc., be as above with A satisfying (12). Set

$$\Gamma = \Sigma_0^\infty \alpha^{\lambda_n} \exp[n\varphi(r_n) + a_n \lambda_n \tau_n].$$

Then, for $0 < a < 1$, the series T converges and

(13) $$||B V Pt \leq \& ^{r}Hs^{AN} - P||$$

for $P \in [A_N]$, where $||\cdot||_Q$ is the L^2 norm on $(0,a)$. Observe that $F \to 0$ as $a \to 0$.

Proof: Let $a < a_1 < 1$. By (12), for large enough n we have

$$a_n \tau_n + ({}^n A_n) V(r_n) \leq (\log \alpha_1/\alpha)$$

so that

λ

and, by (11), the series converges. Observe, now, that

$$\|s^{\lambda_N} - P\|_o = \|\Sigma_o^\infty b_n s^{\lambda_n}\| \leq \Sigma_o^\infty |b_n| \|s^{\lambda_n}\|_o .$$

Using (10) and then (9) and the evaluation

$$\|s^\lambda\|_o^2 = \alpha^{2\lambda+1} / (2\lambda+1),$$

we obtain

$$\|sV \ P\|_o \leq \Sigma_o P [\alpha^{2\lambda_n+1}/(2\lambda_n+1)]^{1/2} \|s^{\lambda_N} - P\| / \|s^{\lambda_n} - [\Lambda_n]\|$$

$$\leq \sqrt{\alpha} \|s^{\lambda_N} - P\| 2 \Sigma_o a^{\lambda_n} \exp[n(p(r_n)+ a_n \lambda_n \tau_n]$$

which is just (13). ||

We are now ready, at last, to obtain the desired lower bound for $\|s^{\lambda_N} - [\Lambda_N]\|_1$.

Theorem 1; Let $\Lambda = (\lambda_0, \lambda_1, \dots)$ be an increasing sequence $(0 \le \lambda_0 < \lambda_1 < \dots)$ of reals satisfying (6) and (12) and let $0 < a < 1$ such that $(1 - a^2) = c^2 > 0$ (observe that $c = c(a) \to 1$ as $a \to 0$). Then

$$(14) \qquad \|s^{\lambda_N} - [\Lambda_N]\|_1 \ge c\|s^{\lambda_N} - [\Lambda_N]\|$$

$$\ge c(2\lambda_N + 1)^{-1/2}\exp[-N\varphi(r_N) - a_N\lambda_N\tau_N] > 0.$$

Proof: For any $P \in [\Lambda_N]$ we have

$$\|s^{\lambda_N} - P\|^2 = \|s^{\lambda_N} - P\|_0^2 + \|s^{\lambda_N} - P\|_1^2$$

by the definition of the norms. Using (13) gives

$$\|s^{\lambda_N} - P\|_1^2 \ge (1 - a^2)\|s^{\lambda_N} - P\|^2 \ge c^2\|s^{\lambda_N} - [\Lambda_N]\|.$$

Since this holds for every $P \in [\Lambda_N]$,

$$\|s^{\lambda_N} - [\Lambda_N]\|_1 \ge c\|s^{\lambda_N} - [\Lambda_N]\|$$

which, with (9), gives (14). ||

§3: In order to apply the Theorem above to the prediction problem described earlier it is only necessary to show that the sequence $\Lambda = (n^2 : n = 0, 1, \dots)$ satisfies (6) and (12). Certainly (6) holds and, in fact,

$$(15) \qquad \tau_N = \sum_{n+2}^{\infty} n^{-2} = O(N^{-1}).$$

We have

$$\varphi(r) = |\log(1-r)/(1+2r)|, \quad r_n = n/(n+1)$$

so, for large N, it follows that $U-r_N) = ©(iT^1)$,

(16) $<P(r_N) = \&(\log(l-r_N)) = ©(\log N)$

and

(17) $a_N = <P(r_N)/r_N = ©(\log N)$.

Thus, combining (15), (16) and (17),

$$a\wedge + (N/A_N)<p(r_N) = ©(N''''^1 \log N + (N/N^2) \log N)$$
$$= ©(N''''^1 \log N)$$

which goes to 0 as N —» oo , satisfying (12) .

It follows, therefore, that Theorem 1 may be applied so that, if *a* is small enough that

(18) $ar^2 < 1$,

we have $c^2 = 1 - alT^2 > 0$ and, by (5) and (14),

(19) $$11*\wedge 11 \_< c''^X(2N^2 + l)^{1/2}\exp[N<p(r_N)+ y \wedge]$$
$$= \exp[©(N \log N)] .$$

Substituting this in (3) we see that the factor $a^{2N^2} = \exp[-\&(N^2)]$ dominates and the series converges. We have thus shown the following.

<u>Theorem 2</u>: For T large enough that (18) is satisfied with $a = e^{-T}$, the mapping

$$A: f = u(*,0) i—*u(T,\bullet) = w$$

is a well-defined bounded (using L^2 norms) linear map for solutions u of (1) with $0 < t < T$. I.e., the 'observation and prediction' problem is well-posed. ||

Remark: The above proof does not, of course, show that A is undefined or unbounded for smaller $T > 0$. It would seem of interest to determine whether this notion of a minimal period of observation' is, indeed, a genuine phenomenon or whether it is, perhaps, imposed merely by the exigencies of this particular method of proof. Since $F \to \infty$ as $a \to 1$, (18) is genuinely a restriction on T and it would also be of interest to estimate the minimal T for which (18) is satisfied. The complications involved in estimating the sum for T make the task of estimating a suitable T unfortunately difficult.

§4: A couple of generalizations of Theorem 2 suggest themselves. For example, it is clear that the similar problem for a non-uniform bar

$$(1') \qquad u_t = a(x)u_{xx}, \quad u_x(t,0) = u_x(t,l) = 0$$

could be treated the same way using the expansion

$$(2') \qquad u(t,x) = \sum_n c_n e^{\lambda_n t} v_n(x)$$

where the $\{\lambda_n\}$ and $\{v_n\}$ are the appropriate eigenvalues and eigenfunctions. The asymptotic behavior of the $\{\lambda_n\}$ is known to be similar enough to that of $\{n^2\}$ to ensure the applicability of Theorem 1. The only new problem introduced would be the estimation of an asymptotic lower bound for $\{v_n(0)\}$.

Another example is the treatment of the observation and prediction problem for a solid body, rather than a rod. For a cylindrical body in R^k we have the following result•

Theorem 3: Let & be a 'suitable' region in R^{k-1} and let &* = (0,1)X *. Then, for T large enough (as in Theorem 2), the mapping from $L^2((0,T) \times \&)$ to $L^2(**)$. defined by

A! f = f(t,y) = u(t,0,y)i——>w = w(x,y) = u(T,x,y)

is a well defined,bounded (using L^2 norms) linear map for solutions u of

$$\text{(1'')} \qquad u_t = Au(=u_{xx} + A_y u) \qquad (0 < t < T,\ (x,y) \in \mathcal{D}_*),$$

$$\partial u/\partial \nu = 0 \qquad (0 < t < T,\ (x,y) \in \partial\mathcal{D}_*).$$

Proof: We use the expansion

$$\text{(2'')} \qquad u(t,x,y) = E_{m,n}\, c_{m,n}\, exp[-(n^2 + M_m)\wedge^2 t] v_m(y) \cos nirx$$

where the $[v_m\}$ are normalized solutions of

$$\Delta_y v_m + \mu_m \pi^2 v_m = 0 \qquad (y \in \mathcal{D}),$$

$$\partial v_m/\partial \nu = 0 \qquad (y \in \partial\mathcal{D}).$$

Since the $\{v_{HI}\}$ form an orthonormal sequence we have

$$\langle u(t,0,\cdot), v_m\rangle_{\mathcal{D}} = e^{-u\, ir^z t}\ \Sigma_n\, c_{m,n} e^{-2\ 2.}\ \text{»»*}.$$

Hence, letting I_N be as originally, we have

$$c_{m,N} = \ell_N(s^{-\mu_n} \langle \varphi(s,\cdot), v_m\rangle_{\mathcal{D}})$$

where

$$\varphi(s,y) = f(t,y) = u(t,0,y)$$

on making the substitution $s = e^{-\pi^2 t}$ as before. Thus

Af = Em,nVs -jLt N ^ ') ' 7 ^ n2+/i m vjyjcos mx

and

$$\|Af\|^2 \leq \Sigma_n\ \alpha^{2n^2}\|\ell_n\|^2\ \Sigma_m\ \alpha^{2\mu_m}\|s^{-\mu_m}\langle\varphi, v_m\rangle_{\mathcal{D}}\|_1^2$$

$$\leq \Sigma_n\ \alpha^{2n^2}\|\ell_n\|^2\ \Sigma_m\|\langle\varphi, v_m\rangle_{\mathcal{D}}\|_1^2$$

$$= \Sigma_n\ \alpha^{2n^2}\|\ell_n\|^2\|\varphi\|^2_{(\alpha,1)\times\mathcal{D}} \leq \pi^2\ \Sigma_n\ \alpha^{2n^2}\|\ell_n\|^2\|f\|^2 .$$

This gives the identical estimate (3) for $||A||$ as in Theorem 1 and the same proof now goes through. ||

The result of Theorem 3 suggests a conjecture that a similar prediction problem would be well-posed given observation, for a sufficiently long period, of the restriction to a non-empty relatively open subset $Q \subseteq S\mathcal{D}_*$ where $\mathcal{D}_*$ is a more general domain in R^n (Q here corresponds to $\{0\}\times\mathcal{D}$ in Theorem 3). The methods of this paper, however, seem to afford no direct mode of attack on this more general problem; the proof of Theorem 3 makes essential use of the special nature of $\mathcal{D}_*$ and Θ.

References

[1] Clarkson, J. A., and Erdös, P., "Approximation by Polynomials", Duke Math. J., 10, (1943) pp. 5-11.

[2] Rudin, W., "Real and Complex Analysis,"