Factors affecting convergence in the design of diffractive optics by iterative vector-space methods

Henry Stark, Yongyi Yang, and Damla Gurkan

Department of Electrical and Computer Engineering, Illinois Institute of Technology, Chicago, Illinois 60601

Received February 27, 1998; revised manuscript received July 21, 1998; accepted September 10, 1998

Special-purpose diffractive optical devices are often designed by iterative methods without consideration of convergence properties and related factors that affect performance. We examine the properties of iterative algorithms in a vector-space setting and illustrate, with examples, differences in convergence performance based on starting point, sequential versus parallel projections, and intersecting versus nonintersecting sets.
 [S0740-3232(99)00201-X]
OCIS codes: 050.1970, 220.3620, 000.3870.

1. INTRODUCTION

The design of complex, i.e., multifunction, high-efficiency, diffractive optical elements (DOE's) has been greatly facilitated by computer-based algorithms that generate the complex transmittance from constraints imposed on the functionality of the DOE. For example, Wood *et al.*[1] used the computer-intensive algorithm of simulated annealing to optimize the diffraction integral that transforms the lens pixel pattern to fit the output pattern. Hatakoshi and Nakamura[2] successfully designed optical branching networks by computer control of the appropriate phase-modulating function. A common procedure is to iteratively apply Fourier- or Fresnel-domain constraints followed by space-domain constraints. Among the more sophisticated techniques along these directions is the work by Bernhardt *et al.*,[3] who used an iterative Fresnel-transform loop to design digital holograms in which coding, binarization, and phase adjustments are sequentially applied. The same group demonstrated the iterative design of pure-phase DOE's that are able to perform several optical functions.[4] Algorithms of the Gerchberg–Saxton[5] type are frequently used in optical design. For example, Kress and Lee[6] used such an algorithm for designing a phase-only computer-generated hologram for optical interconnects.

Although there exists a large literature on iterative design of DOE's by alternating constraints between the source plane and the field plane,[1–4,6–22] there have been correspondingly fewer works that view these iterative algorithms within a strict mathematical framework that considers convergence. Some exceptions to the rule are the work by Piestun and his co-workers[21,22] and Catino and his co-workers.[11–13] In particular, Piestun and Shamir,[21] to the best of the authors' knowledge, were the first to impose Fresnel field constraints within a vector-space setting to design optics for nondiffracting beams. In a vector-space setting, each iteration consists of a series of projections onto convex or nonconvex constraint sets. The difference in the convergence behavior of the algorithm for these two types of sets was considered by Levi and Stark.[23]

In this paper the authors attempt to demonstrate, with examples, the problems associated with convergence in iterative optical design by projections in a vector-space setting. In particular, we address the following: What is the difference between strong convergence, weak convergence, and summed-distance error (SDE) convergence, and when can the designer expect to see each? What are the hazards of imposing nonconvex constraints, especially when there are more than two such constraints? What modification in the projection algorithm will allow SDE convergence even when there are many nonconvex constraints? What is the influence, if any, of the starting point in the iteration? And finally, what happens when constraints are inconsistent? Are reasonable solutions still possible?

2. THEORY

It is well known that vector-space methods, sometimes called projection methods, are a class of (mostly) iterative algorithms in which the operators are projectors and the iterates exhibit convergence properties that depend on the number and topologies of the sets on which the projectors project. In a typical problem, it is desired to design a DOE that is subject to certain constraints in the space domain, the Fourier domain, or the Fresnel domain. Each constraint defines a set C_i whose elements all share that particular constraint. The correct design must meet all the constraints and hence be an element in the intersection of all the sets. To reach the intersection, one can use the sequential projection algorithm

$$t_{n+1}(\mathbf{x}) = P_1 P_2 \ldots P_M t_n(\mathbf{x}), \tag{1}$$

where P_i, $i = 1,\ldots,M$ is the operator that projects onto constraint set C_i, $t_n(\mathbf{x})$ is the nth estimate of the DOE phase profile, $\mathbf{x}$ is a position variable, and M is the number of constraints. Since Eq. (1) is usually realized on a

computer (a finite-state machine), the following remarks apply:

1. If all the C_i, $i = 1,\ldots,M$ are convex and have a nonempty intersection, $t_n(\mathbf{x})$ converges strongly to a point in the intersection $C_0 = \cap_{i=1}^{M} C_i$.

The reader will recall that a set C_i is convex if and only if the convex sum of any two of its elements is also in the set. Thus for any $0 \leqslant \mu \leqslant 1$ and any two vector elements $\boldsymbol{y}_1 \in C_i$, $\boldsymbol{y}_2 \in C_i$, the convex sum $\mu\boldsymbol{y}_1 + (1-\mu)\boldsymbol{y}_2$ will also be in the set. If a set does not obey this property, it is said to be nonconvex.

At this point we remind the reader that given a Hilbert space of engineering interest, e.g., the space of square-integrable functions, the space of square-summable sequences, or the Euclidean space R'', strong convergence of a vector sequence $\boldsymbol{x}_n$ to $\boldsymbol{x}$ implies that $\lim_{n\to\infty}\|\boldsymbol{x}_n - \boldsymbol{x}\| = 0$. Such convergence is also called convergence in the norm. In contrast, weak or inner product convergence written as $\boldsymbol{x}_n \overset{w}{\to} \boldsymbol{x}$ implies that for every z in the Hilbert space, $\lim_{n\to\infty}\langle \boldsymbol{x}_n, \boldsymbol{z}\rangle = \langle \boldsymbol{x}, \boldsymbol{z}\rangle$, where $\langle\cdot\rangle$ denotes the inner (or dot) product. In general, assuming a nonempty intersection and convex constraint sets, Eq. (1) converges weakly to a point in the intersection. A well-known result is that weak convergence in a finite-dimensional space (as one would encounter in a finite-state machine such as a digital computer) always implies strong convergence.[24]

The practical implication of remark 1 is that one can always expect a feasible solution from applying Eq. (1), i.e., one that obeys all the constraints if every constraint statement describes a convex set and the intersection of the sets is not empty. However, if the intersection is empty, no feasible solution exists, and the user is asking for a design that violates the laws of physics. A feasible solution is not necessarily an optimal solution. An optimal solution requires some criterion function that is extremized over all feasible solutions. Also, one can sometimes obtain a reasonable solution even if a feasible one does not exist. We shall return to this point later.

2. If at least one C_i is nonconvex and $M = 2$, then the $t_n(\mathbf{x})$ have the property that the summed distance from $t_n(\mathbf{x})$ to the constraint sets never increases. The summed distance of an element $\boldsymbol{y}$ from its constraint sets is defined as

$$J(\boldsymbol{y}) = \sum_{i=1}^{M} \|\boldsymbol{y} - P_i\boldsymbol{y}\|. \tag{2}$$

In the context of attempting a solution using projection methods, $J(\boldsymbol{y})$ in Eq. (2) is sometimes called the summed-distance error (SDE). For $M = 2$, Eq. (2) yields

$$J(\boldsymbol{y}) = \|\boldsymbol{y} - P_1\boldsymbol{y}\| + \|\boldsymbol{y} - P_2\boldsymbol{y}\|. \tag{3}$$

Consider now Eq. (1) for $M = 2$; SDE convergence means that the iterates in Eq. (1) satisfy

$$J(t_{n+1}) \leqslant J(t_n) \qquad \text{for } n = 1, 2,\ldots. \tag{4}$$

Thus SDE convergence means that the summed distance as defined in Eq. (2) always decreases (or at least never increases). Although this is clearly a desirable property, it does not imply that $J(t_n) \to 0$ with increasing n. Thus the introduction of nonconvex constraints such as magnitude constraints on Fresnel or Fourier fields, phase-only transmittance, binarization, and still others may not lead to a feasible solution, an observation made in numerical simulations.

The fact that SDE convergence applies only to $M = 2$ in the sequential projection algorithm in Eq. (1) was shown by Levi and Stark[23] in their studies of the phase-recovery problem. Indeed for $M > 2$, it is not difficult to construct examples where SDE convergence fails.

3. In some problems, $M > 2$ by necessity. For example, consider the design of a DOE that must satisfy (i) binary-phase only, (ii) constraint on DOE size, and (iii) a magnitude constraint on far-field performance. Constraints (i) and (iii) are nonconvex, and here $M = 3$. In this case, use of Eq. (1) may cause the iterates to diverge or possibly meander to some point that is neither an SDE limit point nor a feasible solution. Nevertheless by using a parallel projection algorithm, one can still obtain SDE convergence even when $M > 2$. The result is stated as a theorem.

Parallel generalized projection algorithm (PGPA) theorem. Let H be a Hilbert space with elements $\boldsymbol{x}$, $\boldsymbol{y}$, etc., with inner product $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ and norm $\|\boldsymbol{x}\|$. For every $\boldsymbol{x}_0 \in H$ and every choice of positive constants $w_1, w_2, \ldots, w_m$ such that $\Sigma_{i=1}^{m} w_i = 1$, the sequence $\{\boldsymbol{x}_n\}$ generated by

$$\boldsymbol{x}_{n+1} = \sum_{i=1}^{m} w_i P_i \boldsymbol{x}_n \tag{5}$$

will satisfy

$$J(\boldsymbol{x}_{n+1}) \leqslant J(\boldsymbol{x}_n),$$

where, in this case, the (parallel) SDE is given by

$$J(\boldsymbol{x}_n) = \left[\sum_{i=1}^{m} w_i \|P_i\boldsymbol{x}_n - \boldsymbol{x}_n\|^2\right]^{1/2}. \tag{6}$$

This remarkable theorem, henceforth called the PGPA theorem, first appeared in the open literature in a paper by Kotzer *et al.*[25] in their work on correlation-filter design. Its derivation is based on combining the ideas of product spaces[26,27] and the Levi–Stark theorem.[23] It was derived independently by Yang[24] without knowledge that Kotzer *et al.* had already published it earlier in 1995.

The PGPA theorem says that one can obtain SDE convergence for more than two nonconvex constraints, i.e., $M > 2$, provided that instead of the sequential algorithm in Eq. (1) we use the parallel algorithm of Eq. (5). Figure 1 shows the various possibilities in solving a problem by vector-space projections. Figure 2 illustrates the progress of the iterates toward a solution for the sequential and the parallel cases.

In Section 4 we attempt to illustrate some of the issues discussed above with examples involving both Fresnel- and Fourier-plane diffraction. Although Fourier-plane constraints are widely used, Fresnel-plane constraints appear less widely, and hence we review some of the mathematics associated with these constraints in the next section.

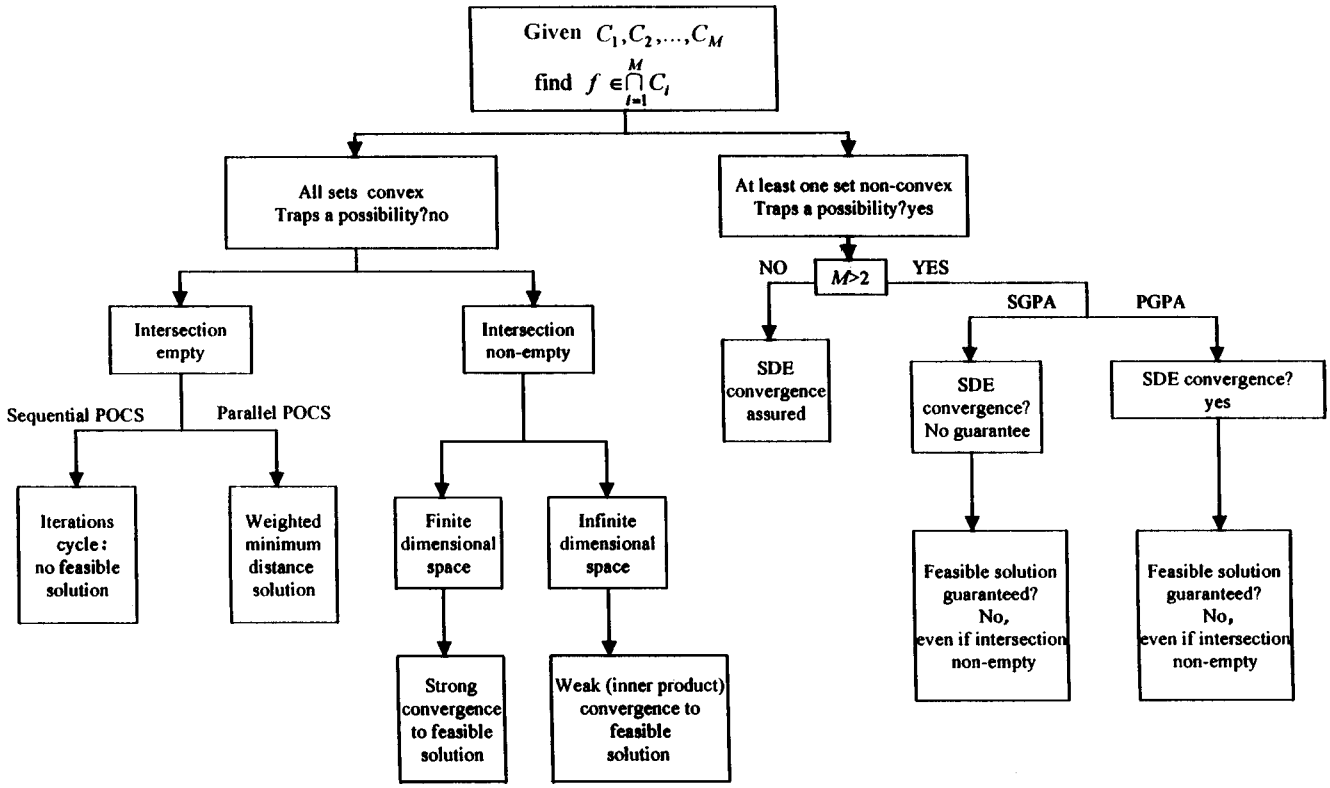


Fig. 1. Block diagram of various possibilities in solving a problem by vector-space projections. POCS stands for projections onto convex sets.

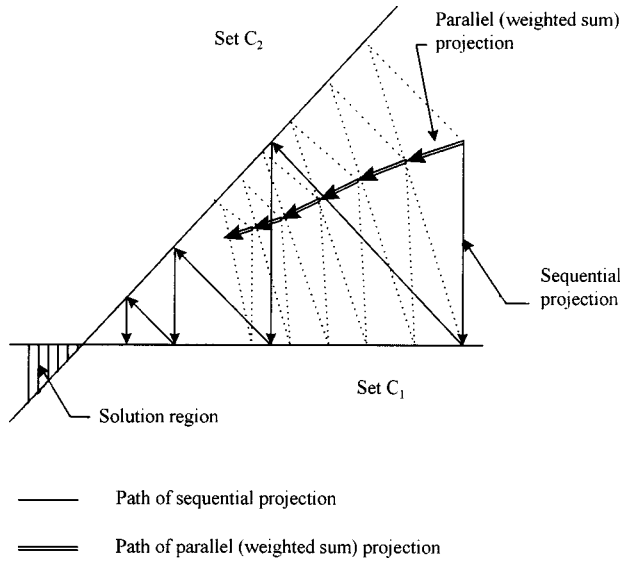


Fig. 2. Progress of the iterates toward a solution for the sequential and parallel cases.

3. FRESNEL-FIELD CONSTRAINTS

In what follows, we assume that scalar-diffraction theory is valid for the given geometries, that all elements are large compared with the wavelength of light, and that all media exhibit nonpolarizing, isotropic properties.

A diffractive element $t(x, y)$ illuminated by a unit-amplitude, quasi-monochromatic plane wave produces a field $g(\xi, \eta)$ in the (ξ, η) plane, a distance z away, according to the Fresnel diffraction law[28]

$$g_z(\xi, \eta) = \frac{1}{\lambda z} \int_R \int t(x, y) \times \exp\left\{ j \frac{\pi}{\lambda z} [(x - \xi)^2 + (y - \eta)^2] \right\} dx\,dy, \quad (7)$$

where λ is the wavelength of radiation, R is the set of points inside the aperture, and nonessential constants have been omitted for simplicity. For further simplicity, we assume a one-dimensional configuration. Then instead of Eq. (7) we write

$$g_z(\xi) = \frac{1}{\sqrt{\lambda z}} \int_{-a/2}^{a/2} t(x) \exp\left[j \frac{\pi}{\lambda z} (x - \xi)^2 \right] dx, \quad (8)$$

where a is the length of the aperture. Either Eq. (7) or Eq. (8) is sometimes called the Fresnel diffraction integral, and we shall write $g_z(\xi) = \mathrm{Fr}[t(x)]$, or simply $g_z = \mathrm{Fr}[t]$.

Constraint sets in the Fresnel case will obviously differ from those in the Fourier case. Nevertheless, Fresnel-field constraints can be cast as Fourier constraints by grouping the quadratic phase term with $t(x)$, a major advantage when one is using the fast-Fourier-transform algorithm on a computer.

In the following we consider a few examples of designing diffractive elements by using vector-space projections, with the aim of seeing how various factors affect convergence.

Example 1. Here we consider a Fresnel-field extension of the classic Fourier-plane/space-domain constraint, alternating orthogonal projection algorithm for designing DOE's that yield prescribed far-field intensities. The problem is stated as follows: Find the transmittance $t(x)$ of a phase-only diffractive element that produces a given field magnitude $M(\xi)$ in the Fresnel region a distance z away, assuming monochromatic, plane-wave illumination. In mathematical terms we seek a $t(x)$ such that

$$\left| \frac{1}{\sqrt{\lambda z}} \int_{-a/2}^{a/2} t(x) \exp\left[j \frac{\pi}{\lambda z} (x - \xi)^2 \right] dx \right| = M(\xi). \quad (9)$$

The sets of interest in the one-dimensional case are

$$C_1 = \{ y(x) \in L^2 : \ |\mathrm{Fr}[y]| = M(\xi) \}, \quad (10)$$

$$C_2 = \{ y(x) \in L^2 : \ |y(x)| = 1 \text{ for } x \in [-a/2, a/2] \text{ and } 0 \text{ elsewhere} \}. \quad (11)$$

Sometimes it is desirable to modify C_1 to C_1' defined as

$$C_1' = \{ y(x) \in L^2 : \ |\mathrm{Fr}[y]| \leqslant c(\xi) \}, \quad (12)$$

where $c(\xi)$ is a bound on the intensity of the Fresnel diffraction pattern. The set C_2 enforces the phase-only condition and is readily shown to be nonconvex. On the other hand, the set C_1' is readily shown to be convex. The projection onto C_1' is given in several places, e.g., Ref. 24, p. 311. The projection onto C_2 is given in Ref. 24, p. 290. The projection onto C_1 requires some discussion. We first note that C_1 is nonconvex because of the magnitude-equality constraint. To find the projection onto C_1 we rewrite Eq. (9) as

$$\left| \int_{-a/2}^{a/2} \tilde{t}(x, z) \exp\{-j\omega x\} dx \right| = M_z(\omega), \quad (13)$$

where

$$\tilde{t}(x, z) \equiv t(x) \Omega(x, z), \quad (14)$$

$$\Omega(x, z) \equiv \exp\left(j \frac{\pi}{\lambda z} x^2 \right), \quad (15)$$

$$\omega \equiv 2\pi\xi / \lambda z, \quad (16)$$

$$M_z(\omega) = \sqrt{\lambda z} M(\lambda z \omega / 2\pi). \quad (17)$$

We adopt the notation that when a lowercase function, say, $h(x)$, is multiplied by $\Omega(x, z)$, the resulting function is denoted by $\tilde{h}(x, z)$ and its Fourier transform; i.e., $F[\tilde{h}]$ is denoted by $\tilde{H}(\omega)$.

To find the projection of an arbitrary $h(x)$ on C_1, we seek to find a $y^*(x)$ such that

$$y^* = \underset{y \in C_1}{\mathrm{argmin}} \|h - y\|, \quad (18)$$

where argmin $f(x)$ refers to the value of x that minimizes the function $f(x)$. But $\|h - y\| = \|(h\Omega - y\Omega)\bar{\Omega}\| = \|h\Omega - y\Omega\| = \|\tilde{h} - \tilde{y}\|$, where the overbar denotes conjugation. Therefore $\min \|h - y\| = \min \|\tilde{h} - \tilde{t}\|$. Consider now the set $\tilde{C}_1$ defined by

$$\tilde{C}_1 = \{ \tilde{y}(x) \in L^2 : \ |\tilde{Y}(\omega)| = M_z(\omega) \}. \quad (19)$$

The projection of an arbitrary $\tilde{h}(x)$ onto $\tilde{C}_1$ is derived in several places[24,29] and is given by

$$\tilde{y}^*(x) = \tilde{P}_1 \tilde{h} \leftrightarrow M_z(\omega) \exp[j \theta_{\tilde{H}}(\omega)], \quad (20)$$

where $\theta_{\tilde{H}}(\omega)$ is the phase of $\tilde{H} = F[\tilde{h}]$. In Fig. 3 the point $\tilde{y}^*$ minimizes the distance from $\tilde{h}$ to $\tilde{C}_1$, and, likewise, the point y^* minimizes the distance from h to C_1. Since $\|h - y^*\| = \|\tilde{h} - \tilde{y}^*\|$, the projection of h onto C_1 is given by

$$y^*(x) = \tilde{y}^* \overline{\Omega(x, z)}. \quad (21)$$

Thus the algorithm

$$t_{k+1}(x) = P_1 P_2 t_k(x), \qquad t_0 \text{ arbitrary}, \quad (22)$$

where

$$P_1 h = \bar{\Omega} F^{-1}\{ M_z(\omega) \exp[j \theta_{\tilde{H}}(\omega)] \}, \quad (23)$$

$$P_2 h = \begin{cases} 0 & x \notin [-a/2, a/2] \\ \exp[j \phi_h(x)] & x \in [-a/2, a/2] \end{cases}, \quad (24)$$

and $\phi_h(x)$ is the phase of $h(x)$, enjoys the SDE property.

Provided that the sets intersect and the algorithm does not settle into a trap (if it does, one might want to choose a new starting point), the iteration in Eq. (22) converges to a $t(x) \in C_1 \cap C_2$.

Example 2. We refer to Fig. 4. It is required that a point source at $x = x_0$ in a plane z_1 units in front of a diffractive element generate a field with magnitude $M(\zeta)$ in

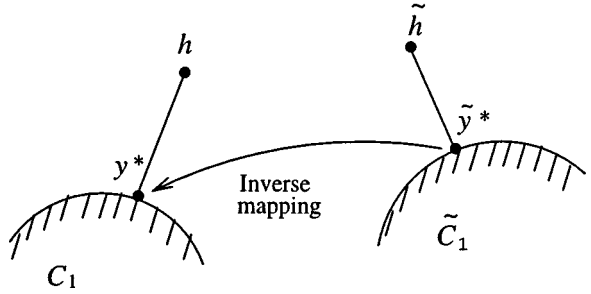


Fig. 3. Point $\tilde{y}^*$, which minimizes the distance from $\tilde{h}$ to $\tilde{C}_1$, permits the computation of y^*, the point that minimizes the distance from h to C_1, according to $y^* = \tilde{y}^* \overline{\Omega(x, z)}$.

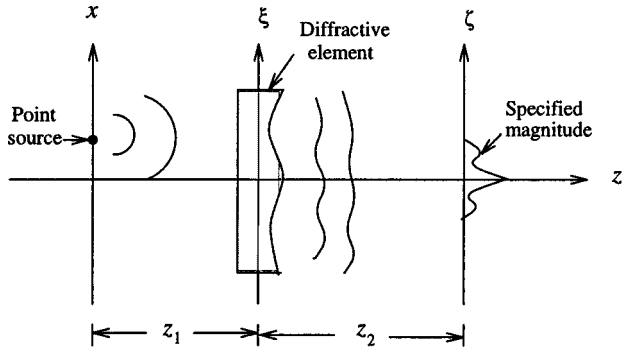


Fig. 4. Diffractive element that produces a specified optical field magnitude when illuminated by a point source a distance x_0 above the axis.

a plane a distance z_2 in back of the element. The analysis is simplified by a twice-repeated application of the Fresnel diffraction integral. Thus the field in the ξ plane immediately to the left of the element is

$$g_{z_1}(\xi) = \frac{1}{\sqrt{\lambda z_1}} \int_{-\infty}^{\infty} \delta(x - x_0) \exp\left[j \frac{\pi}{\lambda z_1} (x - \xi)^2 \right] \mathrm{d}x \tag{25}$$

$$= \frac{1}{\sqrt{\lambda z_1}} \exp\left[j \frac{\pi}{\lambda z_1} (x_0 - \xi)^2 \right], \tag{26}$$

and the field immediately to the right of the diffractive element is $g_{z_1}(\xi) t(\xi)$. Hence in the ζ plane we obtain

$$g_{z_2}(\zeta) = \frac{1}{\sqrt{\lambda z_2}} \int_{-\infty}^{\infty} g_{z_1}(\xi) t(\xi) \exp\left[j \frac{\pi}{\lambda z_2} (\zeta - \xi)^2 \right] \mathrm{d}\xi. \tag{27}$$

After some algebra, and recalling that it is the magnitude $M(\zeta)$ of $g_{z_2}(\zeta)$ that is of interest, we obtain

$$\left| \int_{-a/2}^{a/2} t(\xi) \exp\left\{ j \frac{\pi}{\lambda z} \xi^2 \right\} \exp(-j\omega\xi) \mathrm{d}\xi \right| = M_z(\omega), \tag{28}$$

where

$$z = \frac{z_1 z_2}{z_1 + z_2}, \tag{29}$$

$$\omega = \frac{x_0}{\lambda z_1} + \frac{\zeta}{\lambda z_2}, \tag{30}$$

$$M_z(\omega) = \lambda \sqrt{z_1 z_2} M\left(\lambda \omega z_2 - \frac{z_2}{z_1} x_0 \right). \tag{31}$$

Finally, if we let $\tilde{t}(\xi, z) = t(\xi) \exp\{ j(\pi/\lambda z) \xi^2 \}$, Eq. (28) adopts the same form as Eq. (13), and hence the design of the diffractive element in this example (Fig. 4) is identical to that of Example 1.

4. NUMERICAL SIMULATIONS

A. Influence of the Starting Point

The example we choose here corresponds to the configuration shown in Fig. 4 with the point x_0 chosen as $x_0 = 0$ for simplicity. The goal is to image the point source at $z_1 = 2f$ in front of the DOE into a diffraction-limited image at $z_2 = 2f$ in back of the DOE. Here f is a prescribed constant. In normalized quantities, a diffraction-limited, phase-only lens of diameter $a = 4$ will image a point source into a magnitude $M(\zeta)$ given by

$$M(\zeta) = 4 \left| \frac{\sin 2\pi\zeta}{2\pi\zeta} \right|. \tag{32}$$

We note that a tight upper bound on $M(\zeta)$ is given by $c(\zeta) = 0.85 + 3.15 \,\mathrm{rect}[2\zeta]$. Indeed, $M(\zeta)$ reaches $c(\zeta)$ at several points but never exceeds it. Starting with an arbitrary initialization, can the sequential generalized projection algorithm produce a DOE that images a point source within this bound?

For this problem the appropriate sets are C_1' as in Eq. (12), with $c(\zeta)$ as given above and C_2 as given in Eq. (11) with $a = 4$. By construction, the sets intersect so that feasible solutions exist. However, because C_2 is nonconvex, local stagnation at a trap appears to be a possibility. Because $M = 2$, SDE convergence is ensured.

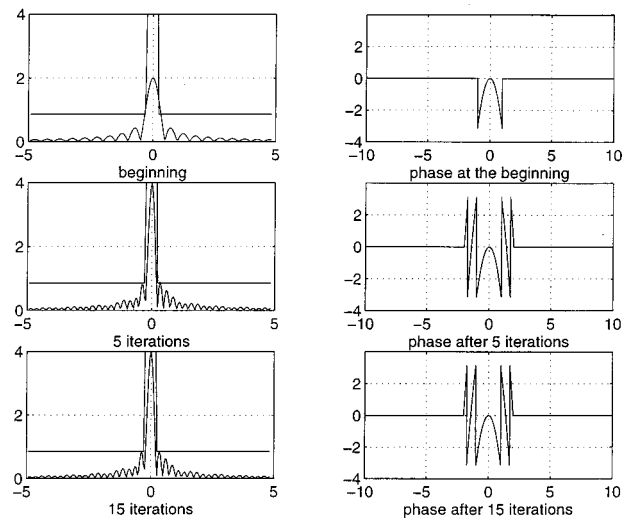


Fig. 5. Evolution of a lens that meets the design constraints when the starting point is a lens of insufficient size. On the left is the image; on the right is the phase of the DOE. The top row is the initial image formed from the initial phase. The middle row shows the same after 5 iterations. The bottom row is a feasible solution obtained after 15 iterations.

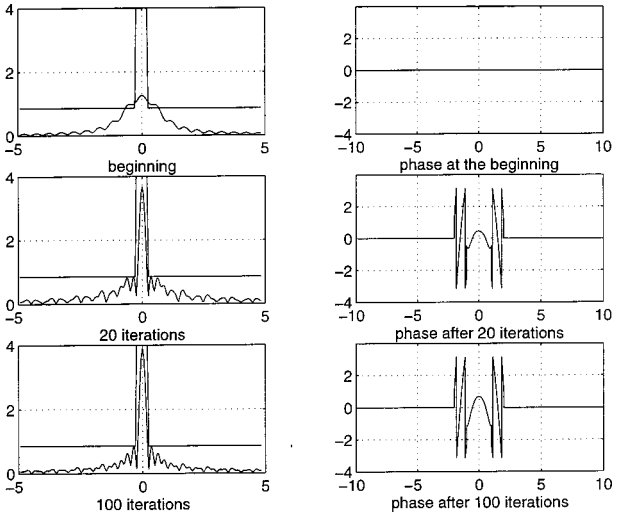


Fig. 6. Evolution of a lens that meets the design constraints when the starting point is a uniform section of glass. On the left is the image; on the right is the phase of the DOE. The top row is the initial image formed from the initial phase. The middle row shows the same after 20 iterations. The bottom row is a feasible solution obtained after 100 iterations. The extra iterations compared with Fig. 5 are due to the poorer starting point.

In Fig. 5 we show the results furnished by a projection algorithm when the initial element is a thin lens of (normalized) diameter $a = 2$. After five iterations the lens has grown to size $a = 4$ and has a phase profile characteristic of a Fresnel lens of diameter $a = 4$. No stagnation was observed.

The influence of the starting point is obviously important. For example, in Fig. 6 the initial object from which the appropriate lens evolves, iteration by iteration, is a uniform section of glass of diameter $a = 2$. After 100 it-

erations the projection algorithm produces a slightly different Fresnel lens than in the previous case but one that nevertheless meets the imaging constraint. Because the starting point is further from the solution region, the number of required iterations is almost seven times larger. Here, too, no stagnation was observed.

In addition to illustrating the effect of the starting point, this example also demonstrates that when one is using nonconvex sets, especially when $M = 2$, stagnation at a trap is often the exception rather than the rule.

B. Serial versus Parallel Projections

In our second example we are motivated to design a DOE that diffracts light away from the origin of the focal plane. Such a DOE is useful in observing the spectrum of film grain noise and other phenomena in the vicinity of the origin. The light diffracted by an ordinary lens is typically orders of magnitude above the noise spectrum in the vicinity of the origin, thus masking the noise there and making it impossible to measure. In theory (and to some extent in practice) it is possible to generate a null along a contour of measure zero, for example by using a π-phase plate near the Fourier lens in which the light exiting the plate in the region $(0, a/2)$ is out of phase by π from the light exiting in the region $(-a/2, 0)$, a being the aperture diameter.[30]

Obtaining a near-null region around the Fourier origin is more difficult. For one thing, the Fourier transform of an object with compact support is an analytic function in the complex plane. It is well known that the zeros of an analytic function are isolated unless the function is identically zero.[31] Hence it is not possible to generate a nontrivial function that furnishes an absolute null region over a measurable interval in the Fourier plane.

On the other hand, it may be possible to design a DOE whose diffraction pattern in the vicinity of the origin has sufficiently low intensity to enable the measurement of weak signals arising from grain noise or other phenomena of interest.

Before attempting a solution by vector-space methods, it is of interest to describe the problem mathematically. We seek to find a pure-phase device $t(x)$ that, when incorporated with a lens of transmittance $\exp[-j(2\pi/\lambda f)^2]\mathrm{rect}[x/a]$, will satisfy the integral equation

$$\left| \int_{-a/2}^{a/2} t(x) \exp\left(-j \frac{2\pi}{\lambda f} x^2 \right) \exp\left[j \frac{\pi}{\lambda f} (x - \zeta)^2 \right] \mathrm{d}x \right|, \tag{33}$$

$$\overset{u \equiv \zeta/\lambda z}{=} \left| \int_{-a/2}^{a/2} t(x) \exp(-j 2\pi u x) \mathrm{d}x \right| \leqslant \epsilon(u) \quad \text{for } u \in D_1, \tag{34}$$

where $\epsilon(u) \geqslant 0$ is a predetermined prescribed bound on the magnitude of the background light and the surface D_1 consists of the region to be examined. In addition to satisfying the inequality in Eq. (34), $t(x)$, being pure phase, must also satisfy $|t(x)| = 1$ for $-a/2 \leqslant x \leqslant a/2$. It appears that solving a problem of this type by analytic means is very difficult.

For the vector-space approach to this problem we define the following three sets:

$$C_1 = \{ y(x) \in L^2 : \ |Y(u)| \leqslant \epsilon(u) \ \text{ for } u \in D_1 \} \tag{35}$$

$$C_2 = \left\{ y(x) \in L^2 : \ \begin{array}{ll} |y(x)| = 1 & \text{for } x \in \left[\dfrac{-a}{2}, \dfrac{a}{2} \right] \\ y(x) = 0 & \text{otherwise} \end{array} \right\}, \tag{36}$$

$$C_3 = \{ y(x) \in L^2 : \ Y(u) = Y(-u) \quad \text{for } u \in D_3 \}. \tag{37}$$

Although there may be other choices of constraint sets, the projectors onto the above are easily computed. We note that C_2 is identical to the set C_2 described in Eq. (11); it is repeated here for the reader's convenience.

In words, C_1 is the set of all functions whose Fourier-transform magnitude cannot exceed $\epsilon(u)$ for all frequencies in the region D_1. C_3 is the set of all functions that yield symmetric diffraction orders in the frequency region D_3. The sets C_1 and C_3 are convex; C_2 is not. The projection of an arbitrary signal $h(x)$ onto C_1 is given by

$$P_1 h(x) \leftrightarrow \begin{cases} \epsilon(u) \exp[j \theta_H(u)] & \text{if } |H(u)| > \epsilon(u) \\ & \text{and } u \in D_1 \\ H(u) & \text{otherwise} \end{cases} . \tag{38}$$

As always, the double-headed arrow means a Fourier-transform pair and $\theta_H(u)$ is the phase of $H(u)$. The projection onto C_2 is given in Eq. (24). Likewise, the projection onto C_3 is readily computed to be

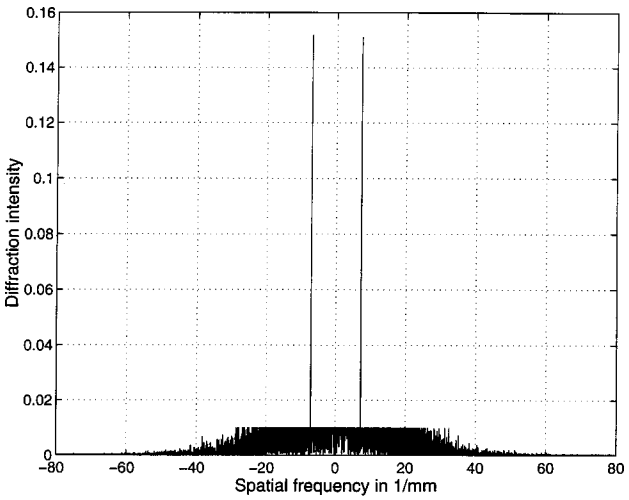
$$P_3 h(x) \leftrightarrow \begin{cases} \frac{1}{2}[H(u) + H(-u)] & \text{if } H(u) \neq H(-u) \\ & \text{and } u \in D_3 \\ H(u) & \text{otherwise} \end{cases} . \tag{39}$$

The algorithm

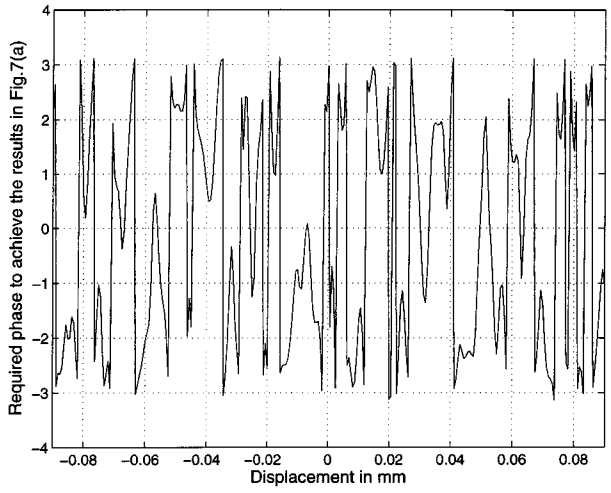
$$t_{n+1} = P_2 P_1 P_3 t_n(\mathbf{x}), \qquad t_0(\mathbf{x}) = 1 \tag{40}$$

was applied with $D_1 = (-\infty, -u_0/2 - \Delta) \cup (-u_0/2, u_0/2) \cup (u_0/2 + \Delta, \infty)$ and $D_3 = (-u_0/2 - \Delta, -u_0/2) \cup (u_0/2, u_0/2 + \Delta)$. For this problem we used the following parameters: $a = 40$ mm, space-domain pixel size $\Delta x = 6.25 \times 10^{-3}$ mm, frequency-domain pixel size $\Delta u = 9.76 \times 10^{-3}$ mm^{-1}, and fast-Fourier-transform size $N = 16{,}384$ points. The set parameters $u_0/2$ and Δ were, respectively, 7.12 mm^{-1} and 0.05 mm^{-1}.

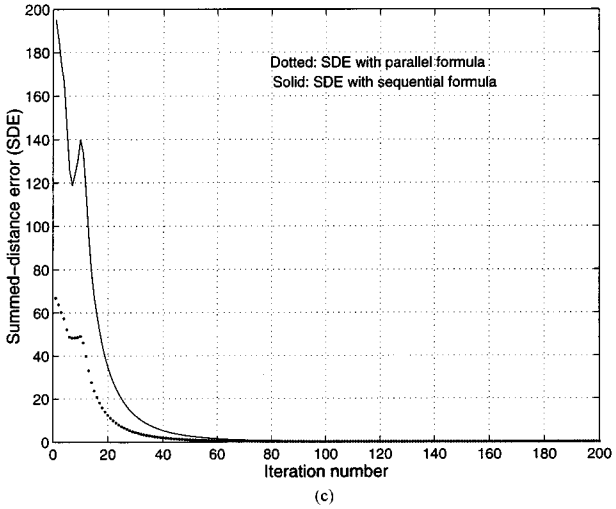
De facto convergence was readily achieved for various starting points when the sequential generalized projection algorithm (SGPA) was used with $\epsilon(u) = \epsilon_0 = 0.01$. In Fig. 7(a) we show the diffraction peaks that are moved away from the origin. The high-frequency and near-dc magnitudes never exceed $\epsilon_0 = 0.01$. For the reader's interest we show in Fig. 7(b) a small portion of the DOE phase. Such fluctuations may raise problems in manufacturing, but these considerations are beyond the scope of this paper. Of interest are the results in Fig. 7(c); there we see the failure of SDE convergence for the SGPA, regardless of which distance error formula is used, i.e., Eq. (3) or Eq. (6). This happens, in this example, because the condition of the Levi–Stark theorem, namely, that $M = 2$, is violated since $M = 3$.



(a)



(b)

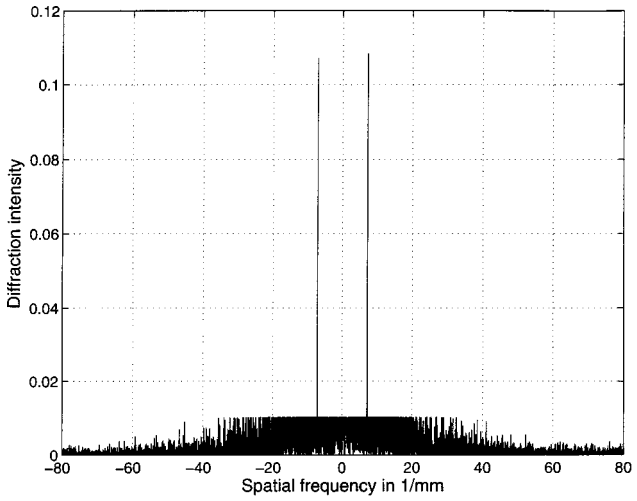


(c)

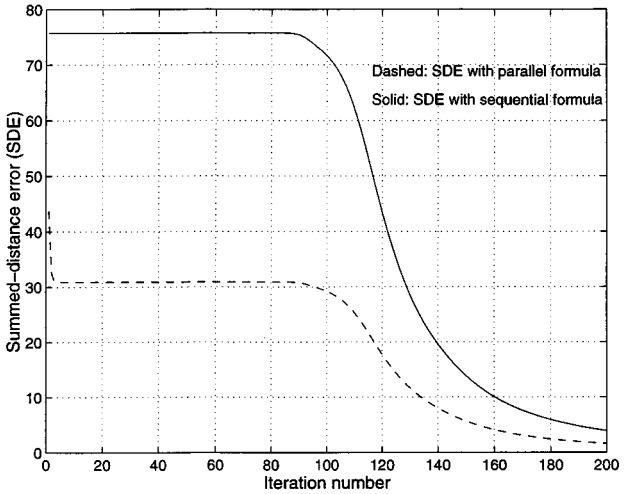
Fig. 7. Results of applying the SGPA to the low-light-level problem at dc after 350 iterations. (a) Diffraction peaks moved away from the origin, (b) a small portion of the highly fluctuating phase of the DOE, (c) SDE behavior with application of the SGPA.

Another interesting and unexpected result is what happens when the PGPA is applied to the same problem. Here, too, we observe a diffraction pattern that meets the constraints [Fig. 8(a)], and, unlike the SGPA, the SDE for the PGPA exhibits SDE convergence [Fig. 8(b)]. The unexpected part of this result is that very early in the iteration the PGPA goes into a temporary quasi-stationary point in which progress toward the solution is exceedingly slow [the flat portion of the SDE curves in Fig. 8(b)]. The appearance of this phenomenon depends on the initialization, a fact we observed by varying the starting point. As in Fig. 7(a), the results shown in Fig. 8(a) are after 350 iterations.

When a more stringent constraint is imposed on the high-frequency and near dc, convergence was not attainable with either the SGPA [Fig. 9(a)] or the PGPA [Fig. 9(b)]. Yet, in one sense, one can say with some justification that the PGPA outperformed the SGPA since, on a SDE formula-by-formula comparison, the SDE for the former was less than that of the latter. However, the PGPA does not yield a DOE with pure phase. [Observe the fluctuations in magnitude of the DOE transmittance in Fig. 9(c)].

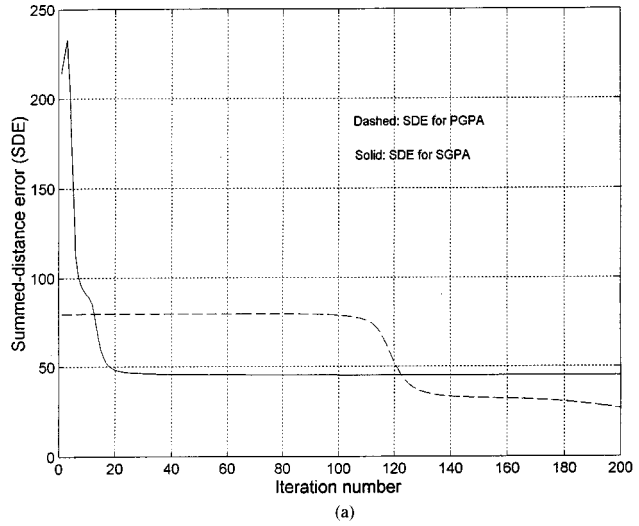


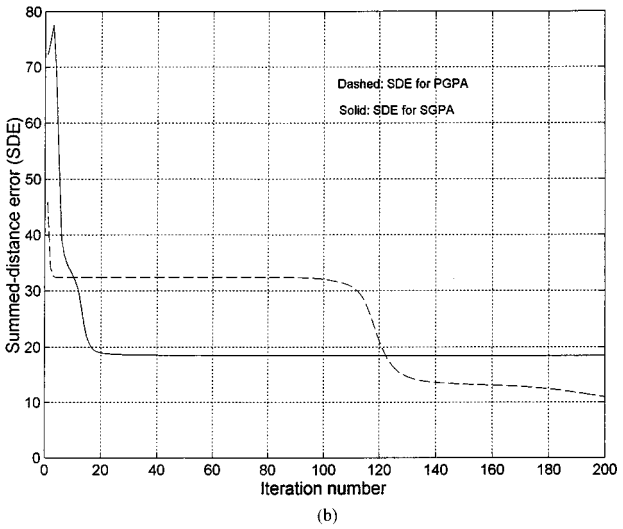
(a)



(b)

Fig. 8. Results of applying the PGPA to the low-light-level problem at dc after 350 iterations. (a) Diffraction peaks moved away from the origin, (b) SDE behavior with application of the PGPA.





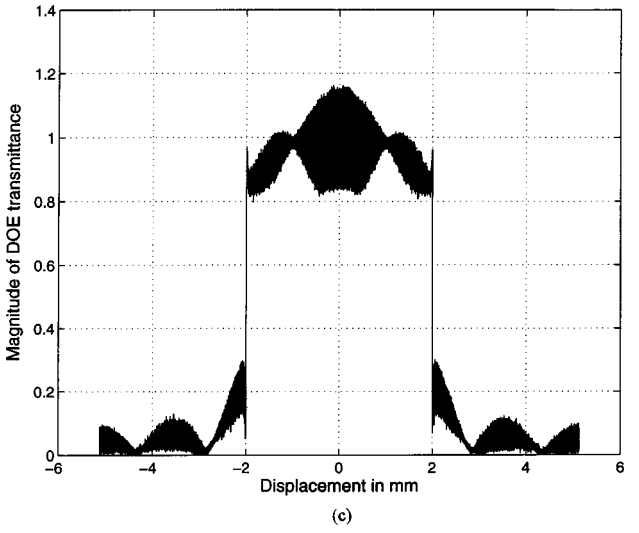


Fig. 9. Failure of both the SGPA and the PGPA to converge strongly when a more stringent constraint is imposed. (a) SDE behavior computed with the SGPA formula, (b) SDE behavior computed with the PGPA formula, (c) resulting magnitude of DOE transmittance. Note the inability of the PGPA to furnish a pure-phase solution.

C. Nonintersecting Sets

In the third and final example, we consider the design of a DOE that yields an increased depth of focus. The configuration is essentially that of Fig. 4 with the point source on axis at $z_1 - \infty$. If we replace the diffractive element with an ordinary lens of focal length f, the image forms at $z_2 = f$. For simplicity we write $z = z_2$ since the subscript is unnecessary.

At a distance z beyond the lens whose diameter is a, it is easily shown that the field $U_z(\zeta)$ is given by

$$U_z(\zeta) = K_z T_z(\zeta), \tag{41}$$

where

$$T_z(\zeta) = \int_{-0.5}^{0.5} \exp[-j\pi b(1 - r)y^2]\exp(-j2\pi r\rho\zeta y)\mathrm{d}y, \tag{42}$$

and

$$K_z = \frac{a}{\sqrt{\lambda z}}, \qquad r = f/z,$$

$$\rho = a/\lambda f, \qquad b = a^2/\lambda f = a\rho,$$

and other unimportant constants have been discarded. For the choice of parameters $a = 25$ mm, $f = 150$ mm and $\lambda = 0.5 \times 10^{-3}$ mm, the field magnitudes for $r = 1.1$ (just before the focal plane), $r = 1$ (at the focal plane), and $r = 0.91$ (just after the focal plane) are shown, respectively in Figs. 10(a), 10(b), and 10(c). We note that a 10% axial displacement from the focal plane yields a 28-dB drop in intensity.

We now wish to increase the depth of focus of our configuration by adding, next to the lens, a second, phase-only element $t(x)$. Then instead of Eqs. (41) and (42) we obtain

$$\tilde{U}_z(\zeta) = K_z \tilde{T}_z(\zeta), \tag{43}$$

where

$$\tilde{T}_z(\zeta) = \int_{-0.5}^{0.5} t(y)\exp[-j\pi b(1 - r)y^2]\exp(-j2\pi r\rho\zeta y)\mathrm{d}y. \tag{44}$$

Next we observe that the image of the point source at infinity has a normalized magnitude height of unity and a sidelobe level of 0.21. Since our aim is to create a set of constraints that lead to nonintersecting sets, we impose the (impossibly) demanding constraints that $t(x)$ have membership in the following sets:

$$C_1 = \{t(x) \in L^2: \quad |\tilde{U}_{f-\epsilon}(\zeta)| \leqslant c(\zeta), \ |\tilde{U}_{f-\epsilon}(0)| = K_f\}, \tag{45}$$

$$C_2 = \{t(x) \in L^2: \quad |\tilde{U}_f(\zeta)| \leqslant c(\zeta), \quad \tilde{U}_f(0) = K_f\}, \tag{46}$$

$$C_3 = \{t(x) \in L^2: \quad |\tilde{U}_{f+\epsilon}(\zeta)| \leqslant c(\zeta), \quad \tilde{U}_{f+\epsilon}(0) = K_f\}, \tag{47}$$

$$C_4 = \{t(x) \in L^2: \quad |t(x)| = \mathrm{rect}(x/a)\}, \tag{48}$$

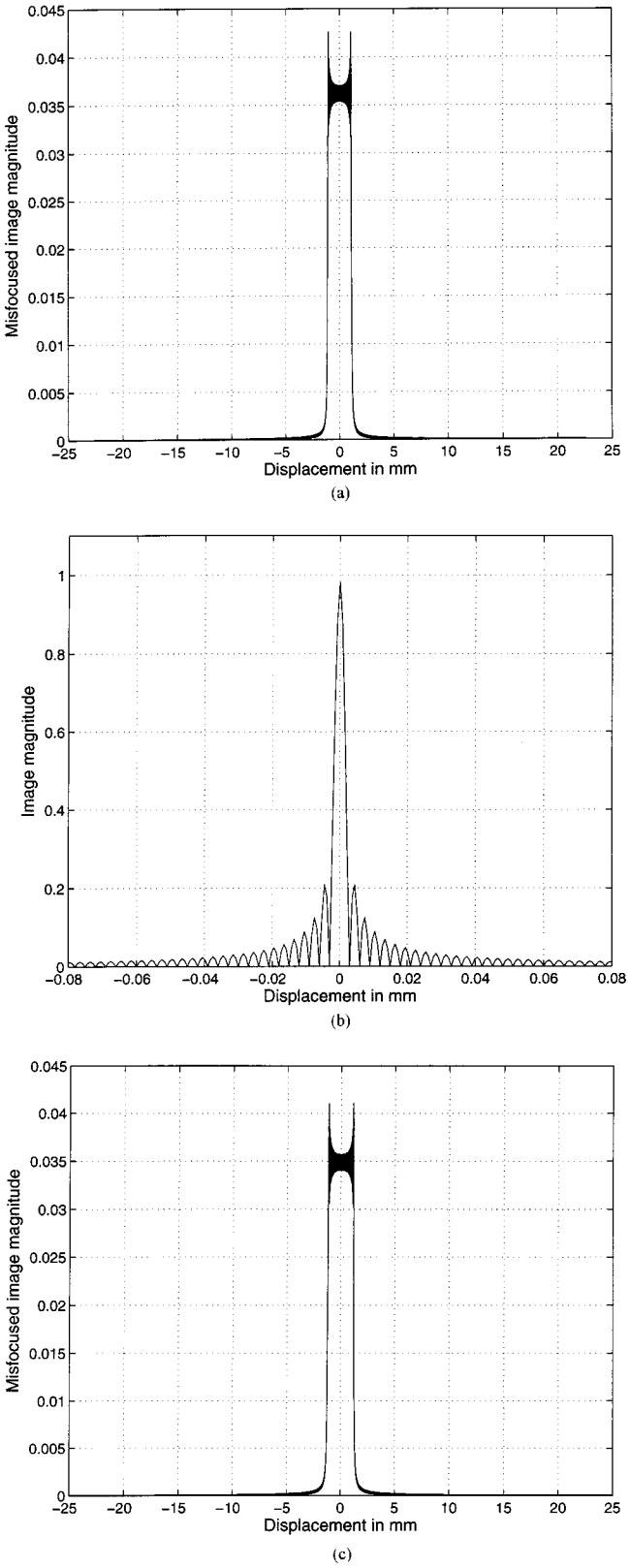


Fig. 10. Image magnitude due to a point source at infinity: 10% misfocusing error (a) in front of image (focal) plane, (b) in focal plane, (c) in back of focal plane.

where $c(\zeta) = K_f[\epsilon + (1 - \epsilon)\mathrm{rect}(\zeta/2\sigma)]$ and $\epsilon = 0.05$, and σ is the distance from the origin to the first null of the image at f.

The results for the SGPA after 250 iterations are shown in Figs. 11(a), 11(b), and 11(c) for $z = f - \epsilon$, $z = f$, and $z = f + \epsilon$, respectively. We note, as expected,

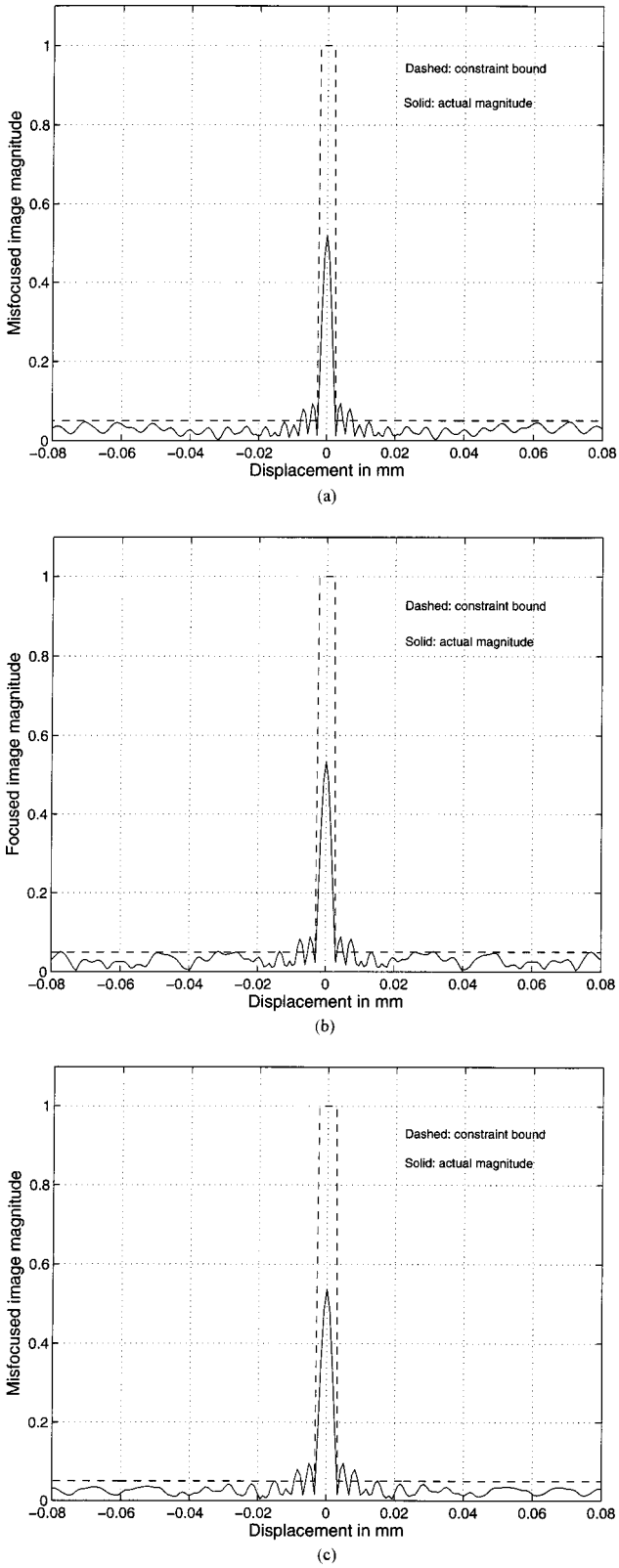


Fig. 11. Designing increased depth-of-focus optics with the SGPA. Magnitude diffraction pattern due to a point source at infinity: 10% misfocusing error (a) in front of image (focal) plane, (b) in focal plane, (c) in back of focal plane.

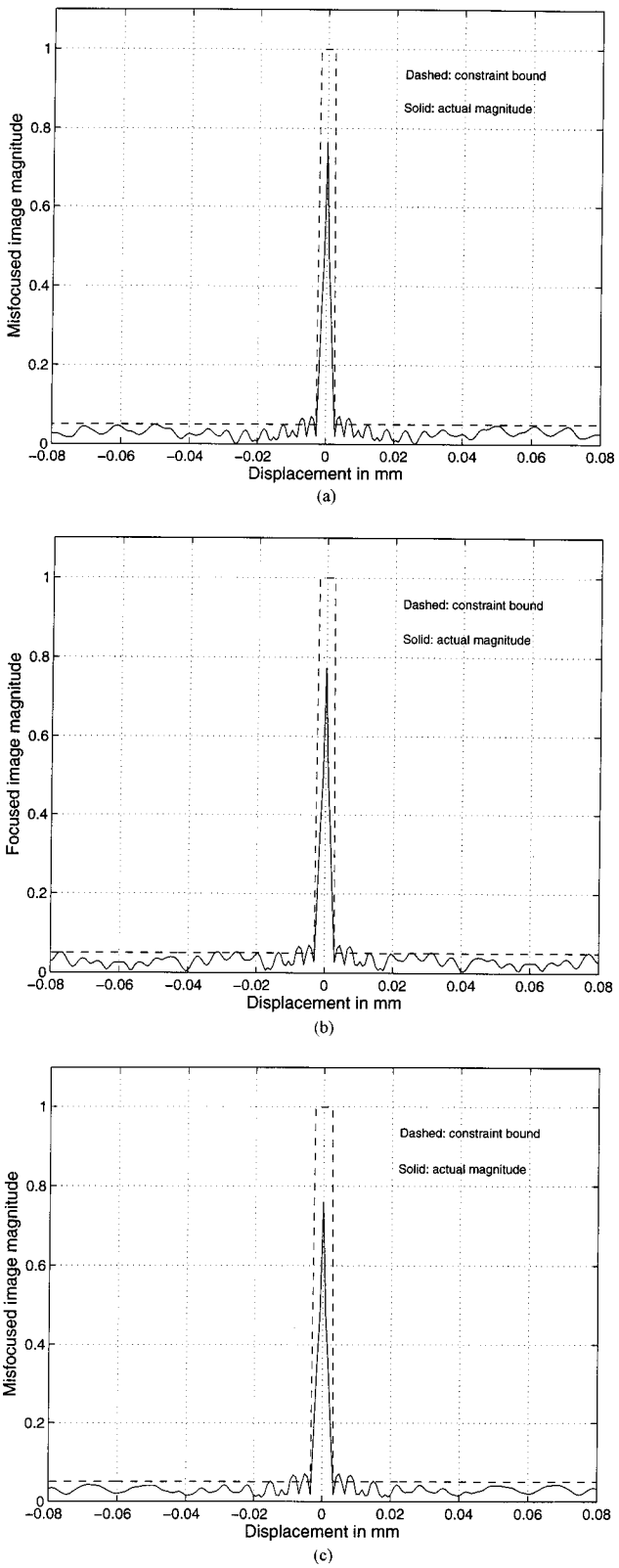


Fig. 12. Designing increased depth-of-focus optics with the PGPA. Magnitude diffraction pattern due to a point source at infinity: 10% misfocusing error (a) in front of image (focal) plane, (b) in focal plane, (c) in back of focal plane.

that although the constraints are not met, the three field-magnitude distributions at $f - \epsilon$, f, and $f + \epsilon$, respectively, are remarkably similar, with a peak-to-maximum sidelobe level of at least 13.5 dB. So, in effect, we have achieved our goal of getting three diffraction-limited images of the point source at infinity in three different planes, albeit at reduced intensity.

The point is that although a feasible solution has not been attained, an entirely reasonable solution is available if a broader interpretation of the results is taken.

We also applied the PGPA to this problem. The results after 250 iterations are shown in Figs. 12(a), 12(b), and 12(c) for $z = f - \epsilon$, $z = f$, and $z = f + \epsilon$, respectively. In one sense the PGPA appears to outperform the SGPA in that, although the three images are still remarkably similar, the central image is higher, the sidelobes are lower, and the peak-to-maximum-sidelobe ratio is 19 dB or better.

On the other hand, since this is the parallel algorithm, $t(x)$ is not quite pure phase, unlike the $t(x)$ produced by the SGPA, which is pure phase because the final step in the iteration is the projection onto the set of pure-phase functions.

5. CONCLUSIONS

The powerful method of vector-space projections was discussed and applied to designing several DOE's. We illustrated the properties of these algorithms with respect to strong and summed-distance-error convergence, serial versus parallel projections, and the properties of solutions when sets do not intersect. In the latter case reasonable solutions are still possible, especially with the PGPA, which yields a weighted minimum-distance solution from the constraint sets.

REFERENCES

1. D. Wood, P. McKee, and M. Dames, "Multiple-imaging and multiple-focussing Fresnel lenses with high numerical aperture," in *Holographics International '92*, Y. N. Denisyuk and F. Wyrowski, eds., Proc. SPIE **1732**, 307–316 (1993).
2. G. Hatakoshi and M. Nakamura, "Grating lenses for optical branching," Appl. Opt. **32**, 3661–3667 (1993).
3. M. Bernhardt, F. Wyrowski, and O. Bryngdahl, "Coding and binarization in digital holography," Opt. Commun. **77**, 4–8 (1990).
4. M. Bernhardt, F. Wyrowski, and O. Bryngdahl, "Iterative techniques to integrate different optical functions in a diffractive phase element," Appl. Opt. **30**, 4629–4635 (1991).
5. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik (Stuttgart) **35**, 235–246 (1972).
6. B. C. Kress and S. H. Lee, "Iterative design of computer generated Fresnel holograms for free-space optical interconnections," in *Optical Computing*, Vol. 7 of 1993 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1993), pp. 22–25.
7. F. Wyrowski and G. M. Morris, guest eds., "Diffractive optics," feature issue of J. Mod. Opt. (April 1993).
8. M. G. Moharam and J. Leger, guest eds., "Diffractive optics modeling," feature issue of J. Opt. Soc. Am. A **12**, 1025–1169 (1995).
9. F. Wyrowski, "Diffractive optical elements: iterative calculation of quantized, blazed phase structures," J. Opt. Soc. Am. A **7**, 961–969 (1990).
10. F. Wyrowski, "Iterative quantization of digital amplitude holograms," Appl. Opt. **28**, 3864–3870 (1989).
11. H. Stark, W. C. Catino, and J. L. LoCicero, "Design of phase gratings by generalized projections," J. Opt. Soc. Am. A **8**, 566–571 (1991).

12. W. C. Catino, J. L. LoCicero, and H. Stark, "Design of continuous and quantized phase holograms by generalized projections," J. Opt. Soc. Am. A **14**, 2715–2725 (1997).
13. W. C. Catino, J. L. LoCicero, and H. Stark, "Design of continuous and quantized amplitude holograms by generalized projections," J. Opt. Soc. Am. A **15**, 68–76 (1998).
14. K. S. Urquhart, P. Marchand, Y. Fainman, and S. H. Lee, "Diffractive optics applied to free-space optical interconnects," Appl. Opt. **33**, 3670–3682 (1994).
15. J. Leger and M. G. Moharam, guest eds., "Diffractive optics," feature issue of Appl. Opt. **36**, 2389–2604 (1995).
16. I. Cindrich and S. H. Lee, eds., *Diffractive and Holographic Optics Technology III*, Proc. SPIE **2689** (1996).
17. *Diffractive Optics and Micro-Optics*, Vol. 5 of 1996 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1996).
18. H. P. Herzig, *Micro-Optics: Elements, Systems and Applications* (Taylor & Francis, London, 1970).
19. A. A. Sawchuk and T. C. Strand, "Digital optical computing," Proc. IEEE **72**, 758–779 (1984).
20. G. M. Morris, "Diffractive optics technology and its applications," short course notes **104**, in *Annual Meeting* (Optical Society of America, Washington, D.C., 1997).
21. R. Piestun and J. Shamir, "Control of wave-front propagation with diffractive elements," Opt. Lett. **19**, 771–773 (1994).
22. R. Piestun, B. Spektor, and J. Shamir, "Diffractive optics for unconventional light distribution," in *Diffractive and Holographic Optics Technology II*, I. Cindrich and S. H. Lee, eds., Proc. SPIE **2404**, 320–327 (1995).
23. A. Levi and H. Stark, "Image restoration by the method of generalized projections with application to restoration from magnitude," J. Opt. Soc. Am. A **1**, 932–943 (1984).
24. H. Stark and Y. Yang, *Vector Space Projections: A Numerical Approach to Signal and Image Processing, Neural Nets, and Optics* (Wiley, New York, 1998).
25. T. Kotzer, J. Rosen, and J. Shamir, "Application of serial and parallel-projection methods to correlation-filter design," Appl. Opt. **34**, 3883–3894 (1995).
26. G. Pierra, "Decomposition through formalization in a product space," Math. Program. **28**, 96–115 (1984).
27. P. L. Combettes, "Inconsistent signal feasibility problem: least-squares solution in a product space, IEEE Trans. Signal Process. **42**, 2955–2966 (1994).
28. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).
29. A. Levi, "Image restoration by the method of projections with applications to the phase and magnitude retrieval problems," Ph.D. dissertation (Rensselaer Polytechnic Institute, Troy, N.Y., 1983).
30. H. Stark, "Theory and measurements of the optical Fourier transform," in *Applications of Optical Fourier Transforms*, H. Stark, ed. (Academic, New York, 1982), pp. 1–40.
31. R. V. Churchill, J. W. Brown, and R. F. Verhey, *Complex Variables and Applications*, 3rd ed. (McGraw-Hill, New York, 1974), p. 167.