

Determining Sets for the Discrete Laplacian*

Aviad Rubinstein[†]
Jacob Rubinstein[†]
Gershon Wolansky[‡]

Abstract. We define a notion of determining sets for the discrete Laplacian in a domain Ω . A set D is called determining if harmonic functions are uniquely determined by providing their values on D , and if D has the same size as the boundary of Ω . It is shown that there exist determining sets that are fairly evenly distributed in Ω . A number of basic properties of determining sets are derived.

Key words. Laplacian, determining sets, uniqueness

AMS subject classifications. 65N06, 05C99

DOI. 10.1137/050645233

I. Background. Harmonic functions play a prominent role in many branches of mathematics, science, and technology. A function u is called harmonic in a domain $\Omega \in \mathbb{R}^n$ if it solves the Laplace equation there. It is well known that a harmonic function is uniquely determined by prescribing its values on the boundary $\partial\Omega$ of Ω . Consider a set D that is in some appropriate sense of the same size as $\partial\Omega$. Such a set D will be called a *determining set* if prescribing the value of a harmonic function on it uniquely determines that function.

We shall consider first a two-dimensional discrete version of the question above. Moreover, we shall study for simplicity the special case where the domain Ω is a square. Several generalizations will be discussed later. A function $u(x, y)$ is said to be harmonic in Ω if it satisfies there the partial differential equation

$$(1.1) \quad \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad (x, y) \in \Omega.$$

The notion of harmonic functions is also well-defined in discrete mathematics: Basically, a function is said to be harmonic over a graph if its value at any vertex v is the average of its values at the nearest neighbors of v . Again, limiting ourselves first to squares, we shall formulate the problem of determining sets for discrete harmonic functions for the special graph consisting of a rectangular grid

$$G = \{(i, j) \mid 1 \leq i \leq N, 1 \leq j \leq N\}$$

*Received by the editors November 15, 2005; accepted for publication (in revised form) August 16, 2006; published electronically May 1, 2007. This work was supported in part by an NSF grant and by an ISF grant.

<http://www.siam.org/journals/sirev/49-2/64523.html>

[†]Department of Mathematics, Indiana University, Bloomington, IN 47405 (the3big6leb9@hotmail.com, jrubinst@indiana.edu).

[‡]Department of Mathematics, Technion, Haifa 32000, Israel (gershonw@math.technion.ac.il).

for some integer $N \geq 3$. Each point in G is connected to its nearest neighbors. There are three types of points in G : The set $I = \{(i, j) \mid 2 \leq i \leq N-1, 2 \leq j \leq N-1\}$ of interior points that consists of all points that have 4 nearest neighbors; the set B of boundary points that consists of the $4(N-2)$ points that have only 3 neighbors; and the remaining points (the four vertices of the square) that have no interior point as a neighbor. A function $u : G \rightarrow R$ is said to be harmonic if

$$(1.2) \quad u(i, j) = \frac{1}{4} (u(i+1, j) + u(i-1, j) + u(i, j+1) + u(i, j-1)) \quad \forall (i, j) \in I.$$

This equation is also well known as the finite difference approximation of the Laplace equation (1.1). Since the four vertices of the square $(1, 1), (1, N), (N, N), (N, 1)$ play no role in (1.2) we eliminate them from G . Equation (1.2) models a variety of transport processes over graphs. In addition, it forms a standard discrete approximation of the continuous Laplace equation.

To obtain a unique discrete harmonic function u , one needs to supplement (1.2) with additional conditions. A common set of such auxiliary data is the Dirichlet condition, where the values of u are given on the boundary points B :

$$(1.3) \quad u(i, j) = g(i, j), \quad (i, j) \in B.$$

The following result summarizes the standard theory of existence and uniqueness of discrete harmonic functions (see, for example, [4]).

THEOREM 1. *The joint problem consisting of (1.2) and (1.3) has a unique solution. Also, the set K of solutions for (1.2) is a linear space of dimension $4(N-2)$.*

Proof. The key idea for the first part of the theorem is the maximum principle: A nonconstant function u satisfying (1.2) cannot attain its maximum at a point $(i, j) \in I$. To justify this principle notice that if $(i, j) \in I$, then $u(i, j)$ is the arithmetic average of its nearest neighbors. Clearly if an average of a set of numbers is greater than or equal to each of the numbers in the set, then all the numbers in the set equal the average. Therefore, if u attains a maximum at some interior point, then the same maximum is also attained by each neighbor of this point. We can continue this process until we cover all the points in the rectangle. Thus u is constant. Exactly the same argument shows that the minimum of u also cannot be obtained at a point in the set I .

Consider next the system (1.2) with the homogeneous Dirichlet condition, i.e., $u(i, j) = 0$ for $(i, j) \in B$. The maximum principle implies at once that the function $u(i, j) = 0$ for all $(i, j) \in I$ is the unique solution. Now, (1.2) forms an inhomogeneous linear algebraic system of $(N-2)^2$ equations in $(N-2)^2$ unknowns (the values of u on I). A standard result in linear algebra states that if the trivial solution $u = 0$ is the only solution of a homogeneous system (which was established in the previous paragraph), then the associated inhomogeneous system has a unique solution.

To prove the second part of the theorem, let $T_k(i, j)$ be the solution of the system (1.2) under the Dirichlet condition

$$(1.4) \quad T_k = \begin{cases} 1 & \text{if } p = p_k, \\ 0 & \text{if } p \neq p_k, \end{cases}$$

where the index of p sweeps over all the points in B . The existence of a unique T_k for each k is guaranteed by the first part of the theorem. The linearity of (1.2) implies that the solution for an arbitrary Dirichlet condition g can be expressed as a linear combination of the form

$$u(i, j) = \sum_{k=1}^{4(N-2)} g_k T_k(i, j),$$

where g_k is the value taken by g at the point $p_k \in B$. Moreover, it is obvious that the functions T_k are linearly independent.

The classical picture presented above is simple and complete. It applies with little change to all elliptic partial differential equations. However, it does not address situations where it is hard to obtain (say, through measurements) information on u at the boundary, but it is possible to obtain certain a priori information on u on some subset of the closure of Ω . Let us describe now a specific physical situation where such a problem arises. In many applications it is necessary to find the phase of the wave. Practically, phase measurements are hard to obtain. On the other hand, it is relatively easy to measure the intensity of a wave. Therefore it would be useful to develop methods for determining the phase of a wave from intensity measurements. Denote the intensity of the wave by $E(x, y, z)$ and its phase by $\phi(x, y, z)$, where z is a point along the optical axis and (x, y) is a point in the plane orthogonal to the optical axis. The transport (Fresnel) equation for paraxial waves is [7], [5]

$$(1.5) \quad \nabla \cdot (E \nabla \phi) = -\frac{\partial E}{\partial z},$$

where ∇ is the two-dimensional gradient operator and the equation is considered in a finite planar domain Ω . In one example of an intensity sensor one measures the intensity (say, using a CCD camera) at two nearby planes orthogonal to the z axis. This measurement is used to estimate both E and $\partial E/\partial z$. One then considers (1.5) as a partial differential equation for the phase ϕ . Solving the equation gives the phase. The difficulty is that the values of ϕ are not known on the boundary. Therefore the elliptic problem (1.5) is not well-posed. Several ideas were proposed to remedy the situation, but none of them provides a rigorous solution (see [6] for a brief review and for an alternative approach that is not based on (1.5) directly).

A commonly used phase sensor is based on the Hartmann–Shack (HS) principle. An HS phase sensor consists of an array of lenslets that convert an incident wave into an array of spots on a screen. The phase slopes, that is, the functions $\partial\phi/\partial x$ and $\partial\phi/\partial y$, are estimated from the spot locations. It is known to be a robust sensor, but its resolution is limited since the lenslets cannot be made too small. An interesting approach to a high resolution phase sensor is to use the data from an HS sensor to replace the unavailable boundary condition for ϕ on $\partial\Omega$. The advantage of the proposed hybrid sensor is that it would enjoy the high resolution obtained by the intensity sensor (the resolution of a CCD camera can be several orders of magnitude larger than the resolution of an HS screen).

2. D-Sets. In the previous section we used the term determining sets to describe any set such that prescribing the values of u on it determines a unique solution for the system (1.2). From now on, we shall give this title only to such sets that are of the same size as the boundary set B . Thus, a set D of $4(N-2)$ points in G will be called a determining set, or a *D-set* for short, if prescribing $u(p) = 0 \forall p \in D$ implies $u = 0$ everywhere in G . Our first example of a D-set is the boundary set B itself. One can check that replacing a point in B by its nearest neighbor in I provides a new D-set. One of the questions we ask is whether there exist more “interesting” D-sets. In particular, an “interesting” D-set is a set that is fairly evenly distributed in G . Other questions of interest include the connection between the different D-sets, the probability that a set of $4(N-2)$ points in G is a D-set, etc. We emphasize that even when the D-set does not consist of the boundary set B , the system (1.2) is considered only for points $(i, j) \in I$.

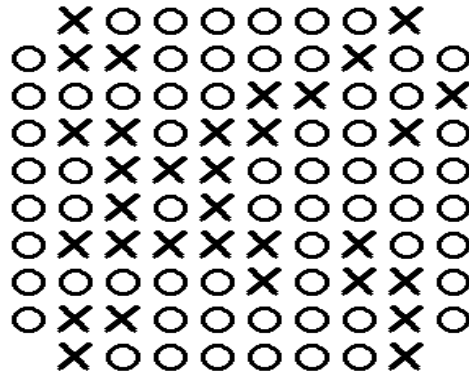


Fig. 1 An example of a D -set. The vertices of the D -set are marked by “x”, while the other vertices are marked as “o”.

The problem of deciding whether a given set is a D -set is equivalent to checking whether the appropriate system of $(N - 2)^2$ homogeneous linear equations derived from (1.2) has a unique solution. Our approach, however, will be more geometrical, since the matrix above could be complicated. Nevertheless, we shall return in the last section to the algebraic approach; it will be shown there that the determination of a D -set requires considering the invertibility of a smaller matrix whose size is just $4(N - 2) \times 4(N - 2)$. An example of a D -set is given in Figure 1. The points in the D -set are marked by x’s, while the other points are marked with o’s.

A serious difficulty in determining whether a given set is a D -set is that in general the sets do not contain points forming a closed orbit. Therefore, the maximum principle that is so essential in establishing uniqueness in elliptic problems may not be applicable. Fortunately, another very simple principle that we call the T -principle can sometimes be invoked: If u vanishes at an interior point p and on three of the nearest neighbors of p , then u must also vanish on the fourth nearest neighbor of p .

We start by constructing a large family of D -sets that are to some extent distributed over the entire G . It is convenient to present the construction algorithm first for a slightly different graph than G . Consider, then, the graph H_2^N , consisting of the points $(i, j) \in \mathbb{Z}^2$ such that $|i| + |j| \leq N$. Each point is connected by an edge to its nearest neighbors. The boundary of H_2^N consists of the set $|i| + |j| = N$. Therefore, a D -set should now consist of $4N$ points. We construct D -sets on shells around the origin. We first choose all the points $|i| + |j| = 1$; then for each shell $|i| + |j| = k$, $k = 2, 3, \dots, N$, we select exactly one point on each edge, except for the forbidden selection of all the four vertices of the shell. We thus define a set of $4N$ points. To see that it is a D -set, we work layer by layer, starting from the seed $(i, j) = (0, 0)$ outward. Indeed, u must vanish at the seed thanks to the maximum principle. By construction u vanishes at all the points in the first layer. Applying repeatedly the T -principle for each successive layer, one can validate the construction. Since the four points selected for each shell are arbitrary, the resulting set can be made well-distributed in H_2^N .

A similar *seed and shell* process applies for the original graph G . Select an arbitrary point as a “seed.” Proceed to construct the D -set on equidistant shells about

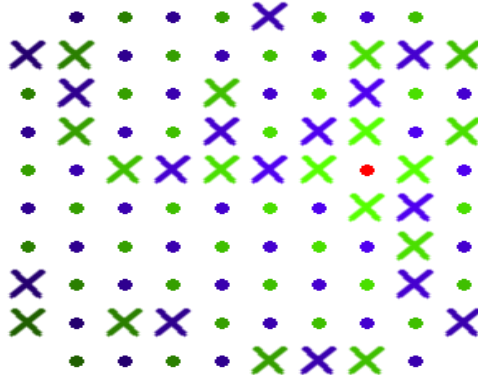


Fig. 2 An example of a D -set generated by the seed and shell method. The seed is marked in red. The vertices of the D -set are marked by “x”, while the other vertices are marked as dots. The layers are marked by alternating colors of green and blue, with the shade turning darker as one recedes from the seed.

the seed. We only select points on the equidistant shells that are in G . It is left to the reader to check that this process indeed generates a D -set. Figure 2 depicts an example of a D -set for G with $N = 10$ that was generated by the seed and shell algorithm. The seed is marked as a red dot. The x’s (points in the D -set) and dots (the other points in G) are marked in successive shells about the seed. The first shell is in green, the next one is in blue, and further shells are drawn in alternating shades of green and blue that get darker as the shell recedes from the seed. Notice that as N gets larger, the D -set becomes sparser, although there is of course some concentration of points near the seed. The reader might find it amusing to look for similar (or other) D -sets for arbitrary rectangular discrete domains or other graphs. To assist the reader in playing with the seed and shell algorithm in rectangular graphs, we wrote an interactive MATLAB code called Tryseed.m. The code can be downloaded from the website <http://php.indiana.edu/~jrubinst>. To run the code it is also necessary to download from the same website the files dot.bmp and ex.bmp. Alternatively one can install other bmp files to mark the D -set elements.

We proceed to analyze some general properties of D -sets. First, recall from Theorem 1 that the space of harmonic functions in G is of dimension $4N - 8$. Therefore, no set of points in G smaller than $4N - 8$ can uniquely determine a harmonic function. Next we argue that a D -set cannot contain a subset that is “too dense.” More precisely, we state that given any D -set A , there is no simply connected subdomain G_0 of G such that the number of points of A in G_0 is larger than the size of the boundary ∂G_0 ; for if there exists a subdomain contrary to the statement, we can replace the points in $A \cap G_0$ by the boundary points of G_0 . The new set A_0 clearly determines uniquely harmonic functions, but its size is smaller than $|B|$, which is a contradiction. One might think that this “sparseness” property of a D -set can be inverted, in the sense that any set of size $|B|$ that satisfies it is a D -set. Unfortunately such a conjecture is false, as can be seen through examples.

We mentioned before that a slight variation of the boundary B is also a D -set. This raises the general question of D -set deformations. We say that a set E_2 of points

in G is a deformation of another set E_1 in G if E_2 is obtained from E_1 by exchanging an arbitrary point of E_1 with one of its nearest neighbors. We prove the following theorem.

THEOREM 2. *Every D-set in G is obtained from B by a chain of deformations. In other words, the set of all D-sets is connected.*

Proof. We construct a basis for the space K . Let D be an arbitrary D-set. For each point $p_i \in D$ let T_i be the *unique* harmonic function that satisfies on D the conditions

$$(2.1) \quad T_i(p) = \begin{cases} 1 & \text{if } p = p_i, \\ 0 & \text{if } p \neq p_i. \end{cases}$$

It is easy to verify that the set of fundamental solutions $\{T_i\}$ thus defined forms a basis for K . Assume that $p_i \notin B$, and let \bar{p}_i be a point such that $\bar{p}_i \notin D$ and $T_i(\bar{p}_i) \neq 0$. Such a point must exist by the maximum principle. One can modify D by replacing p_i by \bar{p}_i , and further divide T_i by $T_i(\bar{p}_i)$ to obtain a new D-set with essentially the same set of fundamental solutions. We next observe that $\forall p \in D \cap I$ there exists an orbit l_p of points in G , connecting p to the boundary B , such that $T_p(x) \neq 0$ for all $x \in l_p$. To justify the last statement notice that if there is no such orbit l_p , then the point p is surrounded by a shell of points on which the associated fundamental solution T_p vanishes. But the maximum principle would then imply $T_p(p) = 0$, which contradicts (2.1). In particular, notice that no point along the orbit l_p belongs to D .

One can thus deform the original set D by finitely many local deformations until the point p_i is replaced by a boundary point. Repeating the process with respect to all the points in the original D that are not boundary points, one obtains that any D-set can be deformed into the boundary B . An example of a sequence of four deformations, along with the values of the associated T functions is given in Figure 3. The reader can watch a cartoon of a deformation chain from an initial D-set all the way to a boundary D-set on the website <http://php.indiana.edu/~jrubinst>.

3. Additional Topics. The analysis above can be extended in several ways. Some are obvious; for example, it is a simple matter to extend the analytical results to any dimension, or to extend the algorithm for constructing “distributed” D-sets to rectilinear domains in Z^2 .

We finish with several extensions of the notion of D-sets and their constructions, suggestions for further research, and with more words on the phase reconstruction problem.

D-Sets in Higher Dimensions. The seed and shell algorithm can also be performed in higher dimensions. We give here a brief outline of how this is done. The goal is to construct a relatively distributed D-set for the graph H_3^N consisting of all points $(i, j, k) \in Z^3$ such that $|i| + |j| + |k| \leq N$. Our algorithm extends the construction above to H_2^N . Again we work in equidistant shells about the origin. However, the selection of points on each shell is more involved now. Let us consider, for example, the part of the n th shell that lies in the first octant, $1 \leq i + j + k \leq n$. Assume by induction that appropriate points have already been selected for all shells $|i| + |j| + |k| = l$ with $l \leq n - 1$, such that these points form a D-set for the graph H_3^{l-1} . Let (i_0, j_0, k_0) be a point such that $i_0 + j_0 + k_0 = n - 1$. We then use repeatedly the principle that if a harmonic function u has determined values at $(i_0 + 1, j_0, k_0)$ and $(i_0, j_0 + 1, k_0)$, then its value is also determined at $(i_0, j_0, k_0 + 1)$ (and similarly

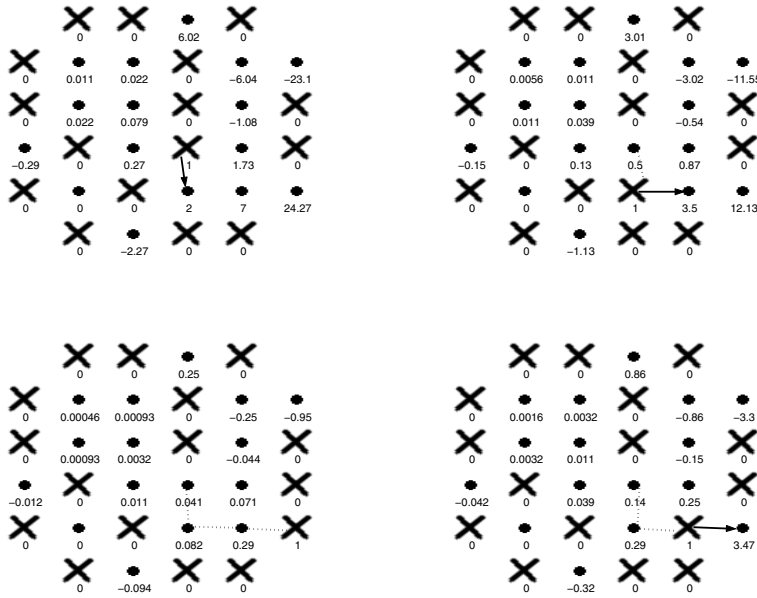


Fig. 3 A sequence of four D-sets connected through simple deformations. The initial set is on the top left corner, and the subsequent deformations are arranged clockwise. The numbers near each vertex are the values of the relevant fundamental solution T there (each value is written up to two significant digits), and the arrows indicate how a point in the D-set is moved.

for other permutations of these triplet). It is convenient to arrange the points on the triangle $|i| + |j| + |k| = n$ as shown in Figure 4. We select an initial point (the “seed”) for the D-set from this triangle. This point makes an inner triangle of “size” 1. We select an arbitrary neighbor of the seed and apply the principle above to imply that the values of u are determined by all the points in the inner triangle of size 2 thus created. We then proceed to fill the triangle $|i| + |j| + |k| = n$ by inner triangles of increasing size, each of them with edges parallel to the main triangle (as shown in Figure 4), where in each of them we select one point and apply the principle above to argue that u is determined in the rest of the respective inner triangle by the smaller inner triangles in $|i| + |j| + |k| = n$, and by the induction hypothesis on H_3^{n-1} . In this way we have generated a D-set for the graph H_3^n . Proceeding by induction we generate the full D-set for H_3^N .

How Many D-Sets Are There? An interesting question relates to the probability that a set is a D-set. Let the function $f(N)$ be the number of D-sets divided by the number of all sets in G of size $4N - 8$. We know that f decays fast as N grows, but we have not determined the rate of decay.

D-Functionals. The problem of solving (1.2) together with an additional $4(N - 2)$ conditions on u at an appropriate D-set is a special case of a more general question involving two systems of linear functionals. We shall formulate that more general problem, prove a statement about it, and then use the statement to derive an efficient method for determining whether a given set of points in G is a D-set.

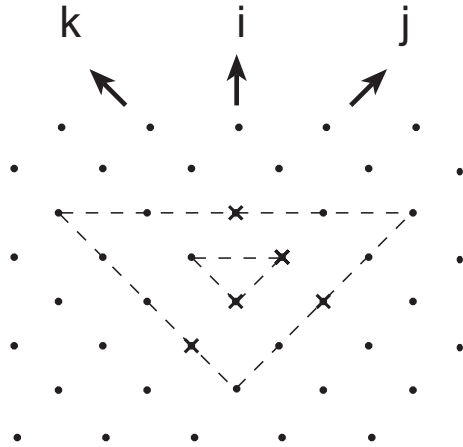


Fig. 4 A sketch for the construction of the D -set for the graph H_3^N . The sketch depicts points in the first octant of $|i| + |j| + |k| = n$. The points marked by “x” are selected into the D -set. The dashed lines form the triangles as explained in the text.

Consider for this purpose two sets of linear equations for a vector x in R^{P+Q} :

$$(3.1) \quad \sum_{j=1}^{P+Q} L_{ij}x_j = h_i, \quad i = 1, \dots, P,$$

$$(3.2) \quad \sum_{j=1}^{P+Q} M_{ij}x_j = h_i, \quad i = P + 1, \dots, P + Q.$$

Here P and Q are positive integers, and L and M are matrices. The matrix L is fixed, and the problem is to characterize all matrices M such that the joint system (3.1)–(3.2) has a unique solution. Such a matrix M will be called a D -functional, in analogy with the notion of a D -set defined earlier.

To make the problem meaningful, we assume that the first system (3.1) is consistent. By the linear algebra arguments that we used in the first section it suffices to prove the existence of a unique solution for the special case where $h_i = 0$ for $i = 1, 2, \dots, P$. Under this setup, the kernel of L is Q -dimensional. Let F_1, F_2, \dots, F_Q be a basis for this kernel, and let F be the matrix whose columns are the basis vectors. The solution of (3.1) can be expressed in the form $x = Fy$ for some vector y in R^Q . Substituting this expression for x in (3.2) gives

$$(3.3) \quad \sum_{j=1}^Q R_{ij}y_j = h_{P+i}, \quad i = 1, \dots, Q,$$

where the $Q \times Q$ matrix R is defined by $R = MF$. We have thus established the following theorem.

THEOREM 3. *A matrix M is a D -functional if and only if the matrix R is invertible.*

An immediate conclusion from Theorem 3 is that to check if a matrix M is a D -functional, it suffices to check the solvability of the $Q \times Q$ system defined by R

(once F is computed), which is much more efficient than checking the solvability of the original $(P + Q) \times (P + Q)$ system defined by (3.1)–(3.2).

D-sets are a special case of D-functionals. In fact, it is possible to use the theory above for D-functionals to define D-sets for any second order elliptic differential equation. Thus, the first P equations represented by the matrix L are the finite difference equations for the elliptic operator that apply to interior points. The points in the D-set candidate are associated with the remaining Q entries of x . In this case the entries of the matrix M are all zeros except for the $Q \times Q$ block of the entries $i, j = P + 1, \dots, P + Q$, where M is the identity matrix.

The boundary set B is a natural candidate to compute the kernel matrix F of the matrix L . We can therefore identify the fundamental vectors $\{F_i\}$ with the fundamental solutions $\{T_i\}$ introduced in (2.1). It follows that the matrix R consists of the values of the fundamental solutions T_i taken on the points of the D-set,

$$(3.4) \quad R_{ij} = T_{Q+i}(x_{P+j}), \quad i, j = 1, \dots, Q.$$

More on Phase Reconstruction. The problem of phase reconstruction was introduced as a motivation for defining the notion of D-sets. In practice, however, a D-set may not be the best way to design a hybrid phase sensor, i.e., a sensor that combines wavefront slope information from an HS sensor and intensity measurements. There are a number of reasons for this, and one has to do with stability. Although describing the phase over a D-set guarantees a unique solution of the Fresnel equation (1.5), the solution may be unstable to measurement errors. In other words, the linear system may be solvable, but its determinant may be very small and thus errors in the data might be amplified. One approach to overcome the problem is to provide the phase data for (1.5) at a set of points larger than a D-set. This might make the problem overdetermined, and therefore a least squares algorithm should be used to extract an optimal solution. The issue is, in fact, more subtle, since the slope and intensity data may come at different resolutions and may also be of different reliability. A stable method to combine slope and intensity information was recently proposed by Barbero, Rubinstein, and Thibos [1]. For simplicity of presentation the problem was formulated slightly differently than (1.5). Thus, consider the problem of estimating a surface $u(x, y)$ from information on its slopes and its Laplacian (and more generally its mean curvature). Denoting the slope measured data by $f(x, y)$ and the measured Laplacian data by $g(x, y)$, the authors introduced the following minimization problem:

$$(3.5) \quad \min \int_G (\Delta u - g)^2 + w |\nabla u - f|^2,$$

where $w(x, y)$ is a weight function that can be used to calibrate the reliability of the slopes versus Laplacian data. Finally, we point out that there are many other approaches to phase reconstruction. A survey of many old and recent methods can be found in [2].

Acknowledgments. We thank J. B. Keller and the reviewers for useful suggestions.

REFERENCES

- [1] S. BARBERO, J. RUBINSTEIN, AND L. THIBOS, *Wavefront sensing and reconstruction from gradient and Laplacian data measured with a Hartmann-Shack set-up*, Opt. Lett., 31 (2006), pp. 1845–1847.

- [2] T.E. GUREYEV, A. POGANI, D.M. PAGANIN, AND S.W. WILKINS, *Linear algorithms for phase retrieval in the Fresnel regime*, Opt. Comm., 231 (2004), pp. 53–70.
- [3] D. MALACARA, *Optical Shop Testing*, 2nd ed., Wiley, New York, 1992.
- [4] Y. PINCHOVER AND J. RUBINSTEIN, *An Introduction to Partial Differential Equations*, Cambridge University Press, Cambridge, UK, 2005.
- [5] F. RODDIER, *Curvature sensing and compensation: A new concept in adaptive optics*, Appl. Opt., 27 (1998), pp. 1223–1225.
- [6] J. RUBINSTEIN AND G. WOLANSKY, *A variational principle in optics*, J. Opt. Soc. Amer. A, 21 (2004), pp. 2164–2172.
- [7] M.R. TEAGUE, *Deterministic phase retrieval: A Green's function solution*, J. Opt. Soc. Amer., 73 (1983), pp. 1434–1441.