Working Paper

Linear Convergence of Epsilon-Subgradient Descent Methods for a Class of Convex Functions

Stephen M. Robinson

WP-96-041
April 1996

IIASA
International Institute for Applied Systems Analysis ▫ A-2361 Laxenburg ▫ Austria
Telephone: +43 2236 807 ▫ Fax: +43 2236 71313 ▫ E-Mail: info@iiasa.ac.at

Linear Convergence of Epsilon-Subgradient Descent Methods for a Class of Convex Functions

Stephen M. Robinson

WP-96-041
April 1996

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

Abstract

This paper establishes a linear convergence rate for a class of epsilon-subgradient descent methods for minimizing certain convex functions on $\boldsymbol{R}^n$. Currently prominent methods belonging to this class include the resolvent (proximal point) method and the bundle method in proximal form (considered as a sequence of serious steps). Other methods, such as the recently proposed descent proximal level method, may also fit this framework depending on implementation. The convex functions covered by the analysis are those whose conjugates have subdifferentials that are locally upper Lipschitzian at the origin, a class introduced by Zhang and Treiman. We argue that this class is a natural candidate for study in connection with minimization algorithms.

Linear Convergence of Epsilon-Subgradient Descent Methods for a Class of Convex Functions*

*Stephen M. Robinson***

1 Introduction

This paper deals with ϵ-subgradient-descent methods for minimizing a convex function f on $\boldsymbol{R}^n$. The class of methods we consider consists of those treated by Correa and Lemaréchal in [3], with the additional restrictions that the minimizing set be nonempty, the stepsize parameters be bounded, and a condition for sufficient descent be enforced at each step. We give a precise description of this class in Section 2.

Currently prominent methods belonging to this class include the resolvent (proximal point) method and the bundle method in proximal form (considered as a sequence of serious steps). The resolvent method was treated by Rockafellar [12, 13] and has since been the subject of much attention. Implementations of the proximal bundle method have been given recently by Zowe [16], Kiwiel [7], and Schramm and Zowe [14], building on a considerable amount of earlier work; see [6] for references. Certain other methods, such as the recently proposed descent proximal level method of Brännlund, Kiwiel, and Lindberg [1], may fit into the class we consider depending on how they are implemented.

We show that the methods we consider will converge with (at least) an R-linear rate in in the sense of Ortega and Rheinboldt [8], in the case when they are used to minimize closed proper convex functions f on $\boldsymbol{R}^n$ that are of a special type: namely, those whose conjugates f^* have subdifferentials that are *locally upper Lipschitzian* at the origin. This means that there exist a neighborhood U of the origin in $\boldsymbol{R}^n$ and a constant μ such that for each $x^* \in U$,

$$\partial f^*(x^*) \subset \partial f^*(0) + \mu \|x^*\| B,$$

where B is the (Euclidean) unit ball. The local upper Lipschitzian property was introducedin [9]; the class of functions whose conjugates have subdifferentials obeying this

*This material is based upon work supported by the U. S. Army Research Office under Grant DAAH04-95-1-0149. Preliminary research for this paper was conducted in part at the Institute for Mathematics and its Applications, Minneapolis, Minnesota, with funds provided by the National Science Foundation, and in part at the Project on Optimization Under Uncertainty, International Institute for Applied Systems Analysis, Laxenburg, Austria.

*Department of Industrial Engineering, University of Wisconsin–Madison, 1513 University Avenue, Madison, WI 53706-1539. Email `smr@cs.wisc.edu`; Fax 608-262-8454; Phone 608-263-6862.

property at the origin has been studied by Zhang and Treiman [15], and we shall call them *ZT-regular with modulus* μ. For the problem of unconstrained minimization of a C^2 function, the standard second-order sufficient condition (that is, positive definiteness of the Hessian at a minimizer) implies that the function is convex if restricted to a suitable neighborhood of the minimizer, that the conjugate this restricted function is finite near the origin, and that ZT-regularity holds. The ZT-regularity condition is therefore a natural candidate for study in connection with minimization algorithms.

The rest of this paper is organized in two sections. Section 2 describes precisely the class of minimization methods we consider, and provides some useful information about their behavior, including convergence. Section 3 then shows that their rate of convergence is at least R-linear if the function being minimized is ZT-regular.

2 Subgradient-descent methods

In this section we describe the class of minimization methods with which we are concerned, and we review some results about their behavior.

Let f be a closed proper convex function on $\boldsymbol{R}^n$, which we wish to minimize. The authors of [3] investigated a class of ϵ-subgradient descent methods for such minimization. These methods proceed by fixing a starting point $x_0 \in \boldsymbol{R}^n$ and then generating succeeding points by the formula

$$x_{n+1} = x_n - t_n d_n^*, \tag{1}$$

where t_n is a positive stepsize parameter and for some nonnegative ϵ_n, d_n^* belongs to the ϵ_n-subdifferential $\partial_{\epsilon_n} f(x_n)$ of f at x_n, defined by

$$\partial_{\epsilon_n} f(x_n) = \{x^* \mid \text{ for each } z \in \boldsymbol{R}^n,\ f(z) \geq f(x_n) + \langle x^*, z - x_n \rangle - \epsilon_n \}.$$

Thus, for $\epsilon_n = 0$ we have the ordinary subdifferential, whereas for positive ϵ_n we have a larger set. For more information about the ϵ-subdifferential, see [10].

In addition to requiring the function f to satisfy certain properties, we shall impose two requirements on the implementation of (1). They are stricter than those imposed in [3], but they will permit us to obtain the convergence rate results that we are after. One of these is that the sequence of stepsize parameters be bounded away from 0 and from ∞: namely, there are t_* and t^* such that for each n,

$$0 < t_* \leq t_n \leq t^*. \tag{2}$$

The other requirement is that at each step a sufficient descent is obtained: specifically, there is a constant $m \in (0, 1]$ such that for each n,

$$f(x_{n+1}) \leq f(x_n) + m(\langle d_n^*, x_{n+1} - x_n \rangle - \epsilon_n). \tag{3}$$

Note that because $d_n^* = -t_n^{-1}(x_{n+1} - x_n)$, the quantity in parentheses in (3) is nonpositive, and in fact negative if $x_{n+1} \neq x_n$ or if $\epsilon_n > 0$, so that we are working with a descent method: that is, one that forces the function value at each successive step to be "sufficiently" smaller than its predecessor. Indeed, if $\epsilon_n = 0$ and if the subgradient is actually a gradient, this is a descent condition very familiar from the literature (for example, see ([4], p. 101). However, the ϵ-descent condition in the general form given here may seem somewhat strange. For that reason, we next show that this condition is satisfied by the two known methods mentioned earlier.

The first of these methods is the resolvent, or proximal point, method in the form appropriate for minimization of f. This algorithm is specified by

$$x_{n+1} = (I + t_n \partial f)^{-1}(x_n);$$

that is, we obtain x_{n+1} by applying to x_n the resolvent J_{t_n} of the maximal monotone operator ∂f. To see that this is in the form (1), note that the algorithm specification implies that there is $d_n^* \in \partial f(x_{n+1})$ such that

$$x_n = x_{n+1} + t_n d_n^*,$$

which is a rearrangement of (1). Further, for each z we have

$$f(z) \geq f(x_{n+1}) + \langle d_n^*, z - x_{n+1} \rangle = f(x_n) + \langle d_n^*, z - x_n \rangle - \epsilon_n,$$

where

$$\epsilon_n = f(x_n) - f(x_{n+1}) - \langle d_n^*, x_n - x_{n+1} \rangle,$$

which is nonnegative because $d_n^* \in \partial f(x_{n+1})$. Therefore $d_n^* \in \partial_{\epsilon_n} f(x_n)$. Moreover, we have

$$f(x_{n+1}) = f(x_n) + \langle d_n^*, x_{n+1} - x_n \rangle - \epsilon_n,$$

so that (3) holds with $m = 1$.

The resolvent method is unfortunately not implementable except in special cases. For practical minimization of nonsmooth convex functions a very effective tool is the well known bundle method, which as is pointed out in [3] can be regarded as a systematic way of approximating the iterations of the resolvent method. The method uses two kinds of steps: "serious steps," which as we shall see correspond to (1), and "null steps," which are used to prepare for the serious steps. Specifically, by means of a sequence of null steps the method builds up a piecewise affine minorant $\hat{f}$ of f. Then a resolvent step is taken, using $\hat{f}$ instead of f:

$$x_{n+1} = (I + t_n \partial \hat{f})^{-1}(x_n), \tag{4}$$

and it is accepted if

$$f(x_n) - f(x_{n+1}) \geq m[f(x_n) - \hat{f}(x_{n+1})]. \tag{5}$$

Now from (4) we see that

$$x_{n+1} = x_n - t_n d_n^*,$$

with $d_n^* \in \partial \hat{f}(x_{n+1})$. Then for each $z \in \boldsymbol{R}^n$ we have

$$f(z) \geq \hat{f}(z) \geq \hat{f}(x_{n+1}) + \langle d_n^*, z - x_{n+1} \rangle = f(x_n) + \langle d_n^*, z - x_n \rangle - \epsilon_n,$$

where we can write ϵ_n as

$$\epsilon_n = [f(x_n) - \hat{f}(x_n)] + [\hat{f}(x_n) - \hat{f}(x_{n+1}) - \langle d_n^*, x_n - x_{n+1} \rangle], \tag{6}$$

which must be nonnegative since $\hat{f}$ minorizes f and $d_n^* \in \partial \hat{f}(x_{n+1})$. In fact, $\hat{f}$ is typically constructed in such a way that $\hat{f}(x_n) = f(x_n)$, so the first term in square brackets is actually zero (this will be the case as long as a subgradient of f at x_n belongs to the bundle). In that case we have from the minorization property and (6)

$$f(x_n) - \hat{f}(x_{n+1}) \geq \hat{f}(x_n) - \hat{f}(x_{n+1}) = \langle d_n^*, x_n - x_{n+1} \rangle + \epsilon_n,$$

so that (5) yields

$$f(x_n) - f(x_{n+1}) \geq m[\langle d_n^*, x_n - x_{n+1} \rangle + \epsilon_n];$$

that is, (3) holds. Therefore the bundle method, if implemented with bounded t_n, fits within our class of methods.

Although our proof of R-linear convergence in Section 3 therefore applies to the bundle method, it must be noted that this analysis takes into account only the serious steps, whereas for each serious step a possibly large number of null steps may be required to build up an adequate approximation $\hat{f}$. Therefore our analysis does not provide a bound on the total work required to implement the bundle method.

We have therefore seen that two well known methods fit into the class we shall analyze. In the analysis we shall need the following theorem, which summarizes the convergence properties of this class.

Theorem 1 *Let f be a lower semicontinuous proper convex function on $\boldsymbol{R}^n$, having a nonempty minimizing set X_*. Let x_0 be given and suppose the algorithm (1) is implemented in such a way that (2) and (3) hold. Then the sequence $\{x_n\}$ generated by (1) converges to a point $x_* \in X_*$, $\{f(x_n)\}$ converges to $\min f$, and*

$$\sum_{n=0}^{\infty} (\|d_n^*\|^2 + \epsilon_n) < \infty. \tag{7}$$

In particular, the sequences $\{\epsilon_n\}$ and $\{\|d_n^*\|\}$ converge to zero.

Proof. Note that for each n we have $\langle d_n^*, x_{n+1} - x_n \rangle = -t_n \|d_n^*\|^2$. From (2) and (3) we obtain

$$m(t_* \|d_n^*\|^2 + \epsilon_n) \leq m(t_n \|d_n^*\|^2 + \epsilon_n) \leq f(x_n) - f(x_{n+1}),$$

so for each $k \geq 1$ we have

$$m \sum_{n=0}^{k-1} (t_* \|d_n^*\|^2 + \epsilon_n) \leq f(x_0) - f(x_k) \leq f(x_0) - \min f,$$

and consequently

$$m \sum_{n=0}^{\infty} (t_* \|d_n^*\|^2 + \epsilon_n) \leq f(x_0) - \min f,$$

which establishes (7). The condition (2) shows that the sum of the t_n is infinite, so that Conditions (1.4) and (1.5) of [3] hold. Moreover, (3) shows that for each n

$$f(x_{n+1}) \leq f(x_n) + m(\langle d_n^*, x_{n+1} - x_n \rangle - \epsilon_n) \leq f(x_n) - m t_n \|d_n^*\|^2,$$

so that Condition (2.7) of [3] also holds. Then Proposition 2.2 of [3] shows that $\{f(x_n)\}$ converges to $\min f$ and that $\{x_n\}$ converges to some element x_* of X_*. $\square$

In this section we have specified the class of methods we are considering, and we have given two examples of concrete methods that belong to this class. Moreover, we have adapted from [3] a general convergence result applicable to this class. In the next section we present the main result of the paper, a proof that the convergence guaranteed by Theorem 1 will under additional conditions actually be at least R-linear.

3 Convergence-rate analysis

In order to prove the main result we need to use a tailored form of the well known Brøndsted-Rockafellar Theorem [2]. We give this next, along with a very simple proof. The technique of this proof is very similar to that given in Theorem 4.2.1 of [5], but this version gives slightly more information and it holds in any real Hilbert space.

Theorem 2 *Let H be a real Hilbert space and let f be a lower semicontinuous proper convex function on H. Suppose that $\epsilon \geq 0$ and that $(x_\epsilon, x_\epsilon^*) \in \partial_\epsilon f$. For each positive β there is a unique y_β with*

$$(x_\epsilon + \beta y_\beta, x_\epsilon^* - \beta^{-1} y_\beta) \in \partial f. \tag{8}$$

Further, $\|y_\beta\| \leq \epsilon^{1/2}$.

Proof. Define a function g on H by

$$g(y) = (1/2)\|y - \beta x_\epsilon^*\|^2 + f(x_\epsilon + \beta y).$$

Then g is lower semicontinuous, proper, and strongly convex; its unique minimizer y_β then satisfies $0 \in \partial g(y_\beta)$, which upon rearrangement becomes (8); justification for the subdifferential computation can be found in, e.g., Theorem 20, p. 56, of [11]. In turn, (8) implies

$$f(x_\epsilon) \geq f(x_\epsilon + \beta y_\beta) + \langle x_\epsilon^* - \beta^{-1} y_\beta, x_\epsilon - (x_\epsilon + \beta y_\beta)\rangle.$$

But the ϵ-subgradient inequality yields

$$f(x_\epsilon + \beta y_\beta) \geq f(x_\epsilon) + \langle x_\epsilon^*, (x_\epsilon + \beta y_\beta) - x_\epsilon\rangle - \epsilon,$$

and by combining these we obtain

$$0 \geq \langle x_\epsilon^* - \beta^{-1} y_\beta, -\beta y_\beta\rangle + \langle x_\epsilon^*, \beta y_\beta\rangle - \epsilon = \|y_\beta\|^2 - \epsilon,$$

which proves the assertion about $\|y_\beta\|$. $\square$

Here is the main theorem, which says that under ZT-regularity and some implementation conditions the ϵ-subgradient descent method is at least R-linearly convergent.

Theorem 3 *Let f be a lower semicontinuous, proper convex function on $\mathbf{R}^n$ that is ZT-regular with modulus $\mu > 0$. Assume that f has a nonempty minimizing set X_*, and that starting from some x_0 the ϵ-subgradient descent method (1) is implemented with (2) and (3) satisfied at each step.*

Then the sequence $\{x_n\}$ produced by (1) converges at least R-linearly to a limit $x_* \in X_*$.

Proof. Consider the step from x_n to x_{n+1}. From (3) we find that $d_n^* \in \partial_{\epsilon_n} f(x_n)$, and by applying Theorem 2 we conclude that there is a unique y with $\|y\| \leq \epsilon_n^{1/2}$ and with

$$(x_n + \mu^{1/2}y, d_n^* - \mu^{-1/2}y) \in \partial f.$$

For any k let u_k be the projection of x_k on the optimal set X_*. We have shown in Theorem 1 that $\|d_n^*\|$ and ϵ_n converge to zero. Therefore there is some N such that for $n \geq N$ the point $d_n^* - \mu^{-1/2}y$ will lie in the neighborhood U associated with the ZT-regularity condition and, as a consequence, we shall have the inequality

$$\|(x_n + \mu^{1/2}y) - u_n\| \leq \mu\|d_n^* - \mu^{-1/2}y\|. \tag{9}$$

Therefore

$$\begin{array}{rcl} \|x_n - u_n\| & \leq & \|(x_n + \mu^{1/2}y) - u_n\| + \mu^{1/2}\|y\| \\ & \leq & \mu\|d_n^* - \mu^{-1/2}y\| + \mu^{1/2}\epsilon_n^{1/2} \\ & \leq & \mu\|d_n^*\| + 2\mu^{1/2}\epsilon_n^{1/2}. \end{array} \tag{10}$$

Next, let $f_* = \min f$; write ϕ_n for $f(x_n) - f_* = f(x_n) - f(u_n)$, and σ_n for μt_n^{-1}. Note that for any real numbers α, β, and γ we have, by applying the Schwarz inequality to $(1, \beta)$ and (α, γ),

$$|\alpha + \beta\gamma| \leq (1 + \beta^2)^{1/2}(\alpha^2 + \gamma^2)^{1/2}. \tag{11}$$

Using (9), (10), and the fact that $d_n^* \in \partial_{\epsilon_n} f(x_n)$ we obtain

$$\begin{array}{rcl} \phi_n & \leq & -\langle d_n^*, u_n - x_n\rangle + \epsilon_n \\ & \leq & \mu\|d_n^*\|^2 + 2\mu^{1/2}\|d_n^*\|\epsilon_n^{1/2} + \epsilon_n \\ & = & (\mu^{1/2}\|d_n^*\| + \epsilon_n^{1/2})^2 \\ & = & (\sigma_n^{1/2}t_n^{1/2}\|d_n^*\| + \epsilon_n^{1/2})^2 \\ & \leq & [(1 + \sigma_n)^{1/2}(t_n\|d_n^*\|^2 + \epsilon_n)^{1/2}]^2 \\ & = & (1 + \sigma_n)(t_n\|d_n^*\|^2 + \epsilon_n), \end{array} \tag{12}$$

where we used in succession the subgradient condition, the Schwarz inequality, and (11). But from (3) we have

$$t_n\|d_n^*\|^2 + \epsilon_n \leq m^{-1}[f(x_n) - f(x_{n+1})],$$

and we also have $f(x_n) - f(x_{n+1}) = \phi_n - \phi_{n+1}$. Therefore (12) yields

$$\phi_n \leq (1 + \sigma_n)m^{-1}(\phi_n - \phi_{n+1}),$$

which, since $t_n \geq t_* > 0$, implies

$$\phi_{n+1} \leq \theta^2\phi_n,$$

with

$$\theta = [1 - m/(1 + \mu t_*^{-1})]^{1/2}.$$

Therefore for fixed N and $n \geq N$ we have

$$\phi_n \leq \kappa\theta^{2n}, \tag{13}$$

with

$$\kappa = \theta^{-2N}\phi_N.$$

Now from Theorem 4.3 of [15] we find that for some $\gamma \geq 0$ and all z with $d(z, X_*)$ sufficiently small the inequality

$$f(z) \geq f_* + \gamma d(z, X_*)^2 \tag{14}$$

holds. We know that $d(x_n, X_*)$ converges to zero, so for all n at least as large as some $N' \geq N$ we have from (14)

$$\xi_n := d(x_n, X_*) \leq \gamma^{-1/2}\phi_n^{1/2} \leq \lambda\theta^n, \tag{15}$$

with

$$\lambda = \gamma^{-1/2}\theta^{-N}\phi_N^{1/2}.$$

Now let $e_n := \|x_n - x_*\|$, where x_* is the unique limit of the sequence $\{x_n\}$, as established in Theorem 1. From Equation (1.3) of [3] we have, for any $y \in \boldsymbol{R}^n$,

$$\|x_{n+1} - y\|^2 \leq \|x_n - y\|^2 + t_n^2\|d_n^*\|^2 + 2t_n[f(y) - f(x_n) + \epsilon_n].$$

If we restrict our attention to points $y \in X_*$ we may simplify this to

$$\|x_{n+1} - y\|^2 \leq \|x_n - y\|^2 + 2t_n[t_n\|d_n^*\|^2 + \epsilon_n - \phi_n].$$

For $j > n \geq N'$ we then use the fact that $t_k \leq t^*$ for all k to obtain the upper bound

$$\|x_j - y\|^2 \leq \|x_n - y\|^2 + 2t^*(\sum_{k=n}^{j-1}[t_k\|d_k^*\|^2 + \epsilon_k] - \phi_n).$$

The condition (3) gives

$$f(x_{k+1}) \leq f(x_k) + m(\langle d_k^*, x_{k+1} - x_k\rangle - \epsilon_k) = f(x_k) - m[t_k\|d_k^*\|^2 + \epsilon_k],$$

from which we conclude that

$$\sum_{k=n}^{j-1}[t_k\|d_k\|^2 + \epsilon_k] \leq m^{-1}[f(x_n) - f(x_j)] \leq m^{-1}\phi_n.$$

Therefore

$$\|x_j - y\|^2 \leq \|x_n - y\|^2 + 2t^*(m^{-1} - 1)\phi_n,$$

and by taking the limit as $j \to \infty$ we find that

$$\|x_* - y\|^2 \leq \|x_n - y\|^2 + 2t^*(m^{-1} - 1)\phi_n.$$

Now set $y = u_n$ to obtain

$$\|x_* - u_n\|^2 \leq \xi_n^2 + 2t^*(m^{-1} - 1)\phi_n.$$

The bounds (13) and (15) now yield, for $n \geq N'$,

$$\|x_* - u_n\| \leq \tau\theta^n,$$

with

$$\tau = (\lambda^2 + 2t^*(m^{-1} - 1)\kappa)^{1/2}.$$

Then we have

$$\|x_n - x_*\| \leq \xi_n + \|x_* - u_n\| \leq (\lambda + \tau)\theta^n,$$

so that $\{x_n\}$ converges at least R-linearly to the limit x_*, as claimed. □

References

[1] U. Brännlund, K. C. Kiwiel, and P. O. Lindberg, "A descent proximal level bundle method for convex nondifferentiable optimization," Preprint, April 1994.

[2] A. Brøndsted and R. T. Rockafellar, "On the subdifferentiability of convex functions," *Proc. Amer. Math. Soc.* **16** (1965) 605–611.

[3] R. Correa and C. Lemaréchal, "Convergence of some algorithms for convex minimization," *Math. Programming* **62** (1993) 261–275.

[4] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic Press, London, 1981).

[5] J.–B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms II,* Grundlehren der mathematische Wissenschaften 306 (Springer-Verlag, Berlin, 1993).

[6] K. C. Kiwiel, *Methods of Descent for Nondifferentiable Optimization* (Lecture Notes in Mathematics No. 1133, Springer-Verlag, Berlin, 1985).

[7] K. C. Kiwiel, "Proximity control in bundle methods for convex nondifferentiable optimization," *Math. Programming* **46** (1990) 105–122.

[8] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).

[9] S. M. Robinson, "Generalized equations and their solutions, Part I: Basic theory," *Math. Programming Study* **10** (1979) 128–141.

[10] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).

[11] R. T. Rockafellar, *Conjugate Duality and Optimization,* CBMS Regional Conference Series in Applied Mathematics No. 16 (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1974).

[12] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM J. Control Opt.* **14** (1976) 877–898.

[13] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Math. Oper. Res.* **1** (1976) 97–116.

[14] H. Schramm and J. Zowe, "A version of the bundle idea for minimizing a nonsmooth function: Conceptual idea, convergence analysis, numerical results," *SIAM J. Optimization* **2** (1992) 121–152.

[15] R. Zhang and J. Treiman, "Upper-Lipschtiz multifunctions and inverse subdifferentials," *Nonlinear Analysis: Theory, Methods, and Applications* **24** (1995) 273–286.

[16] J. Zowe, "The BT-algorithm for minimizing a nonsmooth functional subject to linear constraints," in: F. H. Clarke *et al.,* eds., *Nonsmooth Optimization and Related Topics* (Plenum Publishing Corp., New York, 1989).