Fast Algorithms for Problems on Thermal Tomography

Raymond H. Chan[*] Chun-pong Cheung[†] Hai-wei Sun[‡]

September 20, 1996

Abstract

In this paper, we study an ill-posed, nonlinear inverse problem in heat conduction and hydrology applications. In [2], the problem is linearized to give a linear integral equation, which is then solved by the Tikhonov method with the identity as the regularization operator. We prove in this paper that the resulting equation is well-condition and has clustered spectrum. Hence if the conjugate gradient method is used to solve the equation, we expect superlinear convergence. However, we note that the identity operator does not give good solution to the original equation in general. Therefore in this paper, we use the Laplacian operator as the regularization operator instead. With the Laplacian operator, the regularized equation is ill-conditioned and hence a preconditioner is required to speed up the convergence rate if the equation is solved by the conjugate gradient method. We here propose to use the Laplacian operator itself as preconditioner. This preconditioner can be inverted easily by fast sine-transforms and we prove that the resulting preconditioned system is well-conditioned and has clustered spectrum too. Hence the conjugate gradient method converges superlinearly for the preconditioned system. Numerical results are given to illustrate the fast convergence.

1 Introduction

In this paper, we consider the inverse problem of finding a pair (u, a) for the Cauchy problem

$$\begin{array}{rcll} u_t - \mathrm{div}(a\nabla u) &=& \delta(x-x^*)\delta(t) & \text{on} \quad \mathbb{R}^m \times (0,T), \\ u &=& 0 & \text{on} \quad \mathbb{R}^m \times \{0\}, \end{array} \tag{1}$$

where δ is the Dirac delta function. The unknown coefficient function $a = a(x)$ is assumed to be bounded and measurable and has the expansion $a(x) = 1 + f(x)$, where $f(x)$ is zero outside a bounded domain $\Omega \subset \mathbb{R}^m$. The solution $u(x, t; x^*)$ is assumed to be given for $x, x^* \in \Omega^*$ and $t \in (0, T)$ where Ω^* is a bounded domain in $\mathbb{R}^m$ whose closure does not intersect Ω. This kind of problems arises in the heat conduction and in hydrology applications. The diffusion coefficient $a(x)$ is to be recovered from the boundary measurements of the solution $u(x, t; x^*)$ in (1). In real applications, x^* is the source and x is the sensor.

[*]Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong. Research supported by CUHK Direct Grant 220600680.

[†]Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong. Research supported by CUHK Grants for Overseas Academic Activities.

[‡]Department of Mathematics and Physics, Guangdong University of Technology, Guangzhou, People's Republic of China.

Since (1) is ill-posed and nonlinear, it is difficult to solve. Recently, Elayyan and Isakov have derived in [2] a linearized equation for (1). After linearization, the unknown function f is given as the solution of a linear convolution integral equation. Since the integral equation is also ill-posed, in [2], the equation is solved by using the Tikhonov regularization method with the identity operator as the regularization operator.

In this paper, we first show that if the identity operator is used as the regularization operator, then the resulting operator is well-conditioned with clustered spectrum. Hence if the equation is solved by the conjugate gradient method, then we have superlinear convergence. However, we will illustrate by examples that the identity operator does not give good solution for the original integral equation. In this paper, we therefore use the Laplacian operator as the regularization operator instead. Our numerical examples show that it is a better regularization operator than the identity operator in that the solutions thus obtained are more accurate.

The main drawback in using the Laplacian operator is that the regularized equation becomes ill-conditioned. If the conjugate gradient method is used to solve the equation, the iteration number required for convergence will increase with the size of the discretization matrix. To speed up the convergence rate, we propose to use the Laplacian itself as the preconditioner for the equation. For all 1-dimensional problems or 2-dimensional problems on rectangular domains, the inverse of the Laplacian can easily be found by using fast sine-transforms. We prove that the preconditioned system is well-conditioned and has clustered spectrum. Therefore the convergence rate of the conjugate gradient method for the preconditioned system is expected to be superlinear. Numerical results are given to illustrate this fast convergence.

The outline of the paper is as follows. In §2, we briefly recall the linearization scheme in [2]. In §3, we discuss the identity regularization methods used in [2] and show that the resulting operator has clustered spectrum. Numerical results are given in §4 to illustrate that the Laplacian operator will give much better results than the identity operator. Then in §5, we show that with the Laplacian operator as preconditioner, the resulting preconditioned operator is still well-conditioned and has clustered spectrum. Finally, concluding remarks are given in §6.

2 Linearization of the Inverse Problem

In this section, we review the linearization method proposed by Elayyan and Isakov [2] for the inverse problem (1). We first consider $\tilde{u}$, the zeroth order approximation to u, which satisfies

$$\begin{array}{rcll} \tilde{u}_t - \Delta \tilde{u} & = & \delta(x-x^*)\delta(t) & \text{on} \quad \mathbb{R}^m \times (0,T), \\ \tilde{u} & = & 0 & \text{on} \quad \mathbb{R}^m \times \{0\}. \end{array} \tag{2}$$

We note that (2) can be obtained from (1) by replacing $a(x)$ there by 1, the zeroth order approximation of $a(x)$. We remark that there is a closed form solution in integral form for $\tilde{u}$ in (2), see [2].

Now we let $\tilde{v} = u - \tilde{u}$ be the residual of the approximation. Substituting this expansion into (1), we get

$$\begin{array}{rcll} \tilde{v}_t - \Delta \tilde{v} & = & \operatorname{div}(f\nabla\tilde{u}) + \operatorname{div}(f\nabla\tilde{v}) & \text{on} \quad \mathbb{R}^m \times (0,T), \\ \tilde{v} & = & 0 & \text{on} \quad \mathbb{R}^m \times \{0\}. \end{array} \tag{3}$$

We note that the nonlinear part of (3) is the second term in the right hand side, which can be dropped out as it is small when compared with the first term; see Elayyan and Isakov [2]. Thus we have, to the first order approximation,

$$\begin{array}{rcll} \tilde{v}_t - \Delta \tilde{v} & = & \operatorname{div}(f \nabla \tilde{u}) & \text{on} \quad \mathbb{R}^m \times (0,T), \\ \tilde{v} & = & 0 & \text{on} \quad \mathbb{R}^m \times \{0\}. \end{array} \tag{4}$$

Representing the solution $\tilde{u}$ in (2) in integral form, we can express $\tilde{v}$ in (4) as

$$\tilde{v}(x^*, t) = -\frac{1}{4(4\pi)^m} \int_0^t \int_{\mathbb{R}^m} f(x) \frac{|x^* - x|^2}{(\tau(t-\tau))^{m/2+1}} \exp\left(-\frac{|x^* - x|^2 t}{4\tau(t-\tau)} \right) dx d\tau. \tag{5}$$

Since $u(x, t; x^*)$ is given in $\Omega^* \times (0,T)$ and $\tilde{u}$ in (2) can be solved exactly, the left hand side $\tilde{v} = u - \tilde{u}$ in (5) is known in $\Omega^* \times (0,T)$. Since f is zero outside Ω, we can rewrite (5) as

$$-\frac{1}{4(4\pi)^m} \int_0^t \int_{\Omega} f(x) \frac{|x^* - x|^2}{(\tau(t-\tau))^{m/2+1}} \exp\left(-\frac{|x^* - x|^2 t}{4\tau(t-\tau)} \right) dx d\tau = \tilde{v}(x^*, t), \tag{6}$$

where the right hand side $\tilde{v}(x^*, t)$ is known in $\Omega^* \times (0,T)$.

For simplicity, we set $t = T$. Then the integral equation (6) can be written as

$$\mathcal{K} f = b, \tag{7}$$

with $f \in L^2(\Omega)$, $b \in L^2(\Omega^*)$, and

$$(\mathcal{K} f)(y) = \int_{\Omega} k(y - x) f(x) dx.$$

Here the convolution kernel function $k(x)$ is given by

$$k(x) = -\frac{1}{4(4\pi)^m} \int_0^T \frac{|x|^2}{(\tau(T-\tau))^{m/2+1}} \exp\left(-\frac{|x|^2 T}{4\tau(T-\tau)} \right) d\tau. \tag{8}$$

Thus $\mathcal{K}$ is an operator from $L^2(\Omega)$ into $L^2(\Omega^*)$. The uniqueness of the solution f in (7) has been shown in [2].

Equation (7) is a Fredholm integral equation of the first kind. Since $\mathcal{K}$ maps any Sobolev space on Ω into the space of functions that are analytic in a neighborhood of Ω^*, (7) is ill-posed, see [2]. Regularization method is therefore needed to solve (7). In the next two sections, we will study the Tikhonov regularization method for this problem.

3 Regularizations by the Identity Operator

In the Tikhonov approach for solving the ill-posed equation (7), the equation is replaced by the regularized equation

$$(\alpha \mathcal{Q} + \mathcal{K}^* \mathcal{K}) f = \mathcal{K}^* b, \tag{9}$$

where α is the regularization parameter and $\mathcal{Q}$ is the regularization operator. In [2], $\mathcal{Q}$ is chosen to be the identity operator $\mathcal{I}$.

Since $\mathcal{K}$ is a convolution kernel, the discretized matrix $\mathbf{K}$ of $\mathcal{K}$ by the rectangular quadrature rule is a Toeplitz matrix for 1-dimensional problem and a block-Toeplitz-Toeplitz-block matrix for 2-dimensional problem, see for instance Chan and Ng [1]. Thus the matrix-vector multiplication of $\mathbf{K}$ with any vector is of order $O(n \log n)$ operations, where n is the size of the matrix $\mathbf{K}$, see for instance Chan and Ng [1]. Hence if the conjugate gradient method is applied to solving the discrete equation of (9), the cost per iteration is of order $O(n \log n)$ operations, see for instance Stoer and Bulirsch [4, p. 606]. As for the convergence rate, we have

Theorem 1 *The spectrum of the regularized operator $\alpha\mathcal{I} + \mathcal{K}^*\mathcal{K}$ is clustered around α. In particular, if the conjugate gradient method is applied to solving the regularized system, then the convergence rate will be superlinear.*

Proof: We first note that any operator $\mathcal{K}$ with kernel function $k \in L^2(\Omega \times \Omega^*)$ is a compact operator from $L^2(\Omega)$ to $L^2(\Omega^*)$, see for instance Hackbusch [3, Theorem 3.2.7]. Since the kernel function $k(x)$ in (8) is smooth, $\mathcal{K}$ is a compact operator from $L^2(\Omega)$ to $L^2(\Omega^*)$. Hence $\mathcal{K}^*\mathcal{K}$ is also compact from $L^2(\Omega)$ to $L^2(\Omega)$. Therefore its spectrum is clustered around 0. In particular, the spectrum of the regularized operator $\alpha\mathcal{I} + \mathcal{K}^*\mathcal{K}$ is clustered around α. Thus if the conjugate gradient method is applied to solving the regularized system, then the convergence rate will be superlinear. $\square$

Thus the conjugate gradient method for solving (7) will converge in a fixed number of iterations independent of n, the size of the discretization matrix. Since the cost per iteration of the conjugate gradient method is of order $O(n \log n)$ operations, the total cost of solving (7) is also of order $O(n \log n)$ operations.

4 Regularization by the Laplacian Operator

Although the identity-regularized equation converges superlinearly, we find that the regularized solution so obtained is not good in general. In the following, we propose to use the Laplacian operator as the regularization operator, i.e., the regularization equation is given by

$$(\alpha\Delta + \mathcal{K}^*\mathcal{K})f = \mathcal{K}^*b. \tag{10}$$

We now illustrate the differences between the two regularization operators by examples. In the computation, the discretization matrix $\mathbf{K}$ is obtained by applying the rectangle rule to $\mathcal{K}$. The regularized equations are solved by the conjugate gradient method with stopping criteria $\|\mathbf{r}_q\|_\infty / \|\mathbf{r}_0\|_\infty < 10^{-7}$, where $\mathbf{r}_q$ is the residual after the qth iteration. The initial guess is chosen to be the zero vector. All computations were done by Matlab on an IBM 43P-133 workstation. The best regularization parameter α were obtained experimentally by trying different values of α. We found that it is rather robust – change in the first decimal place usually gives the same result. We note that since the entries of $\mathbf{K}$ are very small (of magnitude smaller than 10^{-4} and 10^{-7} respectively for the 1-dimensional and 2-dimensional problems we tried), α thus computed are small too.

We start with two one-dimensional problems with $\Omega = [0, 2]$, $\Omega^* = [3, 5]$, $T = 4$ and $n = 256$. The true solutions $f(x)$ are chosen to be $\sin(\pi x/2)$ and $x^3 + 3x^2 + 2x$. The right hand side vector $\mathbf{b} = \mathbf{K}\mathbf{f}$ of the discrete equation is obtained accordingly. Then 10% random noise vector $\mathbf{n}$ is added to the right hand side $\mathbf{b}$. More precisely, $\|\mathbf{n}\|_\infty / \|\mathbf{b}\|_\infty = 0.1$ and we solve

$$(\alpha\mathbf{Q} + \mathbf{K}^*\mathbf{K})\mathbf{f} = \mathbf{K}^*(\mathbf{b} + \mathbf{n}), \tag{11}$$

where $\mathbf{Q}$ is either the identity matrix or the discrete Laplacian.

In Figure 1, the solid line, the dash line and the dot line show the original image, the result from the identity regularization and the result from Laplacian regularization respectively. In the first example, the value of α for both regularization operators are 10^{-7}. In the second example, $\alpha = 10^{-9}$ for the identity regularization and $\alpha = 10^{-10}$ for Laplacian regularization.

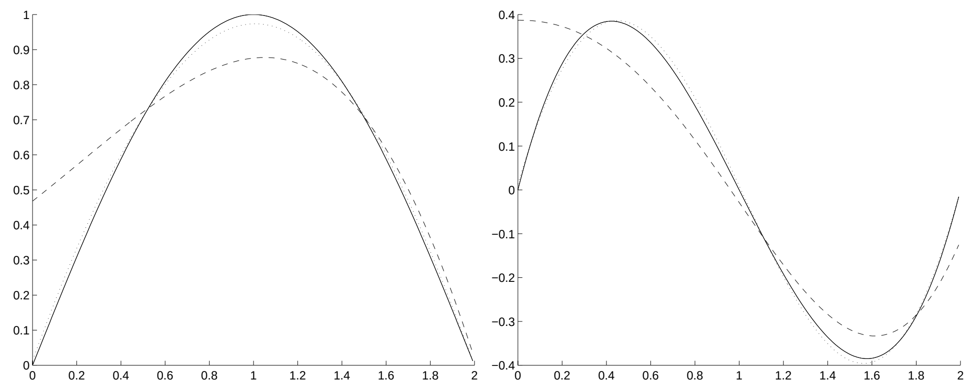

Figure 1. (left) $f(x) = \sin(\pi x/2)$, (right) $f(x) = x^3 - 3x^2 + 2x$

Next we consider the two-dimensional case. We choose $\Omega = [0,2]^2$, $\Omega^* = [3,5]^2$, $T = 4$ and $n = 64^2$. The test functions are $f(x,y) = \sin(\pi x)\sin(\pi y)$ and $f(x,y) = (x^3 - 3x^2 + 2x)(e^{2y^3-4y^2} - 1)$. Ten percent relative noise vector $\mathbf{n}$ is added to the right hand side vector $\mathbf{b}$ as in the 1-dimensional case, see (11). Figure 2 shows the original images for both test functions. Figures 3 and 4 are the images recovered by the identity regularization method and the Laplacian regularization method respectively.

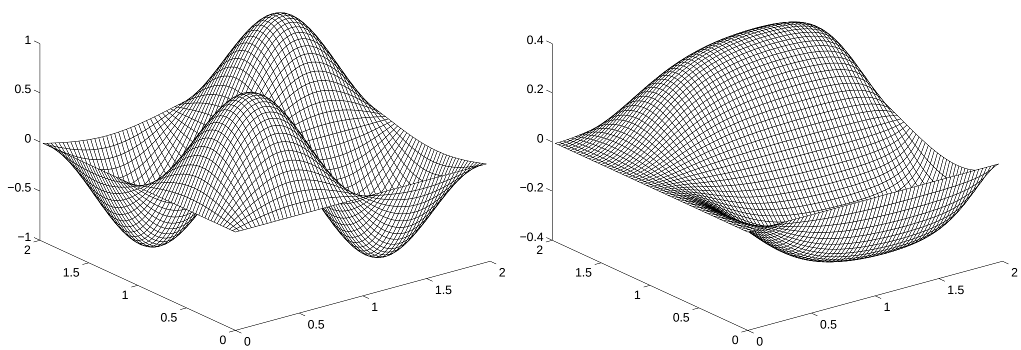

Figure 2. Original Image: (left) $\sin(\pi x)\sin(\pi y)$, (right) $(x^3 - 3x^2 + 2x)(e^{2y^3-4y^2} - 1)$

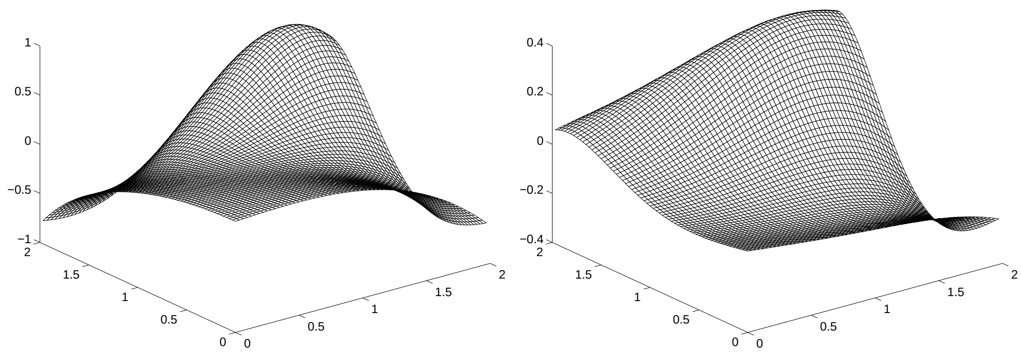

Figure 3. Identity Regularization with (left) $\alpha = 5 \cdot 10^{-14}$, (right) $\alpha = 10^{-13}$

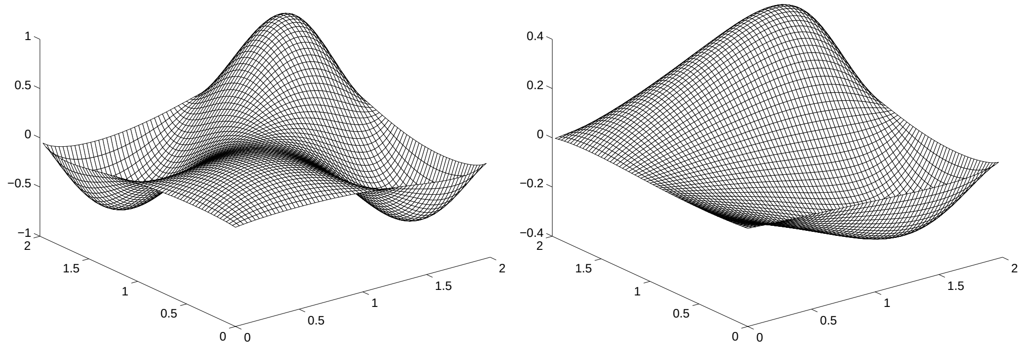

Figure 4. Laplacian Regularization with (left) $\alpha = 3 \cdot 10^{-15}$, (right) $\alpha = 10^{-13}$

From the examples, we see that the solutions obtained by the Laplacian regularization are much better than those from the identity regularization. Although we have shown only two examples here, we emphasize that we have similar results for other examples we tested.

5 Preconditioning with the Laplacian

In this section, we consider solving (10) by the conjugate gradient method. As mentioned previously, the cost of multiplying $\mathbf{K}$ to any vector is of order $O(n \log n)$ operations where n is the size of the matrix. Since the discrete Laplacian is a band-matrix with either 3 or 5 non-zero bands, the cost of multiplying it to any vector is of $O(n)$ operations. Hence if the conjugate gradient method is applied to solving the discrete equation of (10), the cost per iteration is of order $O(n \log n)$ operations.

However, since the Laplacian operator is ill-conditioned, we expect the Laplacian regularized equation (10) to be ill-conditioned. In particular, the number of iterations required for convergence will increase with increasing matrix size, see Tables 1 and 2 below for instance. Here we propose to use $\alpha\Delta$ as preconditioner for the equation and prove that the resulting preconditioned equation is well-conditioned. We note that for all 1-dimensional problems or 2-dimensional problems on rectangular domains, the inverse of the Laplacian operator can be obtained quite efficiently by using fast sine-transforms in $O(n \log n)$ operations, see for instance Stoer and Bulirsch [4, p.592]. Thus the cost per iteration of the preconditioned conjugate gradient method is still of the order $O(n \log n)$ operations. As for the convergence rate, we have

Theorem 2 *The spectrum of the preconditioned operator $\mathcal{I} + \alpha^{-1}\Delta^{-1}\mathcal{K}^*\mathcal{K}$ is clustered around* 1. *In particular, if the conjugate gradient method is applied to solving the preconditioned system, then the convergence rate will be superlinear.*

Proof: From the proof of Theorem 1, we already know that the operator $\mathcal{K}^*\mathcal{K}$ is compact from $L^2(\Omega)$ to $L^2(\Omega)$. Since Δ^{-1} is a bounded operator on $L^2(\Omega)$ (see for instance, Gilbarg and Trudinger [5, p.101]), $\Delta^{-1}\mathcal{K}^*\mathcal{K}$ is still a compact operator from $L^2(\Omega)$ to $L^2(\Omega)$. Therefore its spectrum is clustered around 0. In particular, the spectrum of the regularized operator $\mathcal{I} + \alpha^{-1}\Delta^{-1}\mathcal{K}^*\mathcal{K}$ is clustered around 1. Thus if the conjugate gradient method is applied to solving the regularized system, then the convergence rate will be superlinear. $\square$

Thus the conjugate gradient method for solving the preconditioned system will converge in a fixed number of iterations independent of n. Since the cost per iteration of the preconditioned conjugate gradient method is of order $O(n \log n)$ operations, the total cost of solving (10) is also of order $O(n \log n)$ operations.

In the following, we illustrate the fast convergence of our preconditioned systems by using examples considered in §4. We apply the conjugate gradient method to (10) with or without the preconditioners. The numbers of iterations required for convergence are listed in Tables 1 and 2, where "No" and "$\alpha\Delta$" indicate whether preconditioning is used. The regularization parameter α is as given in §4 and is not changed with n. (We find from our numerical results that the best α is basically unchanged w.r.t n.) The stopping tolerance is again set to $\|\mathbf{r}_q\|_\infty / \|\mathbf{r}_0\|_\infty < 10^{-7}$. The initial guess is chosen to be the zero vector. All computations were done by Matlab on an IBM 43P-133 workstation. We note that the number of iterations is roughly constant if $\alpha\Delta$ is used as the preconditioner whereas the number is increasing for the non-preconditioned system.

$f(x)$	$\sin(\pi x/2)$		$x^3 - 3x^2 + 2x$	
n	No	$\alpha\Delta$	No	$\alpha\Delta$
32	33	3	28	5
64	64	3	63	5
128	128	3	116	5
256	256	3	230	5
512	512	3	441	5
1024	1024	3	849	5

Table 1. Numbers of the Iterations for 1-D Examples.

$f(x,y)$	$\sin(\pi x)\sin(\pi y)$		$(x^3 - 3x^2 + 2x)(e^{2y^3-4y^2} - 1)$	
n	No	$\alpha\Delta$	No	$\alpha\Delta$
8^2	18	12	16	10
16^2	31	10	29	10
32^2	58	13	52	10
64^2	92	10	95	13
128^2	185	10	195	10

Table 2. Numbers of the Iterations for 2-D Examples.

6 Conclusing Remarks

In this preliminary report, we consider solving the integral equation (6) which is obtained by linearizing (1). We find that the Laplacian operator is a better regularization operator than the identity operator. To speed up the convergence rate of the regularized equations, we propose to use the Laplacian operator itself as preconditioner and prove that the resulting preconditioned systems have clustered spectra. We will compare our proposed operator with other regularization operators (such as the TV-norm operator) and methods in future work.

References

[1] R. Chan and M. Ng, *Conjugate Gradient Method for Toeplitz Systems*, SIAM Review, to appear.

[2] A. Elayyan and V. Isakov, *On Thermal Tomography*, SIAM J. of Appl. Math., to appear.

[3] W. Hackbusch, *Integral Equations*, ISNM Vol 120, Birkhäuser Verlag, 1995, Basel.

[4] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, 1991, Berlin.

[5] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd Ed., Springer-Verlag, 1983, Berlin.