The Annals of Statistics
1989, Vol. 17, No. 2, 556–572

CONVERGENCE RATES FOR REGULARIZED SOLUTIONS OF INTEGRAL EQUATIONS FROM DISCRETE NOISY DATA

BY DOUGLAS W. NYCHKA AND DENNIS D. COX[1]

North Carolina State University and University of Illinois, Urbana

Given data $y_i = (Kg)(u_i) + \varepsilon_i$ where the ε's are random errors, the u's are known, g is an unknown function in a reproducing kernel space with kernel r and K is a known integral operator, it is shown how to calculate convergence rates for the regularized solution of the equation as the evaluation points $\{u_i\}$ become dense in the interval of interest. These rates are shown to depend on the eigenvalue asymptotics of KRK^*, where R is the integral operator with kernel r. The theory is applied to Abel's equation and the estimation of particle size densities in stereology. Rates of convergence of regularized histogram estimates of the particle size density are given.

1. Introduction. Integral equations often provide a crucial link between observations on a system and a function g that characterizes the state of the system. We consider an observational model of the form

$$y_{in} = \int_0^1 k(u_{in}, v)g(v)\,dv + \varepsilon_{in}, \qquad i = 1, 2, \dots, n. \tag{1.1}$$

Here, the kernel function k: $[0,1] \times [0,1] \to \mathbb{R}$ is assumed known, the u's are known points in $[0,1]$ and the ε's are mean zero, random errors. g: $[0,1] \to \mathbb{R}$ is an unknown function that is thought to lie in a Hilbert space of smooth functions $\mathscr{X}$. In particular, $\mathscr{X}$ will be a reproducing kernel Hilbert space (RKHS) with a continuous reproducing kernel r: $[0,1] \times [0,1] \to \mathbb{R}$. This class of $\mathscr{X}$'s encompasses a rich variety of spaces including the Sobolev spaces

$$W_2^k[0,1] = \{ f\colon f^{(j)} \text{ absolutely continuous for } 0 \le j \le k-1 \text{ and } f^{(k)} \in L_2[0,1] \}.$$

Given the observation vector

$$\mathbf{y}_n = (y_{i1}, \dots, y_{in})',$$

the statistical problem is to obtain an estimate of g. If $\mathscr{X}$ is finite dimensional, then this can be treated by standard parametric regression techniques. However, there is frequently no sound basis for assuming a parametric form for g, in which case it is appropriate to apply nonparametric methods, i.e., use an infinite dimensional parameter space $\mathscr{X}$. This makes the estimation problem somewhat more difficult.

Received October 1984; revised May 1988.

[1]Research supported by NSF Grant MCS-82-02560.

AMS 1980 *subject classifications.* Primary 62G05; secondary 62J05, 41A35, 41A25, 47A53, 45L10, 45M05.

Key words and phrases. Method of regularization, approximate solution of integral equations, rates of convergence, nonparametric regression, Abel's equation, particle size distribution.

The type of estimator which is analyzed here is the so-called method of regularization (MOR) estimator, which was first proposed in the integral equation context by Tikhonov (1963) [see also Nashed and Wahba (1974)]. Statistical justifications for such estimators have been given by Li (1982) [see also Speckman (1985)]. The arguments in Kimeldorf and Wahba (1970) can be adapted to show that linear Bayes estimators are MOR estimators. The estimator is obtained by minimization over $h \in \mathscr{X}$ of

$$\lambda\langle h, Wh\rangle_{\mathscr{X}} + \frac{1}{n}\sum_{i=1}^{n}\left(y_{in} - \int_{[0,1]} k(u_{in}, v)h(v)\,dv\right)^2,$$

where $W \in \mathscr{L}(\mathscr{X}) = \{\text{all bounded linear operators } \mathscr{X} \to \mathscr{X}\}$ and $\lambda \in (0, \infty)$. It is assumed that $\langle h, Wh\rangle_{\mathscr{X}}$ is positive semidefinite for $h \in \mathscr{X}$ ($\langle \cdot, \cdot \rangle_{\mathscr{X}}$ denotes the inner product on $\mathscr{X}$), that the null space of W has finite dimension and that $\langle h, Wh\rangle_{\mathscr{X}} = 0$ and $\int_{[0,1]} k(u_{in}, v)h(v)\,dv = 0$, for $1 \le i \le n$, together imply $h \equiv 0$. This latter condition guarantees that the resulting estimate (denoted $\hat{g}_{n\lambda}$) is uniquely defined. MOR estimators are widely used and appear to work well in practice. A typical choice for $\mathscr{X}$ is $W_2^k[0,1]$ with $\langle h, Wh\rangle_{\mathscr{X}} = \int_{[0,1]}(h^{(k)}(v))^2\,dv$ and in this case the null space of W will be the set of polynomial functions with degree less than k. With this specific choice of W, $\langle h, Wh\rangle_{\mathscr{X}}$ is referred to as a roughness penalty while λ is called the smoothing parameter. Note that the smoothing parameter controls the relative weight between the roughness of the estimate and its fit to the data. Some pertinent references are Wahba (1977, 1980), Lukas (1980, 1981, 1988) and O'Sullivan (1986).

The main interest here is to study asymptotic properties of $\hat{g}_{n\lambda}$ as $n \to \infty$ and the u_{ni} become dense in $[0,1]$. In particular, we will show that for appropriate sequences of λ, $\hat{g}_{n\lambda}$ is a consistent estimator and we can obtain upper bounds on the rates of convergence. These results are useful for indicating how the features of the problem (i.e., choice of λ, $\mathscr{X}$, W, the true g and the design points $\{u_{in}\}$) affect the estimation error. Determination of λ is another important issue and these results can also be applied to study the use of generalized cross validation for the adaptive choice of λ [see Cox (1984)].

The next section uses the general results on MOR estimates from Cox (1988) and adapts them to the particular problem of solving integral equations. In Section 3 this theory is applied to the weakly, singular Volterra kernel

$$k(u, v) = \begin{cases} (v - u)^{-1/2}, & v > u, \\ 0, & v \le u. \end{cases}$$

If we let K denote the integral operator

$$(Kg)(u) = \int_u^1 \frac{g(v)\,dv}{\sqrt{v - u}}, \tag{1.2}$$

then for g unknown and h known, $h = Kg$ gives Abel's equation. The statistical problem we address is to estimate g from noisy measurements of h on a discrete set. MOR estimates give one solution to this ill-posed problem.

Abel's equation has a diverse range of applications in the physical sciences [Kosarev (1980) and Bullen (1963)] and in stereological microscopy [Anderssen and Jakeman (1975a, b), Anderssen (1976) and Nychka, Wahba, Goldfarb and Pugh (1984)]. In this latter field one is interested in estimating the probability distribution (or its density) of the radii of spherical bodies embedded in a medium using the cross sectional radii observed on a planar slice of the medium. In the last section of this paper, we solve this stereologic problem using a MOR estimator based on a histogram of the observed cross sectional areas. Hall and Smith (1985) apply kernel density estimates in this context and obtain comparable results. Watson (1971) and Franklin (1981) give a more traditional statistical approach to this problem.

We end this introduction by reporting the specific results for Abel's equation. In this context, we take $\mathscr{X}$ to be the Sobolev space with boundary conditions

$$\mathscr{X} = \{h \in W_2^2[0,1]: h(1) = h^{(1)}(1) = 0\}.$$

One possible MOR estimate of g is obtained by minimizing

$$\lambda \int_0^1 (h^{(2)}(v))^2\, dv + \frac{1}{n}\sum_{i=1}^{n}\left(y_{in} - \int_{u_{in}}^1 \frac{h(v)}{\sqrt{v - u_{in}}}\, dv \right)^2$$

over all $h \in \mathscr{X}$.

(In this case $W \equiv I$ since $\|h^{(2)}\|^2_{L^2[0,1]}$ is a norm for $\mathscr{X}$.)

The following conditions will be assumed throughout this paper.

CONDITION 1. $E\varepsilon_n = 0$ and there are constants $\{S_n\} \subseteq (0,\infty)$ such that

$$E\left(\frac{1}{n}\varepsilon_n'\eta\right)^2 = O\left(\frac{S_n}{n}\right)\eta'\eta$$

uniformly for $\eta \in \mathbb{R}^n$, as $n \to \infty$.

Note that for errors that are uncorrelated $S_n \sim 1/n$.

CONDITION 2. If F_n is the empirical distribution of the design sequence

$$\{u_{jn}: 1 \le j \le n\} \subseteq [0,1],$$

then there is a distribution function F with density f bounded away from 0 and ∞ on $[0,1]$ such that

$$d_n = \sup_{0 \le u \le 1} |F(u) - F_n(u)| \to 0 \quad \text{as } n \to \infty.$$

THEOREM 1.1. *Let K be the integral operator in* (1.2). *An upper bound on the rate of convergence of $\hat{g}_{n\lambda}$ to g is for all $\delta > 0$,*

$$E\|K(\hat{g}_{n\lambda} - g)\|^2_{W_2^q} = O\left[\lambda^{1-2q/5}\|g\|^2_{\mathscr{X}} + S_n\lambda^{-(2q+1)/5}\right]\lambda^{-\delta} \tag{1.3}$$

uniformly for $g \in \mathscr{X}$, $\lambda \in [\lambda_n, \infty)$ *provided for some* $\varepsilon > 0$,

$$(1.4) \qquad \lambda_n \to 0, \qquad d_n^2 \lambda_n^{-(2q/15+2/3+\varepsilon)} \to 0 \quad \text{and} \quad 0 \le q < 2.$$

The proof is given in Section 3. When q is not an integer, W_2^q should be interpreted as the Besov space B_{22}^q [see Cox (1988) and Sections 4.4.2 and 4.6.1 of Triebel (1978) for a discussion of these spaces].

The following result gives the optimal upper bound in (1.3) over the range of $\lambda \in [\lambda_n, \infty)$ where λ_n satisfies (1.4).

COROLLARY 1.1. *If* $g \in \mathscr{X}$, $\lambda^* \sim S_n^{5/6}$, $d_n = O(S_n)$ *and* $S_n \to 0$ *as* $n \to \infty$, *then*

$$E\|K(\hat{g}_{n\lambda^*} - g)\|_{L_2}^2 = O\left(S_n^{5/6+\delta}\right) \quad \text{for all } \delta > 0.$$

Note that λ plays a critical role in the convergence rate and S_n does also, but to a lesser extent. Surprisingly, the asymptotic behavior of the design $\{u_{in}, 1 \le i \le n\}$ is only important for determining the range of λ where these approximations are valid. The fractional constant in the exponent of S_n is largely determined by the asymptotic behavior of the eigenvalues for the integral kernel

$$q(u, v) = \int_0^1 \int_0^1 k(u, x) r(x, y) k(v, y)\, dx\, dy,$$

where k is Abel's kernel and r is the reproducing kernel for $\mathscr{X}$. We have focused on reproducing kernel Hilbert spaces because of this concise characterization.

A relationship between these eigenvalues and the convergence rates of $\hat{g}_{n\lambda}$ to g was originally conjectured in Wahba (1977). Wahba's conjectures are proved in Lukas (1988) and here. Lukas' analysis depends on the assumption that the eigenvalues of q converge to zero at an exact polynomial rate. For Abel's kernel it is not known whether this assumption holds and to investigate this particular integral equation we have weakened the assumption on the rate of eigenvalue decay (see Assumption E, Section 2).

2. Convergence rates for MOR estimators. The first part of this section enlarges the convergence and approximation theorems in Cox (1988) (subsequently referred to as AMORE). This extension is needed for situations where the eigenvalues of Q are not known to decay as an exact power, such as the case for Abel's equation (see Section 3).

First we formulate an abstract version of our problem, state the relevant assumptions, prove an important lemma and then give the main result, Theorem 2.1. The proof of Theorem 2.1 is very similar to that in AMORE and rather than reiterate the entire proof we only state the necessary modifications. The second half of this section translates the fairly abstract assumptions for Theorem 2.1 into a more succinct set of conditions tailored to integral equations.

In the rest of the paper it will be convenient to use the following notation: For $\{a_n\}_{n=1,\infty}$ and $\{b_n\}_{n=1,\infty}$, $a_n \lesssim b_n$ means $a_n = O(b_n)$ as $n \to \infty$. $a_n \simeq b_n$ is equivalent to $a_n \lesssim b_n$ and $b_n \lesssim a_n$. Also, for an operator A with domain $\mathscr{X}$, we will use

$$\mathcal{N}(A) = \{g \in \mathcal{X}: A(g) = 0\}$$

to denote the null space of A.

ASSUMPTION A. Let $\mathcal{Y}_n$ and $\mathcal{X}$ be separable Hilbert spaces with $K_n \in \mathcal{L}(\mathcal{X}, \mathcal{Y}_n)$ for all n. Suppose

$$y_n = K_n g + \varepsilon_n,$$

where ε_n is a random element in $\mathcal{Y}_n$ satisfying $E(\varepsilon_n) = 0$ and $E\langle \eta, \varepsilon_n \rangle_{\mathcal{Y}}^2 \simeq S_n \|\eta\|_{\mathcal{Y}}^2$ uniformly in $\eta \in \mathcal{Y}_n$ for some sequence $\{S_n\} \subseteq (0, \infty)$.

ASSUMPTION B. $W \in \mathcal{L}(\mathcal{X})$ is such that:

(i) $\langle h, Wh \rangle_{\mathcal{X}} \geq 0$ for all $h \in \mathcal{X}$.
(ii) $\dim \mathcal{N}(W) = m < \infty$.
(iii) $\langle h, Wg \rangle_{\mathcal{X}} = \langle Wh, g \rangle_{\mathcal{X}}$ for all g, h in $\mathcal{X}$.

ASSUMPTION C. For all n sufficiently large

$$\mathcal{N}(K_n) \cap \mathcal{N}(W) = \{0\}.$$

The last three assumptions define an operator U that is used to approximate $K_n^* K_n$.

ASSUMPTION D. There is a compact, positive and self-adjoint operator $U \in \mathcal{L}(\mathcal{X}, \mathcal{X})$ with $\mathcal{R}(U)$ dense in $\mathcal{X}$.

From the compactness of U, there is a basis $\{\varphi_\nu\}_{\nu=1,\infty}$ for $\mathcal{X}$ and eigenvalues $0 \leq \gamma_1 \leq \gamma_2 \leq \cdots < \infty$ such that

$$\langle \varphi_\nu, U\varphi_\mu \rangle_{\mathcal{X}} = \delta_{\nu\mu},$$

$$\langle \varphi_\nu, W\varphi_\mu \rangle_{\mathcal{X}} = \gamma_\nu \delta_{\nu\mu},$$

where $\delta_{\nu\mu}$ is Kronecker's delta and for $x \in \mathcal{X}$,

$$x = \sum_{\nu=1}^{\infty} \langle x, U\varphi_\nu \rangle_{\mathcal{X}} \varphi_\nu$$

(see Proposition 2.2 from AMORE).

ASSUMPTION E. There are $0 < r \leq q < \infty$ such that $j^r \lesssim \gamma_j \lesssim j^q$.

For $x \in \mathcal{X}$ let

$$\|x\|_\rho = \left[\sum_{\nu=1}^{\infty} (1 + \gamma_\nu)^\rho \langle x, U\varphi_\nu \rangle_{\mathcal{X}}^2 \right]^{1/2}$$

and $\mathcal{X}_\rho^0 = \{x \in \mathcal{X}: \|x\|_\rho < \infty\}$. Now take $\mathcal{X}_\rho$ to be the completion of $\mathcal{X}_\rho^0$ under $\|\cdot\|_\rho$.

ASSUMPTION F. There exists $s \in (0, 1 - 1/r)$, $\{\rho_1, \rho_2 \cdots \rho_j\} \subset [0, s]$ and $\{d_n\} \subseteq [0, \infty)$ with $d_n \to 0$ such that for all $x_1, x_2 \in \mathscr{X}$,

$$\left|\langle x_1, Ux_2\rangle_{\mathscr{X}} - \langle K_n x_1, K_n x_2\rangle_{\mathscr{Y}_n}\right| \le d_n \sum_{i=1}^{j} \|x_1\|_{\rho_i} \|x_2\|_{s-\rho_i}.$$

REMARKS. These assumptions are identical to those in AMORE except for Assumption E where in AMORE $r = q$.

In this general context, the MOR estimator $\hat{g}_{n\lambda}$ is obtained by the minimization of $\lambda\langle h, Wh\rangle_{\mathscr{X}} + \|K_n h - y_n\|^2_{\mathscr{Y}_n}$ and one can show by standard arguments that $\hat{g}_{n\lambda} = (\lambda W + U_n)^{-1} K_n^* y_n$, where $K_n^* \in \mathscr{L}(\mathscr{Y}_n, \mathscr{X})$ is the adjoint of K_n and $U_n = K_n^* K_n$. The main idea behind our asymptotic representation is to approximate U_n by the "continuous" version U (see Assumption F).

The following quantity figures prominently in our development, and is related to the variance of $\hat{g}_{n\lambda}$:

$$C(\lambda, \rho) = \sum_{j>m} \gamma_j^{\rho}(1 + \lambda\gamma_j)^{-2}. \tag{2.1}$$

LEMMA 2.1. *Fix $0 \le \rho < 2 - 1/r$. Then $C(\lambda, \rho) < \infty$ and as $\lambda \to 0$,*

$$\lambda^{-(r\rho+1)/q} \lesssim C(\lambda, \rho) \lesssim \lambda^{-(\rho+1/r)}, \tag{2.2}$$

$$\lambda^{-(\rho+1/r)}/C(\lambda, \rho) \lesssim \lambda^{-\varepsilon(\rho)} \quad \textit{with } \varepsilon(\rho) = (1 - r/q)(\rho + 1/r). \tag{2.3}$$

Also

$$C(\lambda, \rho) \doteq \lambda^{-2} \quad \textit{as } \lambda \to \infty. \tag{2.4}$$

PROOF. The case $q = r$ was already treated in Theorem 2.4 of Cox (1988), so we only sketch the proof. Define the function

$$\omega(u) = \omega(u; \rho) = u^{\rho}(1 + u)^{-2}.$$

Then

$$C(\lambda, \rho) = \lambda^{-\rho} \sum_{j>m} \omega(\lambda\gamma_j; \rho).$$

Note that ω has its maximum at $u_0 = \rho/(2 - \rho)$. To make Assumption E explicit, suppose

$$\gamma_j \ge M_0 j^r, \qquad \forall\, j,$$

and let

$$n(\lambda) = \max\left\{m, \left[(u_0/\lambda M_0)^{1/r} + 1\right]\right\}.$$

Then, since $\omega(u)$ is decreasing for $u > u_0$,

$$\begin{aligned} C(\lambda, \rho) &\le \lambda^{-\rho}\left[\omega(u_0; \rho) n(\lambda) + \sum_{j \ge n(\lambda)} \omega(\lambda M_0 j^r; \rho)\right] \\ &\le \lambda^{-\rho}\left[\omega(u_0; \rho) n(\lambda) + (M_0\lambda)^{-1/r} \int_{u_0^{1/r}}^{\infty} x^{\rho r}(1 + x^r)^{-2}\, dx\right]. \end{aligned}$$

Note that the integral is finite if $\rho < 2 - 1/r$. If this is the case, then both terms within the brackets in the last expression are $O(\lambda^{-1/r})$ so the upper bound in (2.2) follows. For the lower bound, suppose $\gamma_j \le M_1 j^q$, $j > m$, and use

$$\lambda^{-\rho}\omega(\lambda\gamma_j;\rho) \ge M_0^\rho j^{r\rho}/(1+\lambda M_1 j^q)^2 = \left(M_0^\rho/M_1^{r\rho/q}\right)\lambda^{-r\rho/q}\omega(M_1\lambda j^q; r\rho/q),$$

valid for $j > m$. Hence,

$$C(\lambda,\rho) \ge M\lambda^{-r\rho/q}\sum_{j>m}\omega(M_1\lambda j^q; r\rho/q), \tag{2.5}$$

where $0 < M < \infty$. Now as $\lambda \downarrow 0$,

$$(M_1\lambda)^{1/q}\sum_{j>m}\omega(M_1\lambda j^q; r\rho/q) \to \int_0^\infty x^{r\rho}(1+x^q)^{-2}\,dx.$$

If this limit is used in (2.5), the lower bound in (2.2) is obtained. (2.3) follows directly from this result while (2.4) is established just as in AMORE. □

THEOREM 2.1. *Let $0 \le \rho < 2 - s - 1/r$ and suppose $\rho < \beta \le \rho + 2$ and $\lambda_n \to 0$ as $n \to \infty$. Then under Assumptions* A–E,

$$E\|\hat{g}_{n\lambda} - g\|_\rho^2 \lesssim \min\{1, \lambda^{\beta-\rho}\}\|g\|_\beta^2 + S_n\{C(\lambda,\rho) + m\}$$

uniformly for $\lambda \in [\lambda_n, \infty)$ and $g \in \mathscr{X}_\beta$ for each of the different conditions:

(i) $1 - 1/r \le \rho < 2 - s - 1/r$, $s \le \beta \le 2$, $d_n^2\lambda_n^{-(s+1/r)} \to 0$, $d_n\lambda_n^{-(s+\varepsilon(\rho))} \to 0$.
(ii) $0 \le \rho < 1 - 1/r$, $s \le \beta \le 2$, $d_n^2\lambda_n^{-(s+2\varepsilon(\rho)+1/r)} \to 0$.

PROOF. If $r = q$ in Assumption E, then the theorem follows directly from estimates of the bias of $\hat{g}_{n\lambda}$ in Theorems 4.3(i) and (ii) and 2.3 in AMORE and the approximation of the variance is a consequence of Lemma 4.4 and Theorem 4.5 of AMORE. We will complete the proof by describing the modifications to the proofs of these theorems and lemmas when $r < q$.

The proof of Theorem 2.3 of AMORE is not affected and because Theorem 4.3 only requires upper bounds on $C(\lambda,\rho)$ this proof also does not need to be altered.

We now argue that both Lemma 4.4 and Theorem 4.5 in AMORE hold for the rates on λ_n specified above. For Lemma 4.4 under case (i) we must have

$$S_n d_n \frac{C(\lambda,\rho+s)}{C(\lambda,\rho)} \to 0 \quad \text{for } \lambda \in [\lambda_n, \infty).$$

This follows by Lemma 2.1 and the hypotheses on λ_n.

For case (ii) considering the second upper bound derived for the expression at line (4.5) of AMORE, we must have

$$\frac{d_n C^{1/2}(\lambda, s-\rho_i)C(\lambda,\rho+\rho_i/2)}{C(\lambda,\rho)} \lesssim \frac{d_n\lambda_n^{-(s-\rho_i+1/r)/2}\lambda_n^{-(\rho+\rho_i/2+1/r)}}{\lambda_n^{-(\rho+1/r-\varepsilon(\rho))}}$$
$$= d_n\lambda_n^{-(s+2\varepsilon(\rho)+1/r)/2} \to 0$$

by (2.3) and thus under case (ii) this result will still hold. For Theorem 4.5 of AMORE the estimates of $E\|(\lambda W + U)^{-1}T_n^*\varepsilon_n\|_{\rho_k}^2$ will hold provided that $d_n^2\lambda_n^{-(s+1/r)} \to 0$ as $n \to \infty$. This estimate remains valid because only upper bounds on $C(\lambda, \cdot)$ are used in the argument. Substituting these estimates into the expression at line (4.6) of AMORE, Theorem 4.5 of AMORE will follow if

$$d_n^2\left[\lambda_n^{-(\rho+s+2/r)}\right]/C(\lambda_n, \rho) \to 0 \quad \text{as } n \to \infty.$$

This convergence is a consequence of (2.3) and the hypotheses on λ_n. □

Although Theorem 2.1 restricts attention to the case when $g \in \mathscr{X}_\beta$ and $\beta \le \rho + 2$, the extension to larger β is simple. If $g \in \mathscr{X}_\alpha$ for some $\alpha > \rho + 2$, then the theorem will hold with $\beta = \rho + 2$. In other words, after a certain point increasing the smoothness of g will not improve the convergence rate. This saturation phenomenon is well known in approximation theory. (In nonparametric density estimation this effect is illustrated by the upper limit of $\frac{4}{5}$ for the convergence rate of the mean integrated squared error of a density estimate using a nonnegative kernel.)

This section ends by giving a more convenient set of conditions that imply Assumptions A–F for the model (1.1). Note that Conditions 1 and 2 have already been introduced in Section 1.

CONDITION 3. $\mathscr{X} \subseteq L^2$ is a reproducing kernel Hilbert space with a continuous reproducing kernel r. Let

$$R(h)(u) = \int_{[0,1]} r(u, v)h(v)\, dv$$

and assume $\mathscr{N}(R) = \{0\}$.

CONDITION 4. $K \in \mathscr{L}(L_2[0,1])$ with $\mathscr{N}(K) = \{0\}$.

Note that $R \in \mathscr{L}(L_2)$ is a compact operator by Mercer's theorem, so

$$Q = KRK^* \in \mathscr{L}(L_2)$$

is also compact and hence has a spectral decomposition with eigenvalues

$$\mu_1 \ge \mu_2 \ge \mu_3 \ge \cdots > 0, \tag{2.6}$$

where $\mu_j \to 0$ as $j \to \infty$. (Positivity of the eigenvalues follows as $\langle h, Qh\rangle_{L_2} = \langle K^*h, RK^*h\rangle_{L_2} > 0\ \forall\, h \ne 0$ by Conditions 3 and 4.)

CONDITION 5. The eigenvalues (2.6) of Q satisfy

$$j^{-q} \lesssim \mu_j \lesssim j^{-r}.$$

CONDITION 6. There is a $p \ge 0$ such that $1/p < 1 - 1/r$ and $K \in \mathscr{L}(\mathscr{X}, W_2^p)$.

CONDITION 7. $W \in \mathscr{L}(\mathscr{X})$ and W is self-adjoint and nonnegative definite. If $\tilde{K} \in \mathscr{L}(\mathscr{X}, L_2(F))$ denotes the restriction of K to $\mathscr{X}$, then there is a $0 < M < \infty$

such that

$$\|(W + \tilde{K}^*\tilde{K})h\|_{\mathscr{X}} > M\|h\|_{\mathscr{X}} \quad \forall\, h \in \mathscr{X}.$$

Note that in Condition 4, K is regarded as having domain L_2, whereas $\tilde{K}$ has domain $\mathscr{X}$. This distinction is important because $\tilde{K}^*$ and K^* are different.

We identify $\mathscr{Y}_n$ with $\mathbb{R}^n$ equipped with the inner product $(1/n)\eta'\zeta$ for $\eta, \zeta \in \mathbb{R}^n$, and let $K_n(h)' = (Kh(u_{1n}), Kh(u_{2n}), \dots, Kh(u_{nn}))$. In the following discussion it will be shown that Conditions 1–7 imply the more abstract Assumptions A–F.

Note that

$$\frac{1}{n}\|K_n(h)\|^2_{\mathbb{R}^n} \le \sup|K(h)|^2$$

and thus the continuity of K_n is a consequence of $W_2^1 \subset C^0$ and Condition 6. The remaining parts of Assumptions A and B follow from Conditions 1 and 7 and the discussion in Section 1. The discussion of Assumption C will be deferred until Assumptions D–F have been established.

In order to deal with Assumption D, we first motivate the choice of the limiting operator U. Referring to Assumption F, we wish to approximate (for $h_1, h_2 \in \mathscr{X}$),

$$\begin{aligned} \langle K_n h_1, K_n h_2\rangle_{\mathscr{Y}_n} &= \frac{1}{n}\sum_{i=1}^{n}\left(\int k(u_i, v)h_1(v)\,dv\right)\left(\int k(u_i, v)h_2(v)\,dv\right) \\ &= \int\!\!\int h_1(v)w_n(v, v')h_2(v')\,dv\,dv', \end{aligned} \tag{2.7}$$

where the kernel w_n is given by

$$w_n(v, v') = \int k(u, v)k(u, v')\,dF_n(u).$$

If we replace w_n by its continuous analog

$$w(v, v') = \int k(u, v')k(u, v)\,dF(u),$$

then the last expression in (2.7) becomes

$$\int\!\!\int h_1(v)w(v, v')h_2(v')\,dv\,dv' = \langle \tilde{K}h_1, \tilde{K}h_2\rangle_{L_2(F)} = \langle h_1, \tilde{K}^*\tilde{K}h_2\rangle_{\mathscr{X}},$$

where $\tilde{K}$ is given in Condition 7. Hence,

$$U = \tilde{K}^*\tilde{K} \tag{2.8}$$

is the correct choice.

The compactness of $\tilde{K}$ follows from the composition $\tilde{K} = J \circ \overline{K}$, where by Condition 6, $\overline{K} \in \mathscr{L}(\mathscr{X}, W_2^p)$ and J is the compact embedding operator from W_2^p to L_2 for $p \ge 1$. U must also be compact since $U = \tilde{K}^*\tilde{K}$, where $\tilde{K}^*$ is continuous.

Assumption D now follows from the compactness of U and the fact that $\mathscr{N}(U) = \{0\}$.

Before discussing Assumption E we state some properties for the integral operator R from Condition 3. Because r is a continuous reproducing kernel, $R^{1/2} \in \mathscr{L}(L_2)$ is a well defined, self-adjoint square root of R that is also compact [Riesz and Nagy (1955), pages 242–246]. Also, $R^{1/2}$: $L_2 \to \mathscr{X}$ is an isometric isomorphism between these two Hilbert spaces.

The next lemma verifies Assumption E.

LEMMA 2.2. *Under Conditions* 1–7, $1/\mu_j \simeq \gamma_j$.

PROOF. From the discussion in AMORE Section 2, the eigenvalues $\alpha_j = 1/(1 + \gamma_j)$ can be generated by successive minimizations of the Rayleigh quotient

$$\langle x, Ux\rangle_{\mathscr{X}}/\langle x, (U + W)x\rangle_{\mathscr{X}}.$$

Using Condition 7, the boundedness of dF/du away from 0 and ∞ (Condition 2) and the isometric isomorphism between L_2 and $\mathscr{X}$, there are constants $0 < C_1, C_2 < \infty$ such that for all $x \in \mathscr{X}$,

$$C_1 \frac{\langle f, A^*Af\rangle_{L_2}}{\|f\|_{L_2}^2} \le \frac{\langle x, Ux\rangle_{\mathscr{X}}}{\langle x, (U + W)x\rangle_{\mathscr{X}}} \le C_2 \frac{\langle f, A^*Af\rangle_{L_2}}{\|f\|_{L_2}^2},$$

with $f = R^{-1/2}x$ and $A = KR^{1/2}$. If $\{\beta_j\}_{j=1,\infty}$ are the eigenvalues of A^*A (including multiplicities), then by applying the mapping principle to these inequalities [see Weinberger (1974), page 57] it follows that $\beta_j \lesssim \alpha_j \lesssim \beta_j$, or in other terms, $1/\beta_j \simeq \gamma_j$. Finally, we note that $\mu_j = \beta_j$ because $\mathscr{N}(A) = \{0\}$ so the eigenvalues of A^*A are the same as those of $AA^* = Q$. □

LEMMA 2.3. *Condition* 6 *implies assumption* F *with* $s = 1/p$, $\rho_1 = 0$, $\rho_2 = 1/p$ *and* $d_n = \sup|F - F_n|$.

PROOF. The proof is given in the remarks following Assumption 4.1 of Cox (1988), so we only sketch it here. For $h_1, h_2 \in \mathscr{X}$, an integration by parts followed by the Cauchy–Schwarz inequality gives

$$\begin{aligned}|\langle h_1, Uh_2\rangle_{\mathscr{X}} - \langle K_n h_1, K_n h_2\rangle_{\mathscr{Y}_n}| &= \left|\int (Kh_1)(Kh_2)\,d(F - F_n)\right| \\ &\le d_n\left(\|Kh_1\|_{W_2^1}\|Kh_2\|_{L_2} + \|Kh_1\|_{L_2}\|Kh_2\|_{W_2^1}\right),\end{aligned}$$

and to complete the proof, we argue that

$$\|Kh\|_{W_2^1} \le M\|h\|_{1/p}, \qquad \forall\, h \in \mathscr{X}$$

for some constant $M \in (0, \infty)$. This inequality can be established from standard interpolation theory and the details are given in Section 3 of AMORE. [See also the introductions to interpolation theory in Triebel (1978), pages 18–27, or

Butzer and Behrens (1967).] Applying the K method of interpolation functor to the pairs of spaces $(\mathscr{X}_0, \mathscr{X}_1)$ and $(L_2(F), W_2^p)$, it follows that $\mathscr{X}_{1/p} = (\mathscr{X}_0, \mathscr{X}_1)_{1/p,2}$ and $W_2^1 = (L_2(F), W_2^p)_{1/p,2}$. Now $K \in \mathscr{L}(\mathscr{X}_1, W_2^p)$ and $K \in \mathscr{L}(\mathscr{X}_0, L^2(F))$ and therefore $K \in \mathscr{L}(\mathscr{X}_{1/p}, W_2^1)$ by the fundamental properties of the interpolation functor. □

We complete this discussion by treating Assumption C. The argument is by contradiction. Suppose that for any n there is a $\zeta_n \in \mathscr{N}(W)$ such that $K_n(\zeta_n) = 0$ but $\zeta_n \neq 0$. Without loss of generality, we can take $\langle U\zeta_n, \zeta_n\rangle = 1$. By Assumption F and by the fact that $\mathscr{N}(W)$ is finite dimensional, one can make n sufficiently large such that

$$|\langle U_n\zeta_n, \zeta_n\rangle_{\mathscr{X}} - \langle U\zeta_n, \zeta_n\rangle_{\mathscr{X}}| < \tfrac{1}{2}.$$

Thus

$$1 = \langle U\zeta_n, \zeta_n\rangle_{\mathscr{X}} < \tfrac{1}{2} + \langle U_n\zeta_n, \zeta_n\rangle_{\mathscr{X}} = \tfrac{1}{2},$$

a contradiction.

One problem we have not dealt with so far in this section is the explicit description of the spaces $\mathscr{X}_\rho$ for $\rho \geq 0$. Only for $\rho = 1$ and $\rho = 0$ is there an easily obtainable equivalent norm, namely,

$$\|h\|_1 \simeq \|h\|_{\mathscr{X}}, \qquad \|h\|_0 \simeq \|Kh\|_{L_2}.$$

The first follows from $\|h\|_1^2 = \langle h, (W + \tilde{K}^*\tilde{K})h\rangle_{\mathscr{X}}$ and Condition 7. The second relation follows since $\|h\|_0^2 = \langle h, Uh\rangle_{\mathscr{X}} = \|Kh\|_{L_2(F)}$ and $L_2(F)$ is equivalent to L_2 by Condition 2. Furthermore, as W and F are not involved in $\|h\|_{\mathscr{X}}$ and $\|Kh\|_{L_2}$, we have that different choices of W and F (subject to Conditions 2, 4 and 7) lead to equivalent $\mathscr{X}_\rho$ norms for $0 \leq \rho \leq 1$. For $\rho = 0, 1$, this follows as above, while for $0 < \rho < 1$, this follows from the fact that the $\mathscr{X}_\rho$ norm is equivalent to the norm on $(\mathscr{X}_0, \mathscr{X}_1)_{\rho,2}$ given by the K-method of interpolation. Finally, we should note that the interpolation argument outlined in Lemma 2.3 generalizes to give the bound

$$\|Kg\|_{W_2^{p\rho}} \leq M\|g\|_\rho \quad \forall\, g \in \mathscr{X} \text{ and } 0 \leq \rho \leq 1.$$

3. Abel's kernel. Theorem 1.1 is proved by verifying Conditions 3–7 of Section 2 and then applying Theorem 2.1. We also discuss the regularization of histograms to estimate a particle size distribution in stereology.

Let

$$\mathscr{X} = W_{2,BC}^k = \{h \in W_2^k[0,1]\colon h^{(j)}(1) = 0, 0 \leq j \leq k-1\}$$

for some $k \geq 2$ and take as an inner product

$$\langle f, h\rangle_{\mathscr{X}} = \langle f^{(k)}, h^{(k)}\rangle_{L_2}.$$

The reproducing kernel for $\mathscr{X}$ with this inner product is

$$r(u,v) = \frac{1}{(k!)^2}\int_0^1 (w-u)_+^{k-1}(w-v)_+^{k-1}\,dw,$$

where $(x)_+ = \max\{x, 0\}$. Now r is continuous on $[0,1]^2$. It is straightforward to show that $\mathscr{N}(R) = \{0\}$ because the basis $\{r(\cdot, u)\colon u \in [0,1]\}$ spans $\mathscr{X}$ [Aubin (1979), page 116] and $\mathscr{X} = W^k_{2,BC}$ is dense in L_2. Thus, Condition 3 holds.

Turning to Condition 4, we will use the following lemma.

LEMMA 3.1. (i) $K \in \mathscr{L}(L_2, L_q)$ *for all* $1 \le q < \infty$.
(ii) $K \in \mathscr{L}(\mathscr{X}, W_2^\delta)$ *for* $\delta < k + \frac{1}{2}$.

PROOF.

$$\begin{aligned}\int_0^1 |Kh|^q\,dv &\le \int_0^1 \prod_{j=1}^q \int_v^1 \frac{|h(u_j)|}{\sqrt{u_j - v}}\,du_j\,dv \quad \text{(by Fubini's theorem)} \\ &\le \int_{[0,1]^q} \psi(\mathbf{u}) \prod_{j=1}^q |h(u_j)|\,d\mathbf{u},\end{aligned} \tag{3.1}$$

where

$$\psi(\mathbf{u}) = \int_{[0,1]} \prod_{j=1}^q \frac{I(u_j > v)}{\sqrt{u_j - v}}\,dv.$$

Also, it is straightforward to show

$$\int_{[0,1]^q} \psi(\mathbf{u})^2\,d\mathbf{u} = \int_0^1 \int_0^1 \left[\int_{\max(v_1, v_2)}^1 \frac{du}{\sqrt{(u - v_1)(u - v_2)}} \right]^q dv_1\,dv_2$$

and from Selby [(1979), page 430],

$$\begin{aligned} &= 2\int_0^1 \int_0^{v_2} \log^q\left(\frac{\sqrt{v_2} + \sqrt{v_1}}{\sqrt{v_2} - \sqrt{v_1}} \right) dv_1\,dv_2 \\ &= 2\int_0^1 \int_0^1 v_2 w \log^q\left(\frac{1 + w}{1 - w} \right) dw\,dv_2 = M < \infty. \end{aligned}$$

Applying the Cauchy–Schwarz inequality to (3.1) now gives

$$\int (Kh)^q\,dv \le \sqrt{M} \left[\int h^2\,du \right]^{q/2}$$

and (i) follows.

For $h \in \mathscr{X}$, integrating by parts l times and applying the boundary conditions for members of $\mathscr{X}$ it is easy to show that

$$D^l Kh = (-1)^l K D^l h \quad \text{for } l \le k \tag{3.2}$$

and from part (i) we have $K \in \mathscr{L}(\mathscr{X}, W_q^k)$ for $1 \le q < \infty$. Now $W_q^k \subset W_2^{k + 1/2 - \varepsilon}$ provided $1/q < \varepsilon$ [Triebel (1978), Section 4.6.1] and thus (ii) holds. □

The first part of Condition 4 now follows from Part (i) of the lemma. In order to prove that the null space of K only contains 0, we will use the inversion formula $Kh = g \Leftrightarrow Kg$ absolutely continuous and $-(1/\pi)DKg = h$ a.e. [Cochran (1972), page 7]. Thus,

$$Kh = 0 \Rightarrow h = 0 \text{ a.e. } \Rightarrow \mathscr{N}(K) = \{0\}.$$

Before verifying Condition 5, it will be useful to cite the relationship of K and R to fractional integration. For $\alpha > 0$, put

$$(I_\alpha h)(u) = \frac{1}{\Gamma(\alpha)} \int_u^1 (v-u)^{\alpha-1} h(v)\, dv.$$

Then $K = \pi^{1/2} I_{1/2}$, while $R = I_k I_k^*$, where $I_k \in \mathscr{L}(L_2)$ if $k \geq \frac{1}{2}$. I_α is an extension of α-times iterated integration to fractional orders [Ross (1975)]. Furthermore, one can show

$$I_\alpha I_\beta = I_{\alpha+\beta}, \qquad \alpha, \beta > 0, \tag{3.3}$$

and hence

$$Q = KRK^* = \pi I_{k+1/2} I_{k+1/2}^*. \tag{3.4}$$

This latter formula is now used to estimate the eigenvalues of Q.

LEMMA 3.2. *The eigenvalues* $\{\mu_j\}$ *of* $I_{k+1/2} I_{k+1/2}^*$ *satisfy*

$$j^{-2(k+1)} \lesssim \mu_j \lesssim j^{-(2k+1)}$$

and hence Condition 5 *holds with* $q = 2(k+1)$ *and* $r = 2k+1$.

PROOF. The upper bound follows from Theorem 3.2 of Faber and Wing (1985) and (3.4). For the lower bound set $M = \|I_{1/2}^*\|_{\mathscr{L}(L_2)}$ and using (3.3) one can show

$$M^{-2} \|I_{k+1}^* h\|_{L_2}^2 \leq \langle I_{k+1/2} I_{k+1/2}^* h, h \rangle_{L_2},$$

where $0 < M < \infty$. Let $\{\nu_j^{(i)}\}$ denote the eigenvalues of $I_i I_i^*$. Then by the mapping principle [Theorem 3.6.1 of Weinberger (1974)] and the above inequalities, $M^{-2}\nu_j^{(k+1)} \leq \mu_j$. $I_i I_i^*$ is the Green's operator for the differential operator D^{2i} with boundary conditions

$$f^{(0)}(1) = f^{(1)}(1) = \cdots = f^{(i-1)}(1) = 0$$

and

$$f^{(i)}(0) = \cdots = f^{(2i-1)}(0) = 0.$$

The standard theory for the asymptotics of the eigenvalues of such operators [e.g., Naimark (1967), Section 4 or Triebel (1978), page 392] yields

$$\nu_j^{(i)} \approx j^{-2i}. \qquad \square$$

Condition 6 follows directly from Lemma 3.1(ii) with $p < k + \frac{1}{2}$ and Condition 7 holds because $W = I$.

We now can apply Theorem 2.1 to obtain general convergence results for the regularization of Abel's equation.

Set

$$\varepsilon(\rho) = \frac{1}{2(k+1)}\left(\rho + \frac{1}{2k+1}\right) \quad \text{and} \quad s > \frac{2}{2k+1}.$$

Note that $W = I$, $m = 0$ and by Lemma 2.1, $C(\lambda, \rho) < \lambda^{-(\rho+1/(2k+1))}$. Now for some $\varepsilon > 0$ let $\nu = 2 - s - 1/r - \varepsilon = 2 - 3/(2k+1) - \varepsilon$. Assume $\lambda_n \to 0$ and also that

1. $2k/(2k+1) \leq \rho < \nu$,

$$d_n^2 \lambda_n^{-(3/(2k+1)+\varepsilon)} \to 0$$

and

$$d_n \lambda_n^{-(2/(2k+1)+\varepsilon(\rho)+\varepsilon)} \to 0$$

or

2. $0 < \rho < 2k/(2k+1)$ and $d_n^2 \lambda_n^{-(3/(2k+1)+\varepsilon+2\varepsilon(\rho))} \to 0$.

From Theorem 2.1 if

$$g \in \mathscr{X}_\beta, \qquad \rho < \beta \leq \rho + 2 \text{ and } s \leq \beta \leq 2,$$

then for $\lambda \in [\lambda_n, \infty)$,

$$E\|\hat{g}_{n\lambda} - g\|_\rho^2 \leq \min\{1, \lambda^{(\beta-\rho)}\}\|g\|_\beta^2 + S_n \lambda^{-(\rho+1/(2k+1))}. \tag{3.5}$$

The interpolation norm in (3.5) can be bounded by a Sobelev norm using the last inequality of Section 2 and Lemma 3.2. For any $\delta > 0$ such that $\rho + \delta \leq 1$ there is an $M < \infty$ so that

$$\|Kh\|_{W_2^{(k+1/2)\rho}} \leq M\|h\|_{\rho+\delta} \quad \text{for all } h \in \mathscr{X}_{\rho+\delta}.$$

Theorem 1.1 now follows with $k = 2$ and $\beta = 1$. Corollary 1.1 is established by minimizing the upper bound in (3.5).

We now describe our proposed estimator for the stereology problem and give upper bounds on its convergence rate. Suppose that the centers of spheres embedded in a medium follow a three-dimensional Poisson process with rate Λ_3 per unit volume and that the radius R of a sphere is a random variable independent of the centers and radii of the other spheres. Let $A_3 = \pi R^2$ denote the equatorial area. We assume A_3 has Lebesgue density f_2, with support on $[0, 1]$ (any other upper bound can be accommodated by rescaling) and let

$$h_3(a) = \Lambda_3 f_3(a).$$

Now suppose that the medium is sliced by a plane whose orientation is independent of the spheres' locations. Spheres that have been cut produce circular cross sections on the plane. If attention is restricted to those cross sections whose centers lie in a region with area A then the number of circular cross sections with area between a_0 and a_1 ($0 \leq a_0 < a_1 \leq 1$) is a Poisson random variable with mean

$$A\int_{a_0}^{a_1} h_2(a)\, da.$$

Moreover

$$h_2(a) = 2(Kh_3)(a),$$

K being Abel's operator in (1.2). This latter equation follows from a change of variables in (5) of Watson (1971). For the histogram bin limits $0 = u_{0n} < u_{1n} < \cdots < u_{nn} = 1$, let m_{in} denote the number of cross sections with areas in $[u_{i-1,n}, u_{in})$. We take as "observations" $Y_{in} = m_{in}/[A(u_{in} - u_{i-1,n})]$ and set $g = 2h_3$. $E(Y_n) = T_n(g)$ where the ith element of $T_n g$ has the form

$$T_{in}(g) = \frac{1}{(u_{in} - u_{i-1,n})} \int_{u_{i-1,n}}^{u_{in}} Kg(v)\, dv.$$

Note that the number of histogram bins is actually the relevant sample size, not the actual number of observed cross sections.

In order to study the MOR estimator in this context we will determine S_n in Condition 1. Because the components T_n are not exactly in the form of evaluation functionals, Assumption E will not follow from the Conditions 1–7 and must be shown directly. Nevertheless we will argue that $U = \tilde{K}^*\tilde{K}$ is still the appropriate asymptotic approximation to $T_n^* T_n$. The other assumptions follow from having verified Conditions 3–7 for Abel's kernel above.

Let $\varepsilon_n = Y_n - EY_n$. Then

$$\begin{aligned} E\left|\frac{1}{n}\varepsilon_n'\eta\right|^2 &= \frac{1}{n^2}\sum_{i=1}^{n} \eta_i^2 \operatorname{Var}(Y_{in}) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \eta_i^2 (A(u_{in} - u_{i-1,n}))^{-2} A \int_{u_{i-1,n}}^{u_{in}} Kg(v)\, dv \end{aligned}$$

and, using the mean value theorem,

$$\le [nA \min(u_{in} - u_{i-1,n})]^{-1} \sup_{0 \le u \le 1} |Kg(u)| \frac{1}{n}\eta'\eta.$$

As already noted above $Kg \in W_2^1$ if $g \in \mathscr{X}$ and hence $\sup|Kg| < \infty$. Thus, $E|(1/n)\varepsilon_n'\eta|^2 \lesssim S_n(1/n)\eta'\eta$ with $S_n = [nA \min_{1 \le i \le n}(u_{in} - u_{i-1,n})]^{-1}$.

We will now argue that Assumption E holds with

$$d_n = \max\left[\sup_{0 \le u \le 1} |F_n(u) - F(u)|, \Delta_n\right], \tag{3.6}$$

where F_n is the empirical distribution function for $\{u_{in}\}$ and

$$\Delta_n = \max_{1 \le i \le n} |u_{in} - u_{i-1,n}|.$$

For $f, g \in \mathscr{X}$, by the mean value theorem we can choose $\{\bar{u}_{in}\}$, $\{u_{in}^*\}$ and $\{\bar{\bar{u}}_{in}\}$ such that

$$\begin{gathered} u_{i-1,n} \le \bar{u}_{in}, u_{in}^*, \bar{\bar{u}}_{in} \le u_{in}, \\ T_{in}(f) = Kf(\bar{u}_{in}) \end{gathered} \tag{3.7}$$

and

$$T_{in}(g) = Kg(u_{in}^*) = Kg(\bar{u}_{in}) + \frac{d}{du}Kg(\bar{\bar{u}}_{in})(\bar{u}_{in} - u_{in}^*).$$

Take $\bar{K}_n \in \mathscr{L}(\mathscr{X}, \mathbb{R}^n)$ to be the operator so that $\bar{K}_n h$ evaluates Kh at the points $\{\bar{u}_{in}\}$. For $f, g \in \mathscr{X}$ it follows that

$$|(T_n^* T_n f, g)_{\mathscr{X}} - (Uf, g)_{\mathscr{X}}|$$
$$\leq |\langle \bar{K}_n^* \bar{K}_n f, g\rangle_{\mathscr{X}} - \langle Uf, g\rangle_{\mathscr{X}}| + \Delta_n \frac{1}{n} \sum_{i=1}^{n} \left| Kf(\bar{u}_{in}) \frac{d}{du} Kg(\bar{\bar{u}}_{in}) \right|.$$

Now Assumption F can be applied to the first term of this expression. Thus,

$$\leq \bar{d}_n(\|f\|_0 \|g\|_{1/p} + \|f\|_{1/p} \|g\|_0) + \Delta_n \sup|Kf| \sup\left|\frac{d}{du} Kg\right|,$$

where $\bar{d}_n = \sup|\bar{F}_n - F|$ and $p < k + \frac{1}{2}$.

For all $\varepsilon > \delta > 0$,

$$\sup|Kf| \lesssim \|Kf\|_{W_2^{1/2+\delta}} \lesssim \|f\|_{(1/2+\varepsilon)/p}$$

and similarly,

$$\sup\left|\frac{d}{du} Kg\right| \lesssim \|g\|_{(3/2+\varepsilon)/p}.$$

Therefore, Assumption F will hold with $s > 4/(2k+1)$, $\rho_1 = 0$ and $\rho_2 > 1/(2k+1)$. Finally, we note that by the triangle inequality and (3.7),

$$\bar{d}_n \leq \sup|F_n - F| + \frac{1}{n} \quad \text{and so} \quad \bar{d}_n \triangleq \sup|F_n - F|.$$

For an example of this application of Abel's equation take $\mathscr{X} = W_{2,BC}^2[0,1]$, $k = 2$ and suppose that

$$\max_{1 \leq i \leq n} (u_{i,n} - u_{i-1,n}) \lesssim \min_{1 \leq i \leq n} (u_{i,n} - u_{i-1,n}) \quad \text{as } n \to \infty.$$

Then

$$d_n \triangleq n^{-1}, \qquad S_n \triangleq A^{-1}$$

and so if $g \in W_{2,BC}^2$, then the upper bound in Theorem 1.1 will hold provided that $n \to \infty$, $A \to \infty$, $\lambda \to 0$ and $n^{-2}\lambda_n^{-(16/15+2q/15+\varepsilon)} \to 0$ for some $\varepsilon > 0$.

Typically, one is interested in the size distribution of the spheres in terms of radius rather than area. If $\hat{g}_{n\lambda}$ is normalized to integrate to 1, then an estimate for the radial density can be obtained from a simple change of variables.

REFERENCES

ANDERSSEN, R. S. (1976). Stable procedures for the inversion of Abel's equation. *Inst. Math. Appl.* **17** 329–342.

ANDERSSEN, R. and JAKEMAN, A. (1975a). Product integration for functionals of particle size distributions. *Utilitas Math.* **8** 111–126.

ANDERSSEN, R. and JAKEMAN, A. (1975b). Abel type integral equations in stereology, II. Computational methods of solution and the random spheres approximation. *J. Microscopy* **105** 135–153.

ARONSZAJN, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404.

AUBIN, J. (1979). *Applied Functional Analysis*. Wiley, New York.

BULLEN, H. E. (1963). *An Introduction to Seismology*. Cambridge Univ. Press, Cambridge.

BUTZER, P. and BEHRENS, H. (1967). *Semigroups of Operators and Approximation*. Springer, New York.

COCHRAN, J. A. (1972). *The Analysis of Linear Integral Equations*. McGraw-Hill, New York.

COX, D. D. (1984). Gaussian approximation of smoothing splines. Technical Report 743, Dept. Statistics, Univ. Wisconsin, Madison.

COX, D. D. (1988). Approximation of method of regularization estimators. *Ann. Statist.* **16** 694–712.

FABER, V. and WING, G. (1985). Singular values of fractional integral operators: A unification of theorems of Hille–Tamarkin and of Chang. Technical Report LA-UR-85-1853, Los Alamos National Laboratory.

FRANKLIN, J. (1981). Confidence limits for stereological estimators with infinite variance. *SIAM J. Appl. Math.* **40** 179–190.

HALL, P. and SMITH, R. L. (1988). The kernel method for unfolding sphere size distributions. *J. Comput. Phys.* **74** 409–421.

KIMELDORF, G. and WAHBA, G. (1970). A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. *Ann. Math. Statist.* **41** 492–502.

KOSAREV, E. L. (1980). Applications of integral equations of the first kind in experimental physics. *Comput. Phys. Comm.* **20** 69–75.

LI, K. C. (1982). Minimaxity of the method of regularization on stochastic processes. *Ann. Statist.* **10** 937–942.

LUKAS, M. (1980). Regularization. In *Numerical Solution of Integral Equations* (R. Anderssen, ed.) 152–182. Sijthoff and Noordoff, Groningen.

LUKAS, M. (1981). Regularization of linear operator equations. Ph.D. dissertation, Australian National Univ.

LUKAS, M. (1988). Convergence rates for regularized solutions. *Math. Comp.* **51** 107–131.

NAIMARK, M. A. (1967). *Linear Differential Operators, Part I.* Ungar, New York.

NASHED, M. and WAHBA, G. (1974). Generalized inverses in reproducing kernel spaces: An approach to regularization of linear operator equations. *SIAM J. Math. Anal.* **5** 974–987.

NYCHKA, D., WAHBA, G., GOLDFARB, S. and PUGH, T. (1984). Cross-validated spline methods for the estimation of three-dimensional tumor size distributions from observations on two-dimensional cross sections. *J. Amer. Statist. Assoc.* **79** 832–846.

O'SULLIVAN, F. (1986). A statistical perspective on ill-posed inverse problems. *Statist. Sci.* **1** 502–518.

RIESZ, R. and NAGY, B. SZ. (1955). *Functional Analysis*. Ungar, New York.

ROSS, B., ed. (1975). *Fractional Calculus and Its Applications. Lecture Notes in Math.* **457**. Springer, New York.

SELBY, S., ed. (1979). *CRC Standard Mathematical Tables*. Central Rubber Co., Cleveland, O.

SPECKMAN, P. (1985). Spline smoothing and optimal rates of convergence in nonparametric regression models. *Ann. Math. Statist.* **13** 970–984.

TIKHONOV, A. N. (1963). Solution of incorrectly posed problems and the method of regularization. *Soviet Math. Dokl.* **4** 1035–1038.

TRIEBEL, H. (1978). *Interpolation Theory, Function Spaces, Differential Operators*. North-Holland, Amsterdam.

WAHBA, G. (1977). Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.* **14** 651–667.

WAHBA, G. (1980). Ill-posed problems: Numerical and statistical methods for mildly, moderately, and severely ill-posed problems with noisy data. Technical Report 595, Dept. Statistics, Univ. Wisconsin, Madison.

WATSON, G. S. (1971). Estimating functionals of particle size distributions. *Biometrika* **58** 483–490.

WEINBERGER, H. (1974). *Variational Methods for Eigenvalue Approximation*. Arrowsmith, Bristol, England.

DEPARTMENT OF STATISTICS
NORTH CAROLINA STATE UNIVERSITY
RALEIGH, NORTH CAROLINA 27695-8203

DEPARTMENT OF STATISTICS
UNIVERSITY OF ILLINOIS
725 SOUTH WRIGHT STREET
CHAMPAIGN, ILLINOIS 61820