DEPARTMENT OF STATISTICS

The University of Wisconsin
Madison, Wisconsin

Technical Report No. 295

April, 1972

MORE ON REGRESSION DESIGN, WITH APPLICATIONS

by

Grace Wahba

Typist: Mary Ann Clark

This work was supported in part by the Wisconsin Alumni Research Foundation, and by the National Science Foundation under Grant GA-18908.

More on Regression Design, with Applications

by

Grace Wahba

Abstract

Earlier results on asymptotically optimal sequences of regression designs are extended. The regression design problem is equivalent to a problem of optimal approximation in a reproducing kernel Hilbert space. From this point of view the results have numerous applications. Two applications are discussed: i) Approximate solution of linear operator equations, ii) Design of indirect sensing experiments.

1. Introduction. Let $\{Y(t), t \in T\}$ be a stochastic process of the form

$$Y(t) = \theta f(t) + X(t) \tag{1.1}$$

where θ is an unknown constant, $f(t)$ is a known function on T, T is a closed, bounded interval which we frequently take to be $[0,1]$, and $\{X(t), t \in T\}$ is a zero mean Gaussian stochastic process with known continuous covariance kernel Q, $E\,X(t)X(t') = Q(t,t')$. The regression design problem is to choose an n-point subset (or "design") $T_n = \{t_1 < t_2 < \dots < t_n, t_i \in T\}$ so that the variance $\sigma^2_{T_n}$ of the Gauss-Markov estimate of θ given $\{Y(t), t \in T_n\}$ is as small as possible.

This problem has been considered by Sachs and Ylvisaker, Wahba, and Hajek and Kimeldorf, [7] [19] [20] [21] [22] [25]. It is known that $\sigma^2_{T_n}$ is bounded away from 0 as $\Delta = \max_i |t_{i+1} - t_i|$ tends to 0 if and only if $f \in \mathcal{H}_Q$, where $\mathcal{H}_Q$ is the (unique) reproducing kernel Hilbert space (RKHS) with reproducing kernel (RK) Q, see [15]. It will be assumed that the reader is familiar with the basic properties of RKHS as given in [15], [25], see also [1].

For each fixed t, let Q_t be the function on T given by

$$Q_t(t') = Q(t,t'). \tag{1.2}$$

By the properties of RKHS, $Q_t \in \mathcal{H}_Q$, and

$$\langle Q_t, f \rangle_Q = f(t), \qquad f \in \mathcal{H}_Q,\ t \in T, \tag{1.3}$$

where $\langle\ ,\ \rangle_Q$ is the inner product in $\mathcal{H}_Q$.

Let P_{T_n} be the projection operator in $\mathcal{H}_Q$ onto the subspace spanned by $\{Q_t, t \in T_n\}$. It is well known that $\sigma_{T_n}^{-2} = \| P_{T_n} f \|_Q^2$ and $\sigma_T^{-2} = \| f \|_Q^2$, where σ_T^2 is the variance of the Gauss Markov estimate of θ given $\{Y(t), t \in T\}$. Hence $\sigma_{T_n}^2$ is minimized by minimizing $\| f - P_{T_n} f \|_Q^2$. From this point of view the problem becomes one of choosing an optimal subspace in $\mathcal{H}_Q$ of the form span $\{Q_t, t \in T_n\}$, for the purpose of approximating the given element f. In this context, the problem has been considered by Karlin, [10] [11].

The purpose of this paper is twofold. Firstly, we extend the results of [7] [19] [20] [21] [22] [25] on asymptotically optimal sequences of designs to a larger class of Q than previously considered. Secondly, we describe two useful applications.

In Section 2 we review the known results on asymptotically optimal sequences of designs. We will extend these results to the regression design problem for nearly all stochastic processes equivalent to $\{X(t), t \in T\}$ of the type considered in [25]. The equivalence classes of processes that we study are described in Section 3. The results in Section 3 are a mild generalization of Kailath and Shepp, [9] [23]. Section 4 gives the extension of the regression design results. In Section 5 we show how these ideas may be used to obtain good approximations to the solutions of certain linear operator equations. A numerical method for finding an approximate solution to a two point boundary problem is given as an example, and the convergence properties of the method are shown to be related to the results of Section 4. In Section 6 we describe the experimental design problem for indirect sensing experiments, and show how the present results apply.

Throughout the paper, we let $\mathcal{H}_Q$ be the RKHS with RK Q, and inner product $\langle \cdot , \cdot \rangle_Q$. We shall always assume that $\int_T \int_T Q^2(t,t')dt\,dt' < \infty$, and will also denote by Q the Hilbert-Schmidt operator [2] on $\mathcal{L}_2[T]$ with Hilbert-Schmidt kernel $Q(t,t')$. It will be helpful to know that $\mathcal{H}_Q = Q^{\frac{1}{2}}(\mathcal{L}_2[T])$, where $Q^{\frac{1}{2}}$ is any square root of Q, and $\| f \|_Q^2 = \inf_{p \in \mathcal{L}_2(T), Q^{\frac{1}{2}}p = f} \| p \|^2_{\mathcal{L}_2[T]}$, where $\| \cdot \|^2_{\mathcal{L}_2[T]}$ is the norm in $\mathcal{L}_2[T]$. (See[16], Chapter 3 or [26] for details).

2. Summary of Previous Results.

The character of the problem of choosing T_n to minimize $\| f-P_{T_n} f \|_Q^2$ depends strongly on the nature of f and whether or not the stochastic process $\{X(t), t \in T\}$ possess quadratic mean derivatives.

Suppose

$$f = \sum_{\nu=1}^{k} c_\nu Q_{s_\nu} \tag{2.1}$$

for some $\{c_\nu\}$, $\{s_\nu\}$, with $s_\nu \in T$. Then if $k \leq n$ and $s_\nu \in T_n$, $\nu = 1,2,\ldots,k$, then $P_{T_n} f = f$ and $\| f - P_{T_n} f \|_Q^2 = 0$. The process $\{X(t), t \in T\}$ possess a ν^{th} quadratic mean derivative at t if the function $Q_t^{(\nu)}(\cdot)$ defined by

$$Q_t^{(\nu)}(\cdot) = (\partial^\nu / \partial s^\nu) Q(s,\cdot)\Big|_{s=t} \tag{2.2}$$

exists and is an element of $\mathcal{H}_Q$. Suppose $\{X(t), t \in T\}$ possesses one quadratic mean derivative for all $t \in T$, and

$$f = Q_{s_*}^{(1)}. \tag{2.3}$$

Then if s_* and $s_*+\Delta$ are in T_n, it can be shown that

$$\lim_{\Delta \to 0} \| f - P_{T_n} f \|_Q^2 = 0 \qquad (2.4)$$

but the limit is not achieved for the elements of T_n distinct. In [7] [10] [11] [19] [20] [21] [22] [25], f's of the form

$$f(s) = \int_T Q(s,t)\, \rho(t)\, dt, \qquad \rho \geq 0 \qquad (2.5)$$

have been considered. This eliminates cases where the solution is either trivial or non existant.

If $\{X(t),\ t \in T\}$ possess exactly $m-1 > 0$ derivatives continuous in quadratic mean, further difficulties arise in studying $\| f - P_{T_n} f \|_Q^2$. The problem may be simplified by letting P_{m,T_n} be the projection operator in $\mathcal{H}_Q$ onto the subspace of $\mathcal{H}_Q$ of dimension $\leq mn$ spanned by

$$\{Q_t^{(\nu)}(\cdot),\ t \in T_n,\ \nu = 0,1,\ldots,m-1\} \ . \qquad (2.6)$$

It can be shown that

$$\inf_{T_{nm}} \| f - P_{T_{nm}} f \|_Q^2 \leq \inf_{T_n} \| f - P_{m,T_n} f \|_Q^2 \leq \inf_{T_n} \| f - P_{T_n} f \|_Q^2 \qquad (2.7)$$

where the infimum is taken over all subsets T_n or T_{mn} of T of the indicated size. The right hand inequality in (2.7) follows because the subspace spanned by $\{Q_t(\cdot),\ t \in T_n\}$ is contained in the subspace spanned by $\{Q_t^{(\nu)}(\cdot),\ t \in T_n,\ \nu = 0,1,2,\ldots,m-1\}$. The left hand inequality follows because

$$\lim_{\Delta\to 0} \operatorname{span}\{Q_{t+\nu\Delta}, \ \nu = 0,1,\ldots,m-1\} \tag{2.8}$$

$$= \operatorname{span}\{Q_t^{(\nu)}, \ \nu = 0,1,\ldots,m-1\} .$$

Further information about the role of derivatives may be found in Karlin [11], especially Theorem 3(i) and Theorem 4, and Sacks and Ylivisaker [21]. In particular, ([11], equation (13), [21], Theorem 4) if $m = 2$ and other conditions are satisfied, the right hand inequality in (2.7) becomes an equality.

Following [21] a sequence T_n^*, $n = 1,2,\ldots$ of designs is said to be asymptotically optimum with derivatives if

$$\lim_{n\to\infty} \frac{\| f - P_{m,T_n^*} f \|_Q^2}{\inf_{T_n} \| f - P_{m,T_n} f\|_Q^2} = 1 \tag{2.9}$$

For h a continuous positive density on $T = [0,1]$, let $T_n = \{t_{0n}, t_{1n}, \ldots, t_{nn}\}$ be defined by

$$\int_0^{t_{in}} h(x)dx = \frac{i}{n}, \qquad i = 0,1,2,\ldots n. \tag{2.10}$$

(For ease of notation we are now letting T_n contain $n+1$ points). Let f be of the form

$$f(t) = \int_0^1 Q(t,s)\rho(s)\,ds \qquad \rho \geq 0, \tag{2.11}$$

then the behavior of $\| f - P_{m,T_n} f\|_Q^2$ for large n, as a function of h is known for various classes of Q and under various regularity conditions on ρ. Using this information a density h^* generating an asymptotically optimum sequence of designs can be found. In this paper we have selected regularity conditions on ρ to ease the proofs, and so shall omit discussion of the weakest regularity conditions under which the following results are known to hold. A sufficient condition is $\rho > 0$, continuous and ρ' bounded.

The known results are

Case 1. $(m = 1)$. Let

i) $\frac{\partial^{\mu+\nu}}{\partial s^{\mu} \partial t^{\nu}} Q(s,t)$ be continuous on the complement of diagonal of the unit square for $\mu+\nu \leq 2$

ii) $\lim_{s \uparrow t} \frac{\partial}{\partial s} Q(s,t) - \lim_{s \downarrow t} \frac{\partial}{\partial s} Q(s,t) = \alpha(t)$ continuous, > 0

iii) $\frac{\partial^2}{\partial t^2} Q(\cdot, t) \in \mathcal{H}_Q$ and $\sup_t \| \frac{\partial^2}{\partial t^2} Q(\cdot,t) \|_Q^2 < \infty$

Then [19]

$$\| f - P_{T_n} f \|_Q^2 = \frac{1}{n^2} \frac{1}{12} \int_0^1 \frac{\rho^2(s) \alpha(s)}{h^2(s)} ds + o\left(\frac{1}{n^2}\right). \quad (2.12)$$

Case 2. Let

$$Q(s,t) = \int_0^s \int_0^t \frac{(s-u)_+^{m-1}}{(m-1)!} \frac{(t-u)_+^{m-1}}{(m-1)!} K(u,v)\, du\, dv \quad (2.13)$$

where K satisfies the conditions i), ii) and iii) placed on Q in case 1, with $\alpha(t) = \alpha$, a positive constant.

Then [21]

$$\| f - P_{m,T_n} f \|_Q^2 = \frac{1}{n^{2m}} \frac{(m!)^2 \alpha}{(2m)!(2m+1)!} \int_0^1 \frac{\rho^2(s)}{h^{2m}(s)} ds + o(\frac{1}{n^{2m}}). \tag{2.14}$$

Case 3. Let

$$Q(s,t) = \sum_{\mu,\nu=1}^{m} \sigma_{\mu\nu} \phi_\mu(s) \phi_\nu(t) + \int_0^1 G(s,u) G(t,u) du \tag{2.15}$$

where G is the Green's function for the differential operator L_m, $L_m f(t) = \sum_{j=0}^{m} a_{m-j}(t) f^{(j)}(t)$ with boundary conditions $f^{(\nu)}(0) = 0$, $\nu = 0,1,\ldots,m-1$, $\{\sigma_{\mu\nu}\}$ are the entries of an $m \times m$ non-negative definite matrix, $\{\phi_\nu\}_{\nu=1}^m$ span the null space of L_m and L_m is such that the $\{\phi_\nu\}_{\nu=1}^m$ are an extended, complete Tchebychev (ECT) system of continuity class C^{2m}.

Then [25], (as a consequence of Lemma 3),

$$\| f-P_{m,T_n} f \|_Q^2 = \frac{1}{n^{2m}} \frac{(m!)^2}{(2m)!(2m+1)!} \int_0^1 \frac{\rho^2(s)\,\alpha(s)}{h^{2m}(s)} ds + o(\frac{1}{n^{2m}}), \tag{2.16}$$

where $\alpha(t) = \dfrac{1}{a_m^2(t)}$.

Here $\dfrac{\partial^{\mu+\nu}}{\partial s^\mu \, \partial t^\nu} Q(s,t)$ exists and is continuous on

the complement of the diagonal of the unit square for $\mu, \nu \leq m$, a property shared by Q of Case 2, and

$$\lim_{s \downarrow t} \frac{\partial^{2m-1}}{\partial s^{2m-1}} Q(s,t) - \lim_{s \uparrow t} \frac{\partial^{2m-1}}{\partial s^{2m-1}} Q(s,t) = (-1)^m \alpha(t). \tag{2.17}$$

Case 4.

$$Q(s,t) = \int_0^1 G(s,u)\, G(t,u)\, du \tag{2.18}$$

where G is as in Case 3 except L satisfies only the weaker conditions $a_0 \neq 0$, $a_{m-j} \in C^j$.

Then [7] , (2.16) holds.

Following [19] , asymptotically optimal sequences of designs may be found from (2.12), (2.14), (2.16) by using a Holder inequality and the fact that $\int_0^1 h(s)ds = 1$ to show that

$$\left[\int_0^1 \frac{\rho^2(t)\,\alpha(t)}{h^{2m}(t)}\, dt\right] \geq \left[\int_0^1 [\rho^2(t)\,\alpha(t)]^{\frac{1}{2m}} dt\right]^{2m+1} \tag{2.19}$$

with equality iff

$$h(s) = \frac{[\rho^2(s)\,\alpha(s)]^{\frac{1}{2m+1}}}{\int_0^1 [\rho^2(u)\alpha(u)]^{\frac{1}{2m+1}} du} \tag{2.20}$$

Thus if

$$\int_0^{t^*_{in}} [\rho^2(u)\,\alpha(u)]^{\frac{1}{2m+1}} du = \frac{i}{n} \int_0^1 [\rho^2(u)\alpha(u)]^{\frac{1}{2m+1}} du \qquad i=1,2,\ldots,n$$

$$t^*_{on} = 0 \tag{2.21}$$

then $T^*_n = \{t^*_{0n}, t^*_{1n}, \ldots t^*_{nn}\}$, $n = 1,2,\ldots,$ is an asymptotically optimal sequence of designs with

$$\| f - P_{m,T_n^*} f \|_Q^2 = \frac{1}{n^{2m}} \frac{(m!)^2}{(2m)!\,(2m+1)!} \left[\int_0^1 [\rho^2(\theta)\,\alpha(\theta)]^{\frac{1}{2m+1}} d\theta\right]^{2m+1} + o\left(\frac{1}{n^{2m}}\right). \tag{2.22}$$

3. Processes Equivalent to m-fold Integrated Weighted Wiener Processes.

It was the original aim of this work to prove that (2.16) holds for all Q which are the covariances of stochastic processes equivalent to a process with Q of Case 3. Such a result might be presumed to be the maximal non-trivial generalization.

To be more precise, let $\{X(t),\ t \in [0,1]\}$ be a zero mean Gaussian stochastic process, which, under the measure μ_0, has the continuous covariance $Q_0(s,t)$, $s,t \in [0,1]$, and under the measure μ_1 has continuous covariance $Q_1(s,t)$, $s,t \in [0,1]$. Now let Q_0 and Q_1 also denote the Hilbert Schmidt operators on $\mathcal{L}_2[0,1]$, with Hilbert-Schmidt kernels $Q_0(s,t)$, and $Q_1(s,t)$ respectively,

$$(Q_i p)(t) = \int_0^1 Q_i(t,s)p(s)ds \quad , \qquad p \in \mathcal{L}_2[0,1),\ i=0,1. \tag{3.1}$$

A version of the Hajek-Feldman Theorem stated in Root [18] says that μ_0 and μ_1 are equivalent measures iff

$$Q_0^{-\frac{1}{2}}\ Q_1\ Q_0^{-\frac{1}{2}} = I - B, \tag{3.2}$$

where $Q_i^{-\frac{1}{2}}$ is the symmetric square root of Q_i^{-1}, $i = 0,1$, B is a Hilbert-Schmidt operator, and $I - B$ is invertible. For simplicity we will say that the kernels $Q_0(s,t)$ and $Q_1(s,t)$ are equivalent if (3.2) holds. Let Q_1 be a kernel equivalent to Q of Case 3 of Section 2. It will be shown that this entails that

$$(-1)^m \alpha(t) = \frac{\partial}{\partial \tau}\left[\frac{\partial^{2m-2}}{\partial s^{m-1}\, \partial t^{m-1}} Q_1(s,t)\Big|_{s=t=\tau} - \int_0^\tau\int_0^\tau \frac{\partial^{2m}}{\partial u^{2m}\, \partial v^{2m}} Q_1(u,v)du\, dv. \right]$$

(3.3)

We have succeeded in proving that (2.16) holds for all Q_1 equivalent to Q of Case 3 with one further condition, reminiscent of iii) of Case 1. The results are stated precisely in Section 4. We next develop some further information about processes equivalent to those of Case 3.

Some of the statements to follow are mild generalizations of statements in Shepp, [23] . We begin with a special case.

Let

$$G_0(s,u) = \frac{(s-u)_+^{m-1}}{(m-1)!} c(u), \tag{3.4}$$

where $(x)_+ = x$, $x \geq 0$, $(x)_+ = 0$ otherwise, and $c(u) \neq 0$. Consider a process $\{X(t), t \epsilon [0,1]\}$ with a realization of the form

$$X(t) = \int_0^t G_0(s,u)c(u)dW(u) = \int_0^t dt \int_0^{t_1} dt_1 \ldots \int_0^{t_{m-1}} c(u)dW(u) \tag{3.5}$$

where W(u) is a Wiener process. Extending [23] , we call X(t) an m-fold integrated weighted Wiener process. The covariance of X is then given by

$$EX(s)X(t) = Q_0(s,t) = \int_0^1 G_0(s,u)G_0(t,u)du \tag{3.6}$$

and the Hilbert Schmidt operator Q_0 may be written

$$Q_0 = G_0 G_0^* \tag{3.7}$$

where G_0^* is the adjoint operator to G_0 with Hilbert Schmidt kernel

$$G_0^*(s,u) = G_0(u,s). \tag{3.8}$$

Now, since

$$Q_0 = Q_0^{\frac{1}{2}} Q_0^{\frac{1}{2}} = G_0 G_0^*, \tag{3.9}$$

$Q_0^{-\frac{1}{2}} Q_1 Q_0^{-\frac{1}{2}}$ is unitarily equivalent to $G_0^{-1} Q_1 G_0^{*-1}$

and

$$Q_0^{-\frac{1}{2}} Q_1 Q_0^{-\frac{1}{2}} = I - B \tag{3.10}$$

with B Hilbert-Schmidt and I-B invertible iff

$$G_0^{-\frac{1}{2}} Q_1 G_0^{*-\frac{1}{2}} = I + A \tag{3.11}$$

for A some Hilbert-Schmidt operator with $I + A$ invertible. Thus, Q_0 and Q_1 are equivalent iff

$$Q_1 = G_0 (I + A) G_0^* \tag{3.12}$$

for A some Hilbert-Schmidt operator with $I + A$ invertible.

We summarize these remarks as

Theorem 3.1. A kernel Q_1 is equivlent to Q_0 of (3.6) iff

$$Q_1(s,t) = \int_0^1 \frac{(s-u)_+^{m-1}}{(m-1)!} \frac{(t-u)_+^{m-1}}{(m-1)!} c^2(u)du + \int_0^1 \int_0^1 \frac{(s-u)_+^{m-1}}{(m-1)!} \frac{(t-v)_+^{m-1}}{(m-1)!} c(u)A(u,v)c(v)dudv \tag{3.13}$$

where

$$\int_0^1 \int_0^1 A^2(s,t)\, ds\, dt < \infty ,$$

and $I + A$ is invertible, A being the operator with Hilbert-Schmidt kernel $A(s,t)$.

Note that

$$\frac{\partial^{2m-2}}{\partial s^{m-1}\,\partial t^{m-1}} Q_1(s,t) = \int_0^{\min(s,t)} c^2(u)du + \int_0^s\int_0^t c(u)A(u,v)c(v)du\,dv \qquad (3.14)$$

and

$$c^2(\tau) = \frac{d}{d\tau}\left\{\frac{\partial^{2m-2}}{\partial s^{m-1}\partial t^{m-1}} Q_1(s,t)\Big|_{s=t=\tau} - \int_0^\tau\int_0^\tau \frac{\partial^{2m}}{\partial u^m\,\partial v^m} Q_1(u,v)\,du\,dv\right\}$$

$$= \alpha(t), \qquad (3.15)$$

also

$$(-1)^m\,\alpha(t) = \lim_{s\downarrow t} \frac{\partial^{2m-1}}{\partial s^{2m-1}} Q_0(s,t) - \lim_{s\uparrow t} \frac{\partial^{2m-1}}{\partial s^{2m-1}} Q_0(s,t), \qquad (3.16)$$

thus verifying (3.3).

Corollary 3.1.

Let

$$Q_1(s,t) = \int_0^1 G_1(s,u)\,G_1(t,u)\,du \qquad (3.17)$$

where

i) $G_1(t,u) = 0,\ t < u$

ii) $\frac{\partial^j}{\partial t^j} G_1(t,u)\Big|_{t\downarrow u} = 0,\quad j = 0,1,\ldots,m-2$

iii) $\frac{\partial^{m-1}}{\partial t^{m-1}} G_1(t,u)\Big|_{t\downarrow u} = c(u)$,

and

iv) $\int_0^1\int_0^1 M^2(t,u)\,dt\,du < \infty$

where

$$M(t,u) = -\frac{1}{c(t)}\,\frac{\partial^m}{\partial t^m} G_1(t,u)\,. \qquad (3.18)$$

Then Q_1 is equivalent to Q_0.

Proof: Letting G_i, $i = 0, 1$ be the (Hilbert-Schmidt) operators defined with kernels G_0 and G_1 of (3.4) and (3.17) and M the (Volterra) operator with Hilbert-Schmidt kernel $M(t,u)$,

we note that

$$(G_0(Mf))(t) = \int_0^t \frac{(t-x)_+^{m-1}}{(m-1)!} dx \int_0^x \frac{\partial^m}{\partial x^m} G_1(x,u) f(u) \; du \tag{3.19}$$

$$= \int_0^t du \; f(u) \int_u^t \frac{(t-x)_+^{m-1}}{(m-1)!} \frac{\partial^m}{\partial x^m} G_1(x,u) dx$$

and since

$$G_1(t,u) = \frac{(t-u)_+^{m-1}}{(m-1)!} c(u) + \int_u^t \frac{(t-x)_+^{m-1}}{(m-1)!} \frac{\partial^m}{\partial x^m} G_1(x,u) dx \tag{3.20}$$

we have

$$G_1 = G_0(I - M). \tag{3.21}$$

Thus

$$Q_1 = G_1 G_1^* = G_0(I+A) G_0^* \tag{3.22}$$

where

$$A = -M - M^* + MM^* \; . \tag{3.23}$$

Since M is Volterra, $(I-M)$ and $(I-M)(I-M^*) = I + A$ are invertible. See, for example Petrovskii, [17] .

It follows from Corollary 3.1 and the properties of Green's functions, that all Q of the form of Case 3, with $\sigma_{ij} = 0$ and $a_0^2(t) = 1/\alpha(t)$ are equivalent to Q_0 of (3.6).

We next extend these results to the class of (Normal) processes equivalent to those with covariance

$$\tilde{Q}_0(s,t) = \sum_{j=0}^{m-1} \phi_j(s)\,\phi_j(t) + \int_0^1 \frac{(s-u)_+^{m-1}}{(m-1)!}\,\frac{(t-u)_+^{m-1}}{(m-1)!}\,c^2(u)\,du \tag{3.24}$$

where $(L_m\phi_j)(s) = a_0(s)\phi_j^{(m)}(s) = 0$, $j = 0,1,\ldots,m-1$, $\phi_j^{(\nu)}(0) = \delta_{\nu-1,j}$, $\nu, j = 0,1,\ldots,m-1$, that is, $\phi_j(s) = s^{j-1}/(j-1)!$. Extending [23], we call such a process an unpinned integrated weighted Wiener process. A process $\{X(t),\ t\in[0,1]\}$ with covariance (3.24) has a representation

$$X(t) = \sum_{\nu=0}^{m-1} X^{(\nu)}(0)\,\phi_\nu(t) + (X(t) - P_{m,0}X(t)),\quad t\in[0,1] \tag{3.25}$$

where $P_{m,0}X(t) = E\{X(t)\,|\,X^{(\nu)}(0),\ \nu = 0,1,\ldots,m-1\}$ and $\{X^{(\nu)}(0)\}_{\nu=0}^{m-1}$ have covariance matrix I_{mxm} where I_{mxm} is the mxm identity matrix. The process $X(t) - P_{m,0}X(t)$ has covariance Q_0 of (3.6). For μ_1 to be equivalent to μ_0, the measure corresponding to $\tilde{Q}_0$ of (3.24), it is necessary and sufficient that $\{X^{(\nu)}(0)\}_{\nu=0}^{m-1}$ exist in q.m. under μ_1 and have a covariance matrix of full rank, and that the process $X(t)-P_{m,0}X(t)$ has a covariance Q under μ_1 which satisfies (3.13). In this case X(t) has a representation of the form

$$X(t) = \sum_{\nu=0}^{m-1} X^{(\nu)}(0)\,\psi_\nu(t) + (X(t) - P_{m,0}X(t)) \tag{3.26}$$

where

$$\psi_\nu(t) = \sum_{j=0}^{m-1} \sigma^{\nu j}\,E\,X(t)\,X^{(\nu)}(0), \tag{3.27}$$

and

$$\{\sigma^{\nu j}\} = \Sigma^{-1}, \quad \Sigma = \{\sigma_{\nu j}\}, \quad \sigma_{\nu j} = E\, X^{(\nu)}(0)\, X^{(j)}(0), \quad \nu, j = 0,1,\ldots,m-1. \tag{3.28}$$

Then,

$$E\, X(s)\, X(t) = \tilde{Q}_1(s,t) = \sum_{j}^{m-1} \psi_j(s)\, \sigma^{jk}\, \psi_k(t) + Q_1(s,t) \tag{3.29}$$

where Q_1 satisfies (3.13), and

$$E\, X(t) X^{(\nu)}(0) = \frac{\partial^\nu}{\partial s^\nu} \tilde{Q}_1(t,s)\Big|_{s=0}, \quad \nu = 0,1,\ldots, m-1. \tag{3.30}$$

The functions $\{\psi_\nu\}_{\nu=0}^{m-1}$ must all be in $\mathscr{H}_{\tilde{Q}_1}$, since by the properties of RKHS [16], any function of the form $h(t) = EZ\, X(t)$ is in $\mathscr{H}_{\tilde{Q}_1}$, if $Z \in$ span $\{X(t),\ t \in [0,1]\}$. If $\tilde{Q}_1$ is equivalent to $\tilde{Q}_0$, then $f \in \mathscr{H}_{\tilde{Q}_1} \Leftrightarrow f \in \mathscr{H}_{\tilde{Q}_0}$. This follows since $\mathscr{H}_{\tilde{Q}_i} = \tilde{Q}_i^{\frac{1}{2}}[\mathscr{L}_2[0,1]$, $i = 1,2$, and $(I+A)^{\frac{1}{2}}[\mathscr{L}_2[0,1]] = \mathscr{L}_2[0,1]$ if $(I+A)$ is invertible. We summarize these remarks in the following

Theorem 3.2. (compare [23], Theorem 8) $\tilde{Q}_1$ is equivalent to $\tilde{Q}_0$ of (3.24) iff

$$\tilde{Q}_1(s,t) = \sum_{i=0}^{m-1} \tilde{\psi}_i(s)\, \tilde{\psi}_i(t) + Q_1(s,t) \tag{3.31}$$

where $\tilde{\psi}_i^{(m-1)}$ abs. cont., $\tilde{\psi}_i^{(m)} \in \mathscr{L}_2[0,1]$, the $m \times m$ matrix Σ with ij th entry $\sigma_{\nu j}$

$$\sigma_{\nu j} = \frac{\partial^{\nu+j}}{\partial s^\nu\, \partial t^j} \left(\sum_{i=0}^{m-1} \tilde{\psi}_i(s)\, \tilde{\psi}_j(t) \right) \Bigg|_{s=t=0} \tag{3.32}$$

is of full rank, and $Q_1(s,t)$ satisfies (3.13).

We have that $\mathcal{H}_{\tilde{Q}_1} = \mathcal{H}_{Q_1} \oplus \text{span}\,\{\tilde{\psi}_i\}_{i=0}^{m-1}$, and if T_n includes the point $t = 0$, then $\| f - P_{m,T_n} f \|^2_{\tilde{Q}_1} = \| P_{Q_1}(f - P_{m,T_n} f) \|^2_{\tilde{Q}_1}$ where P_{Q_1} is the projection in $\mathcal{H}_{\tilde{Q}_1}$ onto the subspace $\mathcal{H}_{Q_1}$. Thus we may and will, without loss of generality consider Q_1 of the form (3.13). This remark holds of course, whatever the rank of the matrix $\sum$.

4. Asymptotically Optimal Experimental Designs For Some Processes Equivalent to Unpinned, Integrated Weighted Wiener Processes.

In this Section we generalize the results of (2.14) and (2.16) to a class of Q's almost including all Q's equivalent to $\tilde{Q}_0$ of (3.25). One additional regularity condition not entailed by equivalence was used in the proof. We have

Theorem 4.1.

Let Q_1 have a representation

$$Q_1(s,t) = \sum_{j=0}^{m-1} \tilde{\psi}_i(s)\tilde{\psi}_i(t)$$

$$+ \int_0^1 \frac{(s-u)_+^{m-1}}{(m-1)!} \frac{(t-u)_+^{m-1}}{(m-1)!} c^2(u)du + \int_0^1\int_0^1 \frac{(s-u)_+^{m-1}}{(m-1)!} \frac{(t-u)_+^{m-1}}{(m-1)!} c(u)A(u,v)c(v)\, du\, dv$$

$$s,t \in [0,1] \qquad (4.1)$$

where

i) $\tilde{\psi}_i^{(m-1)}$ abs. cont., $\tilde{\psi}_i^{(m)} \in \mathcal{L}_2[0,1]$

ii) $c > 0$, c' bounded

iii) $\int_0^1\int_0^1 A^2(u,v)\, du\, dv < \infty$

iv) The function γ_t given by

$$\gamma_t(s) = \int_0^s \frac{\partial}{\partial t} \frac{1}{c(t)} A(t,\eta)\, c(\eta)\, d\eta$$

is well defined and is in the RKHS $\mathcal{H}_{K_1}$ with RK K_1 given by

$$K_1(s,t) = \int_0^{\min(s,t)} c^2(u)\,du + \int_0^s\int_0^t c(u)A(u,v)c(v)\,du\,dv \tag{4.2}$$

and

$$\| \gamma_t \|_{K_1} \le M_1 < \infty$$

where $\| \cdot \|_{K_1}$ is the norm in $\mathcal{H}_{K_1}$.

Suppose

$$f(t) = \int_0^1 Q_1(t,s)\,\rho(s)ds \tag{4.3}$$

with $\rho > 0$, ρ' bounded, and let $T_n = \{0 = t_{0n}, t_{1n}, t_{2n}, \ldots t_{n-1,n}, t_{nn} = 1\}$

with

$$\int_0^{t_{in}} h(u)du = \frac{i}{n} \quad , i = 0,1,\ldots,n \tag{4.4}$$

where

$$\int_0^1 h(u)\,du = 1 \,, \qquad h > 0, \quad h \text{ continuous.}$$

Then

$$\| f - P_{m,T_n} f \|^2_{Q_1} = \frac{1}{n^{2m}} \frac{(m!)^2}{(2m)!(2m+1)!} \int_0^1 \frac{\rho^2(s)\,\alpha(s)}{h^{2m}(s)}\,ds + o\frac{1}{n^{2m}} \tag{4.5}$$

where $\alpha = c^2$.

Remarks:

1. We conjecture that the hypotheses of Theorem 4.1 insure that

$$\| f - P_{T_{nm}} f \|^2_Q \le \| f - P_{m,T_n} f \|^2_Q \,(1 + o(1)\,). \tag{4.6}$$

See (2.7) and [11], Theorem 4. If this conjecture is true, then, since $\| f - P_{m,T_n} f \|_Q^2 \leq \| f - P_{T_n} f \|_Q^2$ always, we have

$$\frac{1}{n^{2m}} \frac{(m!)^2}{(2m)!(2m+1)!} \int_0^1 \frac{\rho^2(s)\alpha(s)}{h^{2m}(s)} \, ds + o\left(\frac{1}{n^{2m}}\right)$$

$$\leq \| f-P_{T_n} f \|_Q^2 \leq \frac{m^{2m}}{n^{2m}} \frac{(m!)^2}{(2m)!(2m+1)!} \int_0^1 \frac{\rho^2(s)\alpha(s)}{h^{2m}(s)} \, ds + o\left(\frac{1}{n^{2m}}\right) \tag{4.7}$$

2. The hypotheses of the Theorem do not include I + A invertible. On the other hand, if I + A is invertible then condition iv) is equivalent to $\gamma_t \in \mathcal{H}_{K_0}$, the RKHS with RK

$$K_0(s,t) = \int_0^{\min(s,t)} c^2(u) \, du, \tag{4.8}$$

where $\mathcal{H}_{K_0} = \{f: f(0) = 0, \ f \text{ abs. cont.}, \ \frac{f'}{c} \in \mathcal{L}_2\}$. Thus if I + A is invertible and (ii) holds then (iv) is equivalent to

$$\int_0^1 \left(\frac{\partial}{\partial t} \frac{1}{c(t)} A(t,\eta)\right)^2 d\eta < \widetilde{M} < \infty . \tag{4.9}$$

Condition (iv) is similar to (iii) of Case 1. This condition is used in the proof of Lemma 4.1 to follow, and we see no way to eliminate it there.

3. The proof below follows closely along the lines of the proof of Theorem 1 of [21], generalized with the aid of [25].

The proof begins with Lemma 4.1.

Lemma 4.1.

Let

$$K_0(s,t) = \int_0^{\min(s,t)} c^2(u)\,du \tag{4.10}$$

$$K_1(s,t) = \int_0^{\min(s,t)} c^2(u)du + \int_0^s\int_0^t c(u)\,A(u,v)c(v)\,du\,dv \tag{4.11}$$

$$f_0(t) = \int_0 K_0(t,u)\,\rho(u)\,du \tag{4.12}$$

$$f_1(t) = \int_0^1 K_1(t,u)\,\rho(u)\,du \tag{4.13}$$

where $\int_0^1\int_0^1 A^2(u,v)\,du\,dv < \infty$, where $c, \rho > 0$, continuous, c', ρ' bounded. Let $\mathcal{H}_{K_i}$, $i = 0,1$, be the RKHS with reproducing kernels K_i, $i = 0,1$, and inner products $\langle\cdot,\cdot\rangle_{K_0}$ and $\langle\cdot,\cdot\rangle_{K_1}$ respectively. Suppose further that, for each t, the function γ_t defined by

$$\gamma_t(s) = \int_0^s \frac{\partial}{\partial t}\frac{1}{c(t)}\,A(t,\eta)\,c(\eta)\,d\eta \tag{4.14}$$

satisfies

$$\gamma_t \in \mathcal{H}_{K_1}, \qquad \|\gamma_t\|_{K_1} \le M_1 < \infty, \quad t \in [0,1]\ . \tag{4.15}$$

Then, there exists an ϵ independent of ρ such that, for sufficiently large n,

$$1 - \epsilon\Delta \leq \frac{\| f_1 - P_{T_n} f_1 \|^2_{K_1}}{\| f_0 - P_{T_n} f_0 \|^2_{K_0}} \leq 1 + \epsilon\Delta. \tag{4.16}$$

where

$$\Delta = \max_i \; | t_{i+1,n} - t_{in} | . \tag{4.17}$$

Here, for i =0, 1, $P_{T_n} f_i$ is the projection of f_i in $\mathcal{H}_{K_i}$ onto the subspace of $\mathcal{H}_{K_i}$ spanned by $\{K_{it},\ t \in T_n\}$, where $K_{it}(t') = K_i(t, t')$.

Proof. For i = 0,1,

$$\left\langle f_i - P_{T_n} f_i,\ f_i - P_{T_n} f_i \right\rangle_{K_i} = \left\langle f_i, f_i - P_{T_n} f_i \right\rangle_{K_i} = \int_0^1 \rho(u)(f_i(u) - P_{T_n} f_i(u))\, du. \tag{4.18}$$

Then

$$\| f_0 - P_{T_n} f_0 \|^2_{K_0} = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \rho(u)(f_0(u) - P_{T_n} f_0(u))\, du$$

$$= \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \rho(u)\, du \int_{t_i}^{t_{i+1}} B_i(u,v)\, \rho(v)\, dv \tag{4.19}$$

where, according to [25] $B_i(u, v)$ is, for $u, v \in [t_i, t_{i+1}]$, the Green's function for the differential operator $L^*_m L_m = g$ with boundary conditions $f(t_i) = f(t_{i+1}) = 0$,

$$(L^*_m L_m f)(t) = \frac{d}{dt} \frac{1}{c^2(t)} \frac{d}{dt} f(t). \tag{4.20}$$

Similarly,

$$\| f_1 - P_{T_n} f_1 \|^2_{K_1} = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \rho(u)\, (f_1(u) - P_{T_n} f_1(u))\, du. \tag{4.21}$$

Since $f_1(u)-P_{T_n}f_1(u) = 0$ for $t = t_1, t_2, \ldots t_n$, and $f_1 - P_{T_n}f_1 \in \mathcal{L}_2[0,1]$, we may write

$$f_1(u)-P_{T_n}f_1(u) = \int_{t_i}^{t_{i+1}} B_i(u,v)\frac{d}{dv}\frac{1}{c^2(v)}\frac{d}{dv}\left(f_1(v)-P_{T_n}f_1(v)\right)dv,\ u\in[t_i, t_{i+1}] \tag{4.22}$$

where B_i is as before. But, since

$$f_1(t)=\int_0^1 K_0(t,u)\rho(u)du + \int_0^1 \rho(u)\,du \int_0^u\int_0^t c(\xi)A(\xi,\eta)c(\eta)d\xi d\eta, \tag{4.23}$$

then

$$\frac{d}{dt}\frac{1}{c^2(t)}\frac{d}{dt}f_1(t) = \rho(t) + \int_0^1 \rho(u)du\int_0^u \frac{\partial}{\partial t}\frac{1}{c(t)}A(t,\eta)c(\eta)\,d\eta \tag{4.24}$$

$$= \rho(t) + \int_0^1 \rho(u)\,\gamma_t(u)\,du. \tag{4.25}$$

By our assumption, $\gamma_t \in \mathcal{H}_{K_1}$, so that (4.25) becomes

$$\frac{d}{dt}\frac{1}{c^2(t)}\frac{d}{dt}f_1(t) = \rho(t) + \int_0^1 \rho(u)\,\gamma_t(u) = \rho(t) + \langle \gamma_t, f_1\rangle_{K_1}. \tag{4.26}$$

Also

$$(P_{T_n}f_1)(t) = (K_1(t,t_1), K_1(t,t_2), \ldots, K_1(t,t_n))\, K_{1,n}^{-1}(f_1(t_1), f_1(t_n) \ldots, f_1(t_n))' \tag{4.27}$$

where $K_{1,n}$ is the n x n matrix with i j th entry $K_1(t_i,t_j)$. Now, for $t \neq t_i$,

$$\frac{d}{dt}\frac{1}{c^2(t)}\frac{d}{dt}K_1(t,t_i) = \frac{d}{dt}\frac{1}{c^2(t)}\frac{d}{dt}\left[\int_0^t\int_0^{t_i} c(\xi)A(\xi,n)c(\eta)d\xi d\eta\right] = \gamma_t(t_i) \tag{4.28}$$

so that, for each fixed $t \notin T_n$,

$$\frac{d}{dt}\frac{1}{c^2(t)}\frac{d}{dt}(P_{T_n} f_1(t)) = (\gamma_t(t_1), \gamma_t(t_2), \ldots, \gamma_t(t_n))K_{1,n}^{-1}(f_1(t_1), f_1(t_2), \ldots, f_1(t_n))'$$

$$= \langle \gamma_t, P_{T_n} f_1 \rangle_{K_1} . \qquad (4.29)$$

Thus, by (4.21),(4.22),(4.26),(4.29),

$$\| f_1 - P_{T_n} f_1 \|^2_{K_1} = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \rho(u)du \int_{t_i}^{t_{i+1}} B_i(u,v)\, \rho(v)\, dv$$

$$+ \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \rho(u)\, du \int_{t_i}^{t_{i+1}} B_i(u,v) \langle \gamma_v, f_1 - P_{T_n} f_1 \rangle_{K_1} dv. \qquad (4.30)$$

Now ρ and B_i are non-negative, so we may write

$$\left| \int_{t_i}^{t_{i+1}} \rho(u)du \int_{t_i}^{t_{i+1}} B_i(u,v) \langle \gamma_v, f_1 - P_{T_n} f_1 \rangle_{K_1} dv \right|$$

$$\leq \int_{t_i}^{t_{i+1}} \rho(u)du \int_{t_i}^{t_{i+1}} B_i(u,v)dv \times M_1 \| f_1 - P_{T_n} f_1 \|_{K_1} , \qquad (4.31)$$

where M_1 is defined in (4.15).

Now, letting

$$\xi_0(t) = \int_0^1 K_0(t,u)\, du \qquad (4.32)$$

it may be shown that [1]

$$\sum_{i=0}^{n} \int_{t_i}^{t_{i+1}} \rho(u)du \int_{t_i}^{t_{i+1}} B_i(u,v)\, dv = \left\langle f_0 - P_{T_n} f_0 ,\ \xi_0 - P_{T_n} \xi_0 \right\rangle_{K_0} . \qquad (4.33)$$

By (2.16),

$$\| \xi_0 - P_{T_n} \xi_0 \|_{K_0} = M_2\left(\frac{1}{n}\left(1 + o\left(\frac{1}{n}\right)\right)\right) \qquad (4.34)$$

for appropriately chosen M_2.
Thus

$$\| f_1 - P_{T_n} f_1 \|^2_{K_1} = \| f_0 - P_{T_n} f_0 \|^2_{K_0} + \theta \frac{M_3}{n} \| f_1 - P_{T_n} f_1 \|_{K_1} \| f_0 - P_{T_n} f_0 \|_{K_0} \qquad (4.35)$$

for some θ with $|\theta| < 1$ and $M_3 = M_1 M_2$, and so

$$\frac{\| f_1 - P_{T_n} f_1 \|^2_{K_1}}{\| f_0 - P_{T_n} f_0 \|^2_{K_0}} = 1 + \theta \frac{M_3}{n}\left(1 + o\left(\frac{1}{n}\right)\right) . \qquad (4.36)$$

Since $\frac{1}{n} \leq \Delta$, the lemma is proved.

Lemma 4.2.

Let

$$Q_i(s,t) = \int_0^s \int_0^t \frac{(s-u)_+^{m-2}}{(m-2)!} \frac{(t-u)_+^{m-2}}{(m-2)!} K_i(u,v)\, dudv \qquad i = 1,2$$

(4.37)

---

[1] Equation (4.33) may be checked by following the argument of Lemma 1 of [25], see equations (3.4), (3.5) and (3.22). Equation (3.4a) there should read $f(t) = EX(t) \int_0^1 X(u)\rho(u)du$.

where K_i, $i = 1,2$ are as in Lemma 4.1 , let

$$f_i(t) = \int_0^1 Q_i(t,u)\ \rho(u)du, \quad i = 1,2. \tag{4.38}$$

Then, there exists an ϵ independent of ρ such that, for sufficiently large n,

$$1 - \epsilon\Delta \leq \frac{\| f_1 - P_{m,T_n} f_1 \|^2_{Q_1}}{\| f_0 - P_{m,T_n} f_0 \|^2_{Q_0}} \leq 1 + \epsilon\ \Delta. \tag{4.39}$$

Here $P_{m,T_n} f_i$ is the projection of f_i in $\mathcal{H}_{Q_i}$, $i = 0,1$ onto the subspace of (2.6) with $Q = Q_i$.

The proof of this Lemma is contained within the proof of Theorem 1 of [21] , p. 2065 Eqns (2.28) to (2.31), where it is shown that (4.16) implies (4.39).

The Theorem now follows by using the proof of Lemma 3 of [25] to show that

$$\| f_0 - P_{m,T_n} f_0 \|^2_{Q_0} = \frac{1}{n^{2m}} \frac{(m!)^2}{(2m)!(2m+1)!} \int_0^1 \frac{\rho^2(s)\alpha(s)}{h^{2m}(s)} ds + o(\frac{1}{n^{2m}}). \tag{4.40}$$

5. Application to the Approximate Solution of Linear Operator Equations.

We will describe a class of linear operator equations for which approximate solutions may be obtained as projections in an RKHS. Some convergence properties of these approximate solutions are then related to the results of Section 4.

Let $\mathcal{H}$ be, for the moment, an arbitrary Hilbert space, with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, let T be a closed, bounded interval, and let $\{\eta_t, t \in T\}$ be a bounded family of elements in $\mathcal{H}$ with the property that every finite subfamily is linearly independent.

Let V be the closure of the span of $\{\eta_t, t \in T\}$ in $\mathcal{H}$, and let $\mathcal{H}_Q$ be the RKHS with reproducing kernel given by

$$Q(t,t') = \langle \eta_t, \eta_{t'} \rangle_{\mathcal{H}}, \qquad t,t' \in T. \tag{5.1}$$

There is an isometric isomorphism between V and $\mathcal{H}_Q$ generated by the correspondance

$$\eta_t \in V \sim Q_t \in \mathcal{H}_Q . \tag{5.2}$$

To see this, note that $\{\eta_t, t \in T\}$ span V, $\{Q_t, t \in T\}$ span $\mathcal{H}_Q$, and

$$\langle \eta_t, \eta_{t'} \rangle_{\mathcal{H}} = Q(t,t') = \langle Q_t, Q_{t'} \rangle_Q . \tag{5.3}$$

Furthermore, $f \in \mathcal{H}_Q$ corresponds to x in V iff

$$\langle Q_t, f \rangle_Q = f(t) = \langle \eta_t, x \rangle_{\mathcal{H}}, \quad t \in T . \tag{5.4}$$

Now let N be the linear operator which maps $\mathcal{H}$ onto $\mathcal{H}_Q$ according to the formula

$$Nx = f \tag{5.5}$$

where

$$(Nx)(t) = f(t) = \langle \eta_t, x \rangle_{\mathcal{H}} . \qquad t \in T \tag{5.6}$$

The null space of N in $\mathcal{H}$ is $V^{\perp}$, since $x \in V^{\perp} \Rightarrow \langle \eta_t, x \rangle_{\mathcal{H}} = 0,\ t \in T$.

A large class of linear operator equations may be set in the form (5.5) and (5.6) including m-point boundary value problems, Fredholm integral equations of the first and second kind, and mixed integro-differential equations. (See [6] [26], [27]).

If x is any element in $\mathcal{H}$ which satisfies (5.5), then the (unique) element in $\mathcal{H}$ of minimum $\mathcal{H}$ norm which satisfies (5.5) is given by $P_V x$, where P_V is the projection operator in $\mathcal{H}$ onto V. To see this, merely note that $x = P_V x + (I - P_V)x$, with $N(I-P_V)x = 0$. For $f \in \mathcal{H}_Q$, denote by $N^+ f$ the element x in $\mathcal{H}$ of minimal $\mathcal{H}$ norm which satisfies (5.5). The operator N^+ so defined has all the appropriate properties of a generalized inverse. (see, for example, Nashed [14]). Let $\hat{x}_n = N_n^+ f$ denote that element of minimum $\mathcal{H}$ norm which satisfies

$$(N(N_n^+ f))(t) \overset{\text{def}}{=} \langle \eta_t, \hat{x}_n \rangle_{\mathcal{H}} = f(t), \quad t \in T_n \tag{5.7}$$

Then $\hat{x}_n$ is given by

$$\hat{x}_n = (\eta_{t_1}, \eta_{t_2}, \ldots, \eta_{t_n})\, Q_n^{-1}\, (f(t_1), f(t_2), \ldots, f(t_n)) = N_n^+ f. \tag{5.8}$$

In the examples cited in [26] [27] and below, (5.8) gives a computationally feasible algorithm for finding an approximate solution to (5.5).

We have

$$N^+ f \in \mathcal{H} \sim f \in \mathcal{H}_Q, \tag{5.9}$$

$$\hat{x}_n \in \mathcal{H} \sim P_{T_n} f \in \mathcal{H}_Q \tag{5.10}$$

under the isomorphism of (5.2), where P_{T_n} is the projection operator in $\mathcal{H}_Q$ onto the subspace spanned by $\{Q_t, t \in T_n\}$.

Thus

$$\| N^+ f - \hat{x}_n \|^2_{\mathcal{H}} = \| f - P_{T_n} f \|^2_Q \quad . \tag{5.11}$$

Therefore, under suitable conditions on f and Q we understand the convergence properties of the algorithm, see (4.7).

As a simple concrete example consider the problem of finding an approximate numerical solution to the 2 point boundary value problem

$$Lx = f \tag{5.12a}$$

$$x(0) = \theta_0$$

$$x(1) = \theta_1$$

where

$$(Lx)(t) = a_0(t)x''(t) + a_1(t)x'(t) + a_2(t)x(t) \quad , \quad t \in [0, 1], \tag{5.12b}$$

and $a_0 > 0$.

We must be able to choose $\mathcal{H} = \mathcal{H}_R$ an RKHS with RK R with the following properties:

(i) the linear functional which maps $x \in \mathcal{H}_R$ into $(Lx)(t)$, is continuous, for each $t \in [0,1]$

(ii) The null space of L is in $\mathcal{H}_R$

(iii) $f \in \mathcal{H}_Q$, where

$$Q(t,t') = \langle \eta_t, \eta_{t'} \rangle_R = L_{(t)} L_{(t')} R(t,t')$$

and $L_{(t)}$ is the linear operator L applied to R considered as a function of t.

Here

$$\eta_t(t') = L_{(t)} R(t,t') . \tag{5.13}$$

and

$$(Lx)(t) = \langle \eta_t , x \rangle_R . \tag{5.14}$$

Since $L(\mathcal{H}_R) = \mathcal{H}_Q$, condition iii) says $\exists\ x \in \mathcal{H}_R \ni Lx = f$. It is possible, however that there is no solution to the boundary value problem. Let R_t be the representer, in $\mathcal{H}_R$, of the evaluation functional at t, $R_t(t') = R(t,t')$. There may fail to be a solution, if, for example, some linear combination of R_0 and R_1 is in $V = \text{span}\ \{\eta_t, t \in T\}$. In that case the system

$$\begin{aligned} \langle \eta_t , x \rangle_R &= f(t),\ t \in T \\ \langle R_0 , x \rangle_R &= \theta_0 \\ \langle R_1 , x \rangle_R &= \theta_1 \end{aligned} \tag{5.15}$$

is overdetermined. Thus assume further that

iv) R_0 and R_1 are linearly independent over V

Since $V^\perp$ is the null space of L, ii), iii) and iv) guarantee the existence of a unique solution $x \in \mathcal{H}_R$ to the boundary value problem.

For this example, let $\hat{x}_n$ be the element in $\mathcal{H}_R$ of minimal $\mathcal{H}_R$ norm which satisfies

$$
\begin{aligned}
\hat{x}_n(0) &= \theta_0 = \langle R_0, \hat{x}_n \rangle_R \\
\hat{x}_n(1) &= \theta_1 = \langle R_1, \hat{x}_n \rangle_R \qquad (5.16) \\
(L\hat{x}_n)(t) &= f(t) = \langle \eta_t, \hat{x}_n \rangle_R \quad , \quad t \in T_n
\end{aligned}
$$

Then $\hat{x}_n$ is the projection of the true solution onto the subspace $\mathcal{H}_R$ spanned by $\{R_0, R_1, \eta_t, t \in T_n\}$. If R is equivalent to an unpinned, r-fold weighted integrated Weiner process with $r - 1 \geq 2$, and $a_0 > 0$, then this subspace is of dimension $n+2$ (see [25] , Theorem 1). Then, using the fact that $\langle \eta_t, R_s \rangle_R = \eta_t(s)$, gives

$$
\hat{x}_n = (R_0, R_1, \eta_{t_1}, \eta_{t_2}, \ldots, \eta_{t_n})
\left(\begin{array}{cc|ccc}
R(0,0) & R(1,0) & \eta_{t_1}(0) & \eta_{t_2}(0) & \eta_{t_n}(0) \\
R(0,1) & R(1,1) & \eta_{t_1}(1) & \eta_{t_2}(1) & \eta_{t_n}(1) \\
\hline
\eta_{t_1}(0) & \eta_{t_1}(1) & & & \\
\vdots & \vdots & & Q_n & \\
\eta_{t_n}(0) & \eta_{t_n}(1) & & &
\end{array}\right)^{-1}
\left(\begin{array}{c}
\theta_0 \\
\theta_1 \\
\hline
f(t_1) \\
\vdots \\
f(t_n)
\end{array}\right) \qquad (5.17)
$$

where Q_n is the n x n matrix with i,j th entry $Q(t_i, t_j)$.

Since $V_n = \{\eta_t, t \in T_n\} \subset \text{span}\{R_0, R_1, \eta_t, t \in T_n\}$, we have

$$
\| x - \hat{x}_n \|_R^2 \leq \| x - P_{V_n} x \|_R^2 = \| f - P_{T_n} f \|_Q^2 \quad . \qquad (5.18)
$$

If the a_i in (5.12b) satisfy $a_i \in C^m$, $a_0 > 0$, and R is chosen to be equivalent to an unpinned m+2 fold integrated Wiener process, then Q satisfies the hypotheses of Theorem 4.1 Then under suitable conditions on f the convergence properties of this algorithm are understood from (4.7).

6. Application to the Design of Indirect Sensing Experiments

Let

$$Y(t) = \int_S K(t,s)\, X(s)\, ds + \epsilon(t), \qquad t \in T \tag{6.1}$$

where S,T are closed, bounded intervals of the real line, $Y(s)$, $s \in S$, and $\epsilon(t)$, $t \in T$ are independent, 0 mean Gaussian stochastic processes with $E\,X(s)\;X(s') = R(s,s')$, $E\,\epsilon(t)\epsilon(t') = P(t,t')$ and $K(t,s)$ is a known Hilbert Schmidt kernel with $K(\cdot, s) \in L_2[T]$, $s \in S$; $K(t,\cdot) \in L_2[S]$, $t \in T$. Let g be a function defined on S such that $\int_S \int_S g(s)\, R(s,s')\, g(s')\, ds\, ds' < \infty$. It is desired to estimate

$$Z = \int_S g(s)\, X(s)\, ds \tag{6.2}$$

based on observations $Y(t)$, $t \in T_n$, and to choose T_n so that the error variance is minimized. A related situation is discussed in [22] . More generally, one would like to estimate

$$Z^\nu = \int_S g_\nu(s)\; X(s)ds, \qquad \nu = 1,2,\ldots,p.$$

Methods of going from the univariate $(p=1)$ case to general finite p are discussed in [20] , we omit further discussion of the case of general p.

Let

$$Q(t,t') = E\,Y(t)\;Y(t') = \int_S \int_S K(t,s)R(s,s')\;K(t',s')\;dsds' + P(t,t'). \tag{6.3a}$$

and suppose $Q(t,t')$ is continuous. Then, the Hilbert-Schmidt operator Q is given by

$$Q = K\,R\,K^* + P. \tag{6.3b}$$

where K^* is the $\mathcal{L}_2$ adjoint of the operator K with Hilbert Schmidt kernel $K(t, s)$.

We assume as usual, that for any finite n, and any distinct $t_1, t_2, \ldots, t_n$ in T, the n x n matrix Q_n with ij th entry $Q(t_i, t_j)$, is strictly positive definite. Then $\hat{Z}_n = E\{Z \mid Y(t),\ t = t_1, t_2, \ldots t_n\}$ is given by

$$\hat{Z}_n = (f(t_1), f(t_2), \ldots, f(t_n))\, Q_n^{-1} (Y(t_1)\ Y(t_2), \ldots, Y(t_n))'. \tag{6.4}$$

where

$$f(t) = E\,ZY(t) = \int_S \int_S K(t,s) R(s,u) g(u)\, ds\, du. \tag{6.5}$$

Furthermore,

$$E(Z - Z_n)^2 = \int_S \int_S g(s) R(s,s')\, g(s')\ \ ds ds' - (f(t_1), f(t_2), \ldots, f(t_n))\, Q_n^{-1} (f(t_1), f(t_2), \ldots, f(t_n))'. \tag{6.6}$$

The right hand side of (6.6) is minimized by choosing the second term as large as possible. Since $f \in \mathcal{H}_Q$, the second term on the right hand side of (6.6) equals $\| P_{T_n} f \|_Q^2$, and is maximized by minimizing $\| f - P_{T_n} f \|_Q^2$.

Next, let

$$y(t) = \int_S K(t,s) x(s) ds + \xi(t), \qquad t \in T \tag{6.7}$$

$$z^\nu = \int_S g_\nu(s)\, x(s)\, ds, \qquad \nu = 1, 2, \ldots p \tag{6.8}$$

where $x \in \mathcal{H}_R$, $\xi \in \mathcal{H}_P$, and where S, T, K, R and P are as before. It follows that $y \in \mathcal{H}_Q$. Here x is considered to be an unknown function describing some physical phenomena which one wishes to investigate by obtaining estimates of the quantitites $\{z^\nu\}_{\nu=1}^p$. $\xi(t)$ is a disturbance, for each t, and $y(t)$, $t \in T_n$ is measured. This set up is common in meteorologial and geophysical experiments, see, for example [3], [4], [5], and is called an indirect sensing experiment. For an example, consider an experiment to determine the particle size distribution of a volume of aerosol fog [4] [8] . x(s) ds is the number of particles of diameter between x(s) and x(s) + ds. A laser beam of unit intensity of monochromatic light is aimed at the fog. K(t,s) is the intensity scattered at angle t by a unit number of particles of diameter s. y(t) is the sum of the total intensity at angle t scattered by the mixture of particles of varying diameters, plus noise. K(t,s) is determined from physical scattering theory.

By adopting a Bayesian point of view and assigning prior (Gaussian) joint distributions to $\{x(s), s \in S\}$ and $\{\xi(t), t \in T\}$, we may sometimes use the preceeding results to obtain a good design T_n. It may be appropriate to set $Ex(s) = m(s) \neq 0$, but this will not affect the choice of optimal design.

Consider next the solution to the problem: Find $\hat{x} \in \mathcal{H}_R$ to minimize

$$J(x) = \|y - Kx\|_P^2 + \|x\|_R^2 = (P^{-\frac{1}{2}}(y-Kx), P^{-\frac{1}{2}}(y-Kx))_{\mathcal{L}_2[T]}$$

$$+ (R^{-\frac{1}{2}} x, R^{-\frac{1}{2}} x)_{\mathcal{L}_2[S]} \qquad (6.9)$$

where K is the Hilbert-Schmidt operator with Hilbert-Schmidt kernel K(t,s) of (6.1), and all the operator square roots are the symmetric square roots. In (6.9) we are also assuming that the null spaces of P and R in $\mathcal{L}_2[T]$ and $\mathcal{L}_2[s]$ respectively are 0. Arguments similar to those below can be carried out without this assumption by using the appropriate generalized inverse, we omit this discussion.

The remainder of this section owes much to [16]. Using standard gradient techniques, see, for example [13], also [16], we find (formally) that the solution must satisfy

$$-K^* P^{-1}(y - K\hat{x}) + R^{-1}\hat{x} = 0 \tag{6.10}$$

or

$$\hat{x} = (R^{-1} + K^* P^{-1} K)^{-1} K^* P^{-1} y. \tag{6.11}$$

Using the operator identity

$$(R^{-1} + K^* P^{-1} K)^{-1} K^* P^{-1} = RK^*(KRK^* + P)^{-1} = RK^* Q^{-1}$$

gives

$$\hat{x} = RK^*(KRK^* + P)^{-1} y \tag{6.12}$$

We next show that $RK^*(KRK^* + P)^{-1}$ is well defined as a bounded linear operator from $\mathcal{H}_Q$ into $\mathcal{H}_R$. This is seen by noting that $y \in \mathcal{H}_Q \Rightarrow y = Q^{\frac{1}{2}} p = (KRK^* + P)^{\frac{1}{2}} p$ for some $p \in \mathcal{L}_2[T]$. Then, since $R^{\frac{1}{2}} K^*(KRK^* + P)^{-\frac{1}{2}}$ is a bounded linear operator from $\mathcal{L}_2[T]$, to $\mathcal{L}_2[S]$,

$$R^{\frac{1}{2}} K^*(KRK^* + P)^{-1} y = R^{\frac{1}{2}} K^*(KRK^* + P)^{-\frac{1}{2}} p \in \mathcal{L}_2[S]. \tag{6.13}$$

Next, since $\mathcal{H}_R = R^{\frac{1}{2}}(\mathcal{L}_2[S])$, we have that

$$R^{\frac{1}{2}}(R^{\frac{1}{2}} K^*(KRK^* + P)^{-1})\, y \in \mathcal{H}_R \tag{6.14}$$

If $y \in \mathcal{H}_Q$, it can easily be checked that

$$J(\hat{x}) < J(\hat{x} + \delta)$$

for any $\delta \in \mathcal{H}_R$ with $\|\delta\|_R^2 \neq 0$. Thus, (6.12) gives the solution $\hat{x} \in \mathcal{H}_R$ of the minimization problem of (6.9) for any $y \in \mathcal{H}_Q$.

We now seek to find an approximate (computable) solution $\hat{x}_n \in \mathcal{H}_R$ to the minimization problem (6.9). Letting

$$\eta_t(s) = \int_S K(t,u)\, R(u,s)\, ds,$$

we have

$$(Kx)(t) = \langle \eta_t, x \rangle_R .$$

Let $\hat{x}_n \in \mathcal{H}_R$ be the solution to the problem: Find $x \in \mathcal{H}_R$ to minimize

$$\sum_{i,j=1}^{n} (y(t_i) - \langle \eta_{t_i}, x \rangle_R) P^{ij} (y(t_j) - \langle \eta_{t_j}, x \rangle_R) + \|x\|_R^2 \qquad (6.15a)$$

where

$$P^{ij} = \left[P_n^{-1}\right]_{ij}, \quad P_n = \{P(t_i, t_j)\} . \qquad (6.15b)$$

The solution $\hat{x}_n$ is given by

$$\hat{x}_n(s) = (\eta_{t_1}(s), \eta_{t_2}(s), \ldots, \eta_{t_n}(s))\, Q_n^{-1} (y(t_1), y(t_2), \ldots, y(t_n))' . \qquad (6.16)$$

Here Q_n is the $n \times n$ matrix with $i\,j$ th entry $Q(t_i, t_j)$,

$$Q(t,t') = \langle \eta_t, \eta_{t'} \rangle_R + P(t,t') = \int_S \int_S K(t,s) R(s,s') K(t',s')\, ds\, ds' + P(t,t'). \qquad (6.17)$$

To study the convergence properties of $\hat{x}_n$ to $\hat{x}$, let η_s^* be that function on T defined, for each s, by

$$\eta_s^*(t) = \eta_t(s) . \qquad (6.18)$$

For each $s \in S$ we know that $\eta_s^* \in \mathcal{H}_Q$. To see this, note that

$$\eta_s^*(t) = E\,X(s)Y(t) = E\hat{X}(s)Y(t) \qquad (6.19)$$

where

$$\hat{X}(s) = E\{X(s) \mid Y(t), t \in T\} . \tag{6.20}$$

and $\{X(s), s \in S\}$ and $\{Y(t), t \in T\}$ are as in (6.1)

Thus $\hat{X}(s)$ is in the Hilbert space spanned by $\{Y(t), t \in T\}$ and so, for each s the function of t defined by $E\hat{X}(s)Y(t)$ is in $\mathcal{H}_Q$, where $Q(t,t') = E\,Y(t)Y(t')$. For future reference, note that this same reasoning gives

$$\langle \eta_s^*, \eta_{s'}^* \rangle_Q = E\hat{X}(s)\hat{X}(s') = (RK^*(KRK^* + P)^{-1}K^*R)(s,s'), \tag{6.21}$$

and so

$$\| \eta_s^* \|_Q^2 = E\hat{X}^2(s) \leq E X^2(s) = R(s,s).$$

We have, for each fixed s, that (6.16) may be written

$$\hat{x}_n(s) = \langle \eta_s^*, P_{T_n} y \rangle_Q = \langle P_{T_n} \eta_s^*, P_{T_n} y \rangle_Q . \tag{6.22}$$

Now, let $\tilde{x}$ be the function on S defined by

$$\tilde{x}(s) = \langle \eta_s^*, y \rangle_Q , \qquad s \in S. \tag{6.23}$$

If

$$y(t) = \int_T Q(t,t')\rho(t')\,dt' \tag{6.24}$$

for some $\rho \in \mathcal{L}_2[T]$, then

$$\tilde{x}(s) = \langle \eta_s^*, y \rangle_Q = \int_T \eta_s^*(t)\rho(t)\,dt = \int_S R(s,u)\,du \int_T K(t,u)\rho(t)\,dt \tag{6.25}$$

and, since $\rho = Q^{-1} y$, we may write $\tilde{x}$ of (6.25) as

$$\tilde{x} = R K^* Q^{-1} y = R K^*(KRK^* + P)^{-1} y = \hat{x} . \tag{6.26}$$

Since the class of y of the form (6.24) with $\rho \in \mathcal{L}_2[T]$ is dense in $\mathcal{H}_Q$, the formula

$$\hat{x}(s) = \langle \eta_s^*, y \rangle_Q = (RK^*(KRK^* + P)^{-1} y)(s) \tag{6.27}$$

is valid for $y \in \mathcal{H}_Q$, and gives the solution to the minimization problem of (6.9). We now have the convergence properties of the approximate solution given by

$$|\hat{x}(s) - \hat{x}_n(s)| = | \langle \eta_s^*, y \rangle_Q - \langle P_{T_n} \eta_s^*, P_{T_n} y \rangle_Q |$$

$$= | \langle \eta_s^* - P_{T_n} \eta_s^*, y - P_{T_n} y \rangle_Q | \leq \| \eta_s^* - P_{T_n} \eta_s^* \|_Q \ \| y - P_{T_n} y \|_Q$$

$$\leq R^{\frac{1}{2}}(s,s) \| y - P_{T_n} y \|_Q^2 . \tag{6.28}$$

Note, for comparison with (6.16) that, in the model (6.1)

$$\eta_t(s) = E X(s) Y(t) = \int_S K(t,u) R(u,s) ds \tag{6.29}$$

and

$$E(X(s) | Y(t), t \in T_n) = (\eta_{t_1}(s), \eta_{t_2}(s), \ldots, \eta_{t_n}(s)) Q_n^{-1} (Y(t_1), Y(t_2), \ldots, Y(t_n)) . \tag{6.30}$$

Thus, the method defined by (6.16) may be expected to give good practical results in finding approximate solutions $\hat{x}_n$ to the equation

$$y(t) = \int_S K(t,s)x(s)\, ds + \xi(t) \qquad (6.31)$$

when the "data" $\{\int_S K(t,s)x(s)ds,\ \ t \in T_n\}$ is contaminated by a "disturbance" $\{\xi(t),\ t \in T_n\}$. This is, in fact, the case. The method is, for suitable choice of R and P, equivalent to the so-called method of regularization, which has been applied in a number of experimental situations, and studied extensively. See, for example [5] , [24] . In [5] , the method is used to approximate the solution of an integral equation involved in the determination of atmospheric temperature profiles. There, $P(t,t')$ is taken as α if $t = t'$ and 0 otherwise in (6.15) The constant α is called the regularization parameter. In [5], a series of numerical experiments were run in an attempt to study the properties of the method with various T_n and α. A test function x is chosen and $\xi(t)$, $t \in T_n$ is chosen by a Monte Carlo procedure. $y(t)$, $t \in T_n$ is then calculated using (6.31) and $\hat{x}_n(s)$ computed using (6.15). Then

$$\sup_s |\hat{x}_n(s) - x(s)| \qquad (6.32)$$

is determined, where x is the original test function. The variation of (6.32) with changes in T_n and α, were studied experimentally.

REFERENCES

[1] Aronszajn, N. (1950). Theory of reproducing kernels. Trans. Amer. Math. Soc. 68, 337-404.

[2] Akhiezer, N.I., and Glazman, I.M. (1963), Theory of Linear Operators in Hilbert Space, vol. 1. Merlynd Nestell, translator. Ungar.

[3] Backus, George (1970), Inference from inadequate and inaccurate data. Proc. Nat. Acad. Sci, 65, 1, 1-7.

[4] Dave, J.V. (1971), Determination of size distribution of spherical polydispersions using scattered radiation data. Applied Optics, 10, 9, 2035-2044.

[5] Glasko, V.B., and Timofeyev, Yu. M., (1968). Possibilities of the Regularization method in solving the problem of atmospheric thermal sounding. Izv., Atmospheric and Oceanic Physics, Vol. 4, No. 12, 1243-1253, translated by Allen B. Kaufman.

[6] Golomb, Michael, (1971), Spline Approximations to the solutions to two point boundary value problems. University of Wisconsin Mathematics Research Center Technical Summary Report #1066.

[7] Hajek, Jaroslav, and Kimeldorf, George, (1972) Regression designs in autoregressive stochastic processes. Florida State University, Statistics Dept. Report M229, Tallahassee, Florida.

[8] Herman, Benjamin M., Browning, Samuel R., and Reagan, John A., (1971), Determination of Aerosol size distributions from Lidar measurements. J. Atmos. Sci. 28, 763-771.

[9] Kailath, Thomas, (1967), On measures equivalent to Wiener measure, Ann. Math. Statist. 38, 261-263.

[10] Karlin, S. (1969). The fundamental theorem of algebra for monosplines satisfying certain boundary conditions, and applications to optimal quadrature formulas. In Proceedings of the Symposium on Approximations, with Emphasis on Splines ed. I.J. Schoenberg. Academic Press, New York.

[11] Karlin, S., (1972), A Class of best non-linear approximation problems, Bull. A.M.S. 78, 43-49.

[12] Kimmeldorf, George, and Wahba, Grace, (1971), Some results on Tchebycheffian Spline functions. J. Math. Anal. Applic., 33, 82-95.

[13] Nashed, M.Z., (1971), Differentiability and Related Properties of Non Linear Operators: Some Aspects of the Role of Differentials in Non-linear Functional Analysis. In Nonlinear Functional Analysis and Applications, L.B. Rall, ed., Academic Press, 103-310.

[14] Nashed, M.Z., (1971), Generalized inverses, normal solvability, and iteration for singular operator equations. In Nonlinear Functional Analysis and Applications, L.B. Rall, ed., Academic Press, 311-359.

[15] Parzen, E. (1961). An approach to time series analysis. Ann. Math. Statist. 32, 951-989.

[16] Parzen, Emanuel, (1971), Statistical inference on time series by RKHS methods, in Proceedings of the 12 th Biennial Canadian Mathematical Society Seminar, Ronald Pyke, ed., 1-37.

[17] Petrovskii, I.G. (1957), Lectures on the Theory of Integral Equations, Graylock Press.

[18] Root, W.L., (1962), Singular Gaussian measures in detection theory, in Time Series Analysis, Proceedings of a Symposium held at Brown University, M. Rosenblatt, ed., Wiley, 292-314.

[19] Sacks, Jerome and Ylvisaker, Donald (1966). Designs for regression with correlated errors. Ann. Math. Statist. 37, 66-89.

[20] Sacks, Jerome and Ylvisaker, Donald (1968). Designs for regression problems with correlated errors; many parameters. Ann. Math. Statist. 39, 49-69.

[21] Sacks, Jerome and Ylvisaker, Donald (1969). Designs for regression problems with correlated errors, III, Ann. Math. Statist., 41, 2057-2074.

[22] Sacks, Jerome, and Ylvisaker, Donald (1971), Statistical designs and integral approximation, in Proceedings of the 12 th Biennial Canadian Mathematical Society Seminar, Ronald Pyke, ed., 115-136.

[23] Shepp, L.A. (1966). Radon-Nikodym derivatives of Gaussian measures. Ann. Math. Statist. 32, 321-354.

[24] Wahba, Grace (1969), On the approximate solution of Fredholm integral equation of the first kind, University of Wisconsin Mathematics Research Center Technical Summary Report No. 990.