SIAM J. APPL. MATH.
Vol. 42, No. 4, August 1982

0036-1399/82/4204-0019 $01.00/0

RECONSTRUCTION FROM PARTIAL INFORMATION WITH APPLICATIONS TO TOMOGRAPHY*

CHARLES L. BYRNE† AND RAYMOND M. FITZGERALD‡

Abstract. The problem of reconstruction from projections in Hilbert space is treated. An axiomatic basis is considered which leads to techniques which provide improved reconstructions by incorporating prior knowledge to tailor the Hilbert space to the problem at hand. When applied to reconstruction of the Fourier transform of a function sampled at finitely many discrete points, the procedures lead to previously derived optimal estimation techniques. When applied to x-ray tomography, the procedures lead to new reconstruction techniques which are shown to include as a special case the minimum energy reconstruction of Logan and Shepp [Duke Math. J., 42 (1975), pp. 645–659].

1. Introduction. Techniques for the reconstruction of mathematical objects from partial information have a variety of important applications. The estimation of a Fourier transform from finitely many samples, the approximation of linear attenuation functions from x-ray data, the extrapolation of time-series beyond the observed values and the reconstruction and enhancement of images in radio astronomy and picture processing share a mathematical foundation that is best understood within the context of Hilbert space approximation theory.

The problems we want to solve are typically ill-posed; the finite data we have gathered do not uniquely specify the function to be reconstructed. For this reason the notion of reconstruction is itself problematic, and we begin in § 2 with a discussion of various properties any reconstruction procedure should exhibit. This axiomatic analysis of reconstruction procedures suggests certain basic principles to be followed in the design of reconstruction techniques. In § 3 we apply these principles to the problem of estimating the Fourier transform of a function from finitely many samples. Our intent here is to illustrate the content of the axioms previously developed as well as to prepare for the tomographic reconstruction algorithms to be discussed in § 4. In the final § 5 we compare tomographic algorithms due to Logan and Shepp with those of § 4.

2. Reconstructions in Hilbert spaces. Suppose that f is an unknown member of a Hilbert space H and that we have only incomplete information about f, namely,

$$x_j = \langle f, g_j \rangle, \qquad j = 1, 2, \cdots, N, \tag{1}$$

where the g_j, $j = 1, \cdots, N$ are known linearly independent members of H and $\langle \cdot, \cdot \rangle$ denotes the inner product. We wish to reconstruct f from the data (1). So that we may better understand how this may be accomplished, we consider five properties of a reconstruction procedure and analyze the consequences of using these properties to give an axiomatic statement of the meaning of reconstruction. We begin by assuming only that a reconstruction method R is a function from H to itself, with $R(f)$ denoting the reconstruction that results when the unknown element is f.

The first property is (a) *linearity*:

$$R(\alpha f + \beta h) = \alpha R(f) + \beta R(h), \tag{2}$$

* Received by the editors January 16, 1981.

† Department of Mathematics, The Catholic University of America, Washington, DC 20064 and Code 5120, U.S. Naval Research Laboratory, Washington, DC 20375.

‡ Code 5120, U.S. Naval Research Laboratory, Washington, DC 20375.

for any scalars α and β and elements f and h of H. The second is (b) *continuity*:

$$\lim_{\|h\|\to 0} R(f+h)=R(f), \tag{3}$$

where $\|\cdot\|$ denotes the norm of H. These two properties, although not absolutely essential, greatly simplify error analysis and make available the elegant machinery of Hilbert space. The third property is (c) *computability*: for each f,

$$R(f)=R(P_G(f)), \tag{4}$$

where $G=\operatorname{span}\{g_1,\cdots,g_n\}$ and P_G denotes the orthogonal projection onto G. Property (c) is necessary if $R(f)$ is to be computed from the data we have. The fourth property is (d) *consistency*: for each f we have

$$\langle R(f), g_j\rangle=\langle f, g_j\rangle, \qquad j=1,2,\cdots,N. \tag{5}$$

Property (d) is not essential either, but, unless one has strong reasons for not assuming that the reconstructed element shares with f the data that have been observed, it is the common practice to accept (d).

Careful consideration of the four axioms above reveals that they have no approximation-theoretic content. Although it does follow from (a), (b), (c) and (d) that R is a projection, that is $R(R(f))=R(f)$, it does not follow that R is an orthogonal projection, so that $R(f)$ need not be the closest element to f in the range of R. Indeed, $R(f)$ can be a considerable distance from f. To remedy this situation we need a fifth property, (e) *optimality*: for any f and h in H, we have

$$\|f-R(f)\|\leqq\|f-R(h)\|. \tag{6}$$

It follows from (a), (b), (c), (d) and (e) that R is the orthogonal projection onto its range $R(H)$. We show that actually $R(H)=G$. From (c) it follows that the dimension of $R(H)$ is at most N, while from (d) we can show that the dimension of $R(H)$ is at least N. Therefore, it must be N. Note that if $P_G(h)=0$, then from (c) it follows that $R(h)=0$. Consequently $R(H)=G$: if m is in $R(H)$ but not in G, then $m-P_Gm$ is not 0 and $m-P_Gm$ is orthogonal to G, hence to $R(H)$. But then

$$\begin{aligned}\|m-P_Gm\|^2&=\langle m-P_Gm, m-P_Gm\rangle\\ &=\langle m, m-P_Gm\rangle-\langle P_Gm, m-P_Gm\rangle=0-0,\end{aligned} \tag{7}$$

so $m-P_Gm=0$.

Summarizing, we have shown that any method of reconstruction that satisfies all five of the properties discussed must reduce to the orthogonal projection onto the range of R, which must be G. Therefore, all such reconstructions will have the form

$$R(f)=\sum_{j=1}^{N} a_j g_j, \tag{8}$$

with the a_j, $j=1,\cdots,N$, chosen so as to satisfy (5). We have the option of discarding any or all of the five properties. However, a more reasonable alternative exists if we are unhappy with reconstructions of the form (8).

In practice, the object to be reconstructed is usually a function of one or more variables. The Hilbert space H in which we choose to place this function is largely up to us. Certain well-known Hilbert spaces, such as $L^2(-\infty,\infty)$, are frequently used. It is the selection of H that affords us the most flexibility in designing reconstruction techniques and at the same time the opportunity to tailor H to the problem at hand so that the reconstruction, whose general form is given in (8), will share broad features of f such as we know them a priori.

In order to illustrate the content of the five properties and the principles of reconstruction design suggested by them, in particular, the importance of the careful selection of the ambient space H, we present in the next section an application of these ideas to the problem of estimating the Fourier transform of a function from finitely many sample values.

3. Fourier transform reconstruction. Suppose we have observed the function $x(t)$ at the values $t=t_1, \cdots, t_N$, and on the basis of this data we wish to estimate the Fourier transform of $x(t)$,

$$X(\omega)=\int_{-\infty}^{\infty} x(t)\, e^{i\omega t}\, dt. \tag{9}$$

The data are insufficient to determine $X(\omega)$ uniquely. Our success in reconstructing $X(\omega)$ will depend upon our ability to incorporate prior knowledge about $X(\omega)$ into the reconstruction method. As an illustration, let us reconstruct $X(\omega)$, assuming $X(\omega)=0$ for $|\omega|>\sigma$, where $\sigma>0$ is some known constant. In this case we say that $x(t)$ is σ-band-limited.

Let us take $X(\omega)$ to be a member of the Hilbert space $H=L^2(-\sigma, \sigma)$. Then

$$x(t_j)=\int_{-\sigma}^{\sigma} X(\omega)\, e^{-i\omega t_j}\, d\omega/2\pi = \langle X(\omega), g_j(\omega)\rangle, \tag{10}$$

where $g_j(\omega)=e^{i\omega t_j}$ for $|\omega|\leqq\sigma$ and 0 for $|\omega|>\sigma$. According to the discussion in the previous section, we should take as our reconstruction of $X(\omega)$,

$$R(X(\omega))=\begin{cases}\sum_{j=1}^{N} a_j\, e^{i\omega t_j}, & |\omega|\leqq\sigma,\\ 0 & \text{otherwise}\end{cases} \tag{11}$$

where the $a_1, \cdots, a_N$ are chosen so that $R(X(\omega))$ is consistent with the data. We then have

$$x(t_k)=\sum_{j=1}^{N} a_j \frac{\sin (t_k-t_j)\sigma}{\pi(t_k-t_j)} \tag{12}$$

for $k=1, 2, \cdots, N$, which must then be solved for $a_1 \cdots a_N$. By property (e) this choice of coefficients minimizes the expression

$$\int_{-\sigma}^{\sigma} |X(\omega)-\sum_{j=1}^{N} a_j\, e^{i\omega t_j}|^2\, d\omega. \tag{13}$$

The estimate $R(X(\omega))$ is the best approximation of $X(\omega)$ of its particular form, in the sense of the $L^2(-\sigma, \sigma)$ norm (13). In [4] it was shown that this estimation procedure is implicit in a number of recently published techniques for band-limited extrapolation (Cadzow [3], Papoulis [7], Kolba and Parks [5]). It also plays an important role in the derivation of time-limited sampling theorems for band-limited signals [1].

Suppose we have additional prior knowledge about $X(\omega)$, such as relative energy concentrations in various bands or that $X(\omega)$ has a noise or clutter component of known statistical form. Let us take $P(\omega)\geqq 0$ to embody these broad features of $X(\omega)$, such as we know them a priori, taking care that $P(\omega)=0$ only if $X(\omega)=0$. Then consider the Hilbert space $H = L^2(-\infty, \infty; 1/P(\omega))$, whose inner product is given by

$$\langle f(\omega), g(\omega)\rangle_P = \int_{-\infty}^{\infty} f(\omega)\overline{g(\omega)}/P(\omega)\, d\omega/2\pi, \tag{14}$$

where the integral is understood to be taken over that part of the line where $P(\omega)$ is not zero. Viewing $X(\omega)$ as an element of this H, we have for $j = 1, \cdots, N$,

$$\begin{aligned} x(t_j) &= \int_{-\infty}^{\infty} X(\omega)\, e^{-i\omega t_j}\, d\omega/2\pi \\ &= \int_{-\infty}^{\infty} X(\omega)[e^{-i\omega t_j}P(\omega)]/P(\omega)\, d\omega/2\pi = \langle X(\omega), e^{i\omega t_j}P(\omega)\rangle_P. \end{aligned} \tag{15}$$

Our reconstruction of $X(\omega)$ must then take the form

$$R(X(\omega)) = \sum_{j=1}^{N} a_j g_j(\omega) = \sum_{j=1}^{N} a_j\, e^{i\omega t_j}P(\omega), \tag{16}$$

and the choice of coefficients is such as to minimize the expression

$$\int_{-\infty}^{\infty} \Big|X(\omega) - \sum_{j=1}^{N} a_j\, e^{i\omega t_j}P(\omega)\Big|^2/P(\omega)\, d\omega/2\pi. \tag{17}$$

The system of equations to be solved is then

$$x(t_k) = \sum_{j=1}^{N} a_j g(t_k - t_j), \qquad k = 1, 2, \cdots, N,$$

where

$$g(t) = \int_{-\infty}^{\infty} P(\omega)\, e^{-i\omega t}\, d\omega/2\pi.$$

Two points are worth noting here. The first is that because $P(\omega)$ has been selected to embody broad features of $X(\omega)$, so will $R(X(\omega))$. The second is that the inverse weighting of the approximation error in (17) makes it possible to increase the sensitivity of the estimator to errors made in approximating small values of $X(\omega)$. This increased sensitivity to the less prominent features will be of special importance in tomographic reconstructions based on these estimators.

In the example discussed earlier, our knowledge that $X(\omega) = 0$ for $|\omega| > \sigma$ justified our use of the space $L^2(-\sigma, \sigma)$. In fact, the same reconstruction could have been obtained had we set $P(\omega) = 1$ for $|\omega| \leqq \sigma$, 0 for $|\omega| > \sigma$ and proceeded according to (16), (17). A detailed discussion of the Fourier transform estimation methods is presented in [2] along with recursive algorithms for the solution of the linear systems involved.

4. Tomographic reconstruction. The basic problem of x-ray tomography is the reconstruction of a cross-section of density distribution (or linear attenuation function) from a finite set of x-ray data. For concreteness let $D = \{(x, y) | x^2 + y^2 \leqq 1\}$, and suppose that for any chord C of D the drop in x-ray intensity I is given by

$$I(\text{final}) = I(\text{initial}) \exp\left(-\int_C f(x, y)\, ds\right), \tag{18}$$

where the integral is the line integral along C. For every point (t, θ), $|t| \leqq 1$, $0 \leqq \theta < \pi$, in D (in polar coordinates) we associate the chord $C(t, \theta)$ having (t, θ) as its midpoint.

This chord then is $\{(x, y)|x \cos \theta + y \sin \theta = t\}$. Parametrically, this chord is given by

(19) $$C(t, \theta) = \{(t \cos \theta - u \sin \theta, t \sin \theta + u \cos \theta) \ |u| \leqq \sqrt{1-t^2}\}.$$

The Radon transform of $f(x, y)$ is defined by

(20) $$P_f(t, \theta) = \int_{u=-\sqrt{1-t^2}}^{+\sqrt{1-t^2}} f(t \cos \theta - u \sin \theta, t \sin \theta + u \cos \theta)\, du,$$

and assigns to each chord $C(t, \theta)$ the integral of $f(x, y)$ along that chord. We see then that our data can be taken to be finitely many values of the Radon transform

(21) $$P_f(t_j, \theta_j), \qquad j = 1, 2, \cdots, N,$$

where the t_j are not necessarily distinct, nor are the θ_j; only the pairs (t_j, θ_j) are distinct.

We define the two-dimensional Fourier transform of $f(x, y)$ as

(22) $$\tilde{f}(r, s) = \iint_D f(x, y)\, e^{ixr+iys}\, dx\, dy.$$

The Fourier transform of $P_f(t, \theta)$ treated as a function of t only is

(23) $$\tilde{P}_f(\omega, \theta) = \int_{-1}^{1} P_f(t, \theta)\, e^{i\omega t}\, dt.$$

It is easily seen that for $-\infty < \omega < \infty$, $0 \leqq \theta < \pi$,

(24) $$\tilde{P}_f(\omega, \theta) = \tilde{f}(\omega \cos \theta, \omega \sin \theta).$$

The problem of reconstructing $f(x, y)$ from the data (21) is not precisely the same as the problem treated in the last section. We do not have values of $\tilde{f}(r, s)$ on which to base our reconstruction. Two approaches seem reasonable. The first involves using the data to estimate $\tilde{P}_f(\omega, \theta)$, along the lines of § 3. From (24) we then have values of the Fourier transform of f, from which we compute $f(x, y)$ itself. The second approach does not use § 3 directly but rather the ideas of § 2. We consider $f(x, y)$ as an element of an appropriate Hilbert space and view the data as finitely many inner product values.

In order to use the first approach it is necessary that the chords $C(t_j, \theta_j)$ be in parallel groups, that is, our data is

(25) $$P_f(t_{j,k}, \theta_j), \qquad k = 1, \cdots, K_j, \quad j = 1, \cdots, M,$$

so that there are M distinct directions $\theta_1, \cdots, \theta_M$, corresponding to which there are parallel chords at the t-values $t_{j,1}, t_{j,2}, \cdots, t_{j,K_j}$. Using the data

(26) $$P_f(t_{j,k}, \theta_j), \qquad k = 1, \cdots, K_j,$$

to estimate $\tilde{P}_f(\omega, \theta_j)$ for each $j = 1, \cdots, M$, we then obtain values of $\tilde{f}(\omega \cos \theta, \omega \sin \theta)$ corresponding to each $\theta = \theta_j$. The problem is then reduced to that of reconstructing $f(x, y)$ from values of its Fourier transform. This problem is analogous to the one-dimensional problem discussed earlier, and we shall omit the details.

The second method employs the ideas of § 2 directly. We consider $f(x, y)$ as a member of a Hilbert space H to be specified below. Once we find elements $e_j(x, y)$ in H with, for $j = 1, \cdots, N$,

(27) $$P_f(t_j, \theta_j) = \langle f(x, y), e_j(x, y) \rangle,$$

we then have the form of our reconstruction:

$$R(f(x, y)) = \sum_{j=1}^{N} a_j e_j(x, y), \tag{28}$$

where the a_j, $j = 1, \cdots, N$ are chosen so as to make $R(f(x, y))$ consistent with the data (21). Which Hilbert space H we choose will affect the functional form of the $e_j(x, y)$, so it must be chosen with some care.

Suppose $f(x, y) \geqq 0$ on D and we are able to construct a prior estimate $h(x, y) > 0$ of $f(x, y)$, based on our knowledge of what the object is that is being x-rayed, what typical features it must have, etc., or perhaps based on a prior tomographic reconstruction from the same data. We then employ $h(x, y)$ the way $P(\omega)$ was used in the last section. Define

$$w(x, y) = \begin{cases} 1/h(x, y) & \text{if } h(x, y) \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

We shall assume that $f(x, y) = 0$, if $h(x, y) = 0$. Consider the space $L^2(D; w(x, y))$ with inner product

$$\langle f(x, y), g(x, y)\rangle = \iint_D f(x, y)\overline{g(x, y)}w(x, y)\, dx\, dy. \tag{30}$$

We need to place some restrictions on $f(x, y)$ in order to find $e_j(x, y)$. We require that the function $P_f(t, \theta_j)$ be σ_j-band-limited for each $j = 1, \cdots, N$. What the values of the $\sigma_1, \cdots, \sigma_N$ are is unimportant for what follows. These values are free parameters that can be altered to improve the performance of computer implementation. With these assumptions it can be easily checked that for $j = 1, \cdots, N$,

$$e_j(x, y) = \frac{\sin (t_j - x \cos \theta_j - y \sin \theta_j)\sigma_j}{\pi(t_j - x \cos \theta_j - y \sin \theta_j)} h(x, y), \tag{31}$$

and so our reconstruction has the form

$$R(f(x, y)) = h(x, y) \sum_{j=1}^{N} a_j \frac{\sin (t_j - x \cos \theta_j - y \sin \theta_j)\sigma_j}{\pi(t_j - x \cos \theta_j - y \sin \theta_j)} \tag{32}$$

where the $a_1, \cdots, a_N$ are chosen so as to make $R(f(x, y))$ satisfy the data (21).

It follows from our discussion in § 2 that these coefficients minimize the expression

$$\iint_D \left|f(x, y) - \sum_{j=1}^{N} a_j e_j(x, y)\right|^2 / h(x, y)\, dx\, dy. \tag{33}$$

If we know a priori that $f(x, y)$ takes on large values in certain places (for example, bone in the x-ray of a head) we can build that into $h(x, y)$, thereby making the approximation error with respect to the smaller values of $f(x, y)$ more significant than it would be if we used say $L^2(D)$ for our H.

Returning to the first method mentioned earlier, if our chords do lie in parallel groups and we choose to use the methods of § 3 directly, we still use our prior $h(x, y)$. Indeed, from $h(x, y)$ we get $\tilde{h}(r, s)$, its Fourier transform, and then from (24) we have for $-\infty < \omega < \infty$, $0 \leqq \theta < \pi$,

$$\tilde{h}(\omega \cos \theta, \omega \sin \theta) = \tilde{P}_h(\omega, \theta). \tag{34}$$

For each fixed direction θ_j, we could employ $|P_h(\omega, \theta_j)|$ as the $P(\omega)$ required to perform the Fourier transform estimation.

The first method for tomographic reconstruction, although it requires parallel groups of chords, has the advantage that the Fourier transform estimates needed for each fixed θ_j are calculated independently of all data not associated with θ_j. The second method has the advantage that it places no restriction whatsoever on the pairs (t_j, θ_j) and so can easily be used with the faster fan-beam method.

5. Relation to a method of Logan and Shepp. In [6] a technique for tomographic reconstruction is given for the case in which $P_f(t, \theta_j)$ is known for all t and for $j = 1, \cdots, N$. This is, of course, equivalent to assuming $\tilde{P}_f(\omega, \theta_j)$ is known for all ω for each $j = 1, \cdots, N$. If we take $H = L^2(D)$, then

(35) $$\tilde{P}_f(\omega, \theta_j) = \iint_D f(x, y) e_{j,\omega}(x, y)\, dx\, dy,$$

where for $j = 1, \cdots, N$,

(36) $$e_{j,\omega}(x, y) = \exp(i\omega(x \cos \theta_j + y \sin \theta_j)).$$

According to § 2 again our best reconstruction in H would be

(37) $$R(f(x, y)) = \sum_{j,\omega} a_{j,\omega} e_{j,\omega}(x, y),$$

where the sum is over $j = 1, \cdots, N$ and all ω. More properly, we should write

(38) $$R(f(x, y)) = \sum_{j=1}^{N} \int_{-\infty}^{\infty} a_j(\omega) e_{j,\omega}(x, y)\, d\omega,$$

where the $a_j(\omega)$ are chosen so as to satisfy the data constraints. In [6] Logan and Shepp discuss so-called "ridge functions" $p_j(x, y)$ of the form

(39) $$p_j(x, y) = p_j(x \cos \theta_j + y \sin \theta_j)$$

and show that the minimum energy reconstruction of $f(x, y)$ has the form

(40) $$f_0(x, y) = \sum_{j=1}^{N} p_j(x, y).$$

In fact, $f_0(x, y) = R(f(x, y))$ of (38), and

(41) $$p_j(x, y) = \int_{-\infty}^{\infty} a_j(\omega) e_{j,\omega}(x, y)\, d\omega$$

for each $j = 1, \cdots, N$.

Our approach has been to assume only finite data and to seek optimal resolution in the reconstruction by tailoring the ambient Hilbert space to the problem at hand rather than relying on larger data sets. When applied to the hypothetical case of continuous data in N directions these techniques provide the same reconstruction as [6].

REFERENCES

[1] C. Byrne and R. Fitzgerald, *Time-limited sampling theorems for band-limited functions*, submitted to IEEE Trans. Inform. Theory.

[2] ———, *Fourier transformation estimation with prior knowledge*, submitted to IEEE Trans. Acoust., Speech, Signal Proc.

[3] J. A. CADZOW, *An extrapolation procedure for band-limited signals*, IEEE Trans. Acoust., Speech, Signal Proc., 27 (1979), pp. 4–12.

[4] R. FITZGERALD AND C. BYRNE, *Extrapolation of band-limited signals: A tutorial*, in Signal Processing: Theories and Applications, M. Kunt and F. de Coulon, eds., North-Holland, New York, 1980, pp. 175–180.

[5] D. P. KOLBA AND T. W. PARKS, *Extrapolation and spectral estimation for band-limited signals*, IEEE International Conference on Acoustics, Speech, and Signal Processing, April 1979, pp. 194–199.

[6] B. F. LOGAN AND L. A. SHEPP, *Optimal reconstruction of a function from its projections*, Duke Math. J., 42 (1975), pp. 645–659.

[7] A. PAPOULIS, *A new algorithm in spectral analysis and band-limited extrapolation*, IEEE Trans. Comput. Syst., 22 (1975), pp. 735–742.