INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **17** (2001) 1307–1320 PII: S0266-5611(01)23556-4

Analytic continuation, singular-value expansions, and Kramers–Kronig analysis

A Dienstfrey[1,3] **and L Greengard**[2]

[1] NIST, Boulder, CO, USA
[2] Courant Insitute, New York, NY, USA

E-mail: andrewd@boulder.nist.gov

Received 3 April 2001
Published 30 August 2001
Online at stacks.iop.org/IP/17/1307

Abstract

We describe a systematic approach to the recovery of a function analytic in the upper half-plane, C^+, from measurements over a finite interval on the real axis, $D \subset R$. Analytic continuation problems of this type are well known to be ill-posed. Thus, the best one can hope for is a simple, linear procedure which exposes this underlying difficulty and solves the problem in a least-squares sense. To accomplish this, we first construct an explicit analytic approximation of the desired function and recast the continuation problem in terms of a 'residual function' defined on the measurement window D itself. The resulting procedure is robust in the presence of noise, and we demonstrate its performance with some numerical experiments.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

In many areas of engineering and applied physics, one encounters time-invariant, causal linear systems, with input $H(t)$, output $S(t)$, and transfer function $F(t)$. Such systems are characterized by the relation

$$S(t) = \int_{-\infty}^{t} F(t-\tau)H(\tau)\,\mathrm{d}\tau. \tag{1}$$

If we assume that $F \in L^2(R^+)$, we can investigate the Fourier transform

$$f(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} F(t)\mathrm{e}^{\mathrm{i}\omega t}\,\mathrm{d}t. \tag{2}$$

[3] Contribution of US Government, not subject to copyright.

By the Paley–Wiener theorem [DM], $f(\omega)$ is analytic for $\omega \in \boldsymbol{C}^+ = \{\omega |\ \mathrm{Im}\,(\omega) > 0\}$ and satisfies the integrability property

$$\sup_{y>0} \left\{ \int_{-\infty}^{\infty} |f(x+\mathrm{i}y)|^2 \,\mathrm{d}x \right\} = \int_{-\infty}^{\infty} |f(x)|^2 \,\mathrm{d}x < \infty. \tag{3}$$

Functions analytic in the upper half plane, and satisfying (3) are referred to as Hardy functions ($f \in \boldsymbol{H}^2(\boldsymbol{R})$). The converse also holds: all Hardy functions may be obtained as Fourier transforms of functions supported on $(0, \infty)$.

It is well known that the real and imaginary parts of Hardy functions are not independent of one another. Rather, applying a limiting procedure to the Cauchy integral representation for analytic functions, one obtains the following integral identities.

Lemma 1 (Kramers–Kronig relations). *Let* $f(\omega) = f_{\mathrm{r}}(\omega) + i f_{\mathrm{i}}(\omega) \in \boldsymbol{H}^2(\boldsymbol{R})$. *Then*

$$\begin{aligned} f_{\mathrm{r}}(\omega) &= -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{\omega - s} f_{\mathrm{i}}(s) \,\mathrm{d}s, \\ f_{\mathrm{i}}(\omega) &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{\omega - s} f_{\mathrm{r}}(s) \,\mathrm{d}s. \end{aligned} \tag{4}$$

Furthermore, the 'time domain' function $F(t)$ in (2) is oftentimes real-valued. In this case, the real and imaginary parts of $f(\omega)$ are respectively even and odd. Incorporating these symmetries into (4) gives [Kra]

$$\begin{aligned} f_{\mathrm{r}}(\omega) &= \frac{2}{\pi} \int_{0}^{\infty} \frac{s}{s^2 - \omega^2} f_{\mathrm{i}}(s) \,\mathrm{d}s, \\ f_{\mathrm{i}}(\omega) &= -\frac{2\omega}{\pi} \int_{0}^{\infty} \frac{1}{s^2 - \omega^2} f_{\mathrm{r}}(s) \,\mathrm{d}s. \end{aligned} \tag{5}$$

Remark 2. The pairs of equations are redundant for, in both (4) and (5), one line implies the other. For a given $\omega \in \boldsymbol{R}$, we refer to the 'local' part of the operator as the integration region where $s \approx \omega$; the rest of the integral we refer to as the 'far-field' contribution.

In applications, the desire to use analyticity in the form of the Kramers–Kronig relations is pervasive. The universal problem in practice is that it is possible to take frequency measurements only over a *finite* data window, $\omega \in D = (\Omega_{\min}, \Omega_{\max})$, while the integrals in (4) and (5) are global, $\omega \in \boldsymbol{R}$. The specific problems that one encounters depend on the data available through experiment:

(1) In some situations it is possible to measure both real and imaginary parts of the complex, causal function $f(\omega) = f_{\mathrm{r}}(\omega) + \mathrm{i}\, f_{\mathrm{i}}(\omega)$. In this case, the natural problems are
   (a) to determine to what extent the measured functions satisfy the Kramers–Kronig relations, i.e. to what degree these relations may be viewed as a constraint;
   (b) to extrapolate the given data to frequencies outside D (or the entire real line) in such a way that the extension is causal and matches the measured functions over the data window.

(2) In other experiments only the real part of the function f_{r} is available. Here, one would like to apply causality in the form of the Kramers–Kronig relations to determine the imaginary part f_{i}. A common example is phase determination (see section 3.2).

We refer to the first type of problem as 'analytic continuation' and the second as 'analytic interpolation'. As mentioned above, both problems are complicated primarily by the fact that the frequency-measurement window is finite. There are two competing theorems here. First, it is well known that analytic continuation is unique. On the other hand, it is a theorem due

to Riesz that the restriction of Hardy functions to any finite interval is dense in L^2 [Par, DM]. Thus, problems 1(a) and 2 are senseless. Problem 1(b) fares only slightly better: it is one of the classical ill-posed problems of mathematical physics. Hence, one must seek solutions to these problems that *parametrize* the reconstruction in some systematic fashion.

Informally speaking, the reason for some hope is that the Hilbert transform is dominated by its local part where the kernel in (4) is singular. In the far-field the kernel is smoothly decaying and, as is straightforward to show, the contribution from distant values of f to the local measurements is not only differentiable, but analytic. In other words, although it is theoretically possible for features outside the data window to have arbitrary influence on the interior, the exertion of this influence requires a great deal of L^2-energy and structure to exist in the far field (the 'tail regions'). In the literature, this idea tends to exist in the form of heuristically chosen regularizations.

In the present paper, we present a systematic analysis of the problem through the use of a singular-value analysis of operators related to (5). By extending the function from the data window smoothly (but not necessarily analytically), we are able to formulate a 'residual problem' which provides

(1) a natural choice of subspaces in which to solve the continuation problem, and
(2) a simple measure of the difficulty of extension.

In section 2, we briefly describe the singular-value expansion for compact operators. Section 3 comprises the bulk of the paper. Here we derive the residual problems associated with the analytic continuation and interpolation problems presented above, describe the constraints that we employ to guarantee the compactness of these residual problems, and discuss some possible discretization techniques. We follow this with numerical examples in section 4. Section 5 contains some concluding remarks, including a discussion of our procedure in relation to existing approaches.

2. Singular-value expansions

The following theorem is a standard result in functional analysis [Con].

Theorem 3 (Singular-value expansion (SVE)). *Let X, Y be two infinite-dimensional Hilbert spaces and $K : X \to Y$ be a compact operator. Then there exists a sequence, $\{u_n, v_n, \lambda_n\} \in X \times Y \times \boldsymbol{R}^+$, such that $\lambda_n \to 0$ decreasingly and*

- *$\{v_n\}$ forms a basis for $N(K)^\perp \subset X$, the orthogonal complement of the null-space of K.*
- *$\{u_n\}$ forms a basis for $\overline{R(K)} \subset Y$, the closure of the range in Y.*
- *For all $f \in X$, $Kf = \sum_1^\infty \lambda_n u_n (f, v_n)_X$.*

The functions $\{u_n, v_n\}$ are referred to as the left and right singular functions respectively, and $\{\lambda_n\}$ as the corresponding singular values.

The above representation for the action of K parallels the singular-value decomposition (SVD) of matrices, $\boldsymbol{K} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^*$. The use of the SVD to solve discrete, least-squares algorithms is well known and an analogous procedure may be employed in the compact (i.e. continuous) case. Our goal is to bring this analysis to bear on appropriately defined problems stemming from the Kramers–Kronig relations (5).

Once the SVE has been obtained for a compact operator K, one may form its pseudo-inverse, $K^\dagger$, mapping $R(K) \cup R(K)^\perp \to N(K)^\perp$. More precisely, the solution to $Kf = g$ is

$$f = K^{\dagger} g = \sum_{n=1}^{\infty} \left(\frac{(u_n, g)_Y}{\lambda_n} \right) v_n. \tag{6}$$

Since the singular values decay to zero, it is clear that the pseudo-inverse is unbounded on $\overline{R(K)}$. The condition number of a finite-dimensional, truncated approximation of K is defined to be the ratio λ_1/λ_N with pseudo-inverse given by the preceding series truncated after N terms. Assuming measurement errors of the order ϵ, it is sensible to choose N so that $\lambda_1/\lambda_N \approx \sqrt{1/\epsilon}$. This limits the amplification of error and provides a definition for the 'effective rank' of the operator K (with specified precision, ϵ) [GRV].

3. The residual problem

In this section, we recast the analytic continuation and interpolation problems from the introduction into a 'residual' form. We will see that the residual problems are naturally posed as simple Fredholm integral equations of the first kind. For ease of presentation we consider the case where the measurement window takes the form $D = (0, \Omega_{\max})$. Recognizing that equations (5) are homogeneous in the frequency variable, we may rescale D to be the interval $(0, 1)$. One sometimes encounters a non-zero, low-frequency cutoff, so that the data window is $(\Omega_{\min}, \Omega_{\max})$. In optical experiments, for example, only a portion of the electromagnetic spectrum is accessible. Generalizing the procedure described below to this and more general data regions is straightforward.

3.1. Analytic continuation

We assume that both the real and imaginary parts of an analytic (i.e. Hardy) function are available in the form $f(\omega) = f_{\mathrm{r}}(\omega) + \mathrm{i} f_{\mathrm{i}}(\omega)$, $\omega \in D$. We first construct *separate* extensions of the real and imaginary parts to the entire real line so that the extended functions agree with the measured data on $D = (0, 1)$ and have sufficient decay at infinity:

$$\begin{cases} f_{\mathrm{r}}(\omega) \\ f_{\mathrm{i}}(\omega) \end{cases}, \qquad \omega \in (0, 1) \rightarrow \begin{cases} f_{\mathrm{r}}^{\mathrm{ext}}(\omega) \\ f_{\mathrm{i}}^{\mathrm{ext}}(\omega) \end{cases}, \qquad \omega \in (0, \infty). \tag{7}$$

The exact manner in which this is done is not essential. In the computations presented here, we define

$$\begin{aligned} f_{\mathrm{r}}^{\mathrm{ext}} &= (a_0 + a_1(\omega - 1) + a_2(\omega - 1)^2)\mathrm{e}^{-\alpha(\omega-1)} \\ f_{\mathrm{i}}^{\mathrm{ext}} &= (b_0 + b_1(\omega - 1) + b_2(\omega - 1)^2)\mathrm{e}^{-\alpha(\omega-1)} \end{aligned}$$

for $\omega \in [1, \infty)$. The value $\alpha > 0$ is selected by the user, and the coefficients a_{i}, b_{i} are chosen to match the first two derivatives:

$$\begin{aligned} a_0 &= f_{\mathrm{r}}(1), & a_1 - \alpha\, a_0 &= f_{\mathrm{r}}'(1), & 2a_2 - 2\alpha\, a_1 + \alpha^2 a_0 &= f_{\mathrm{r}}''(1) \\ b_0 &= f_{\mathrm{i}}(1), & b_1 - \alpha\, b_0 &= f_{\mathrm{i}}'(1), & 2b_2 - 2\alpha\, b_1 + \alpha^2 b_0 &= f_{\mathrm{i}}''(1). \end{aligned}$$

Thus, the functions, $(f_{\mathrm{r}}^{\mathrm{ext}}, f_{\mathrm{i}}^{\mathrm{ext}})$ are in $C^2(R)$ and exponentially decaying.

Next, employing the Kramers–Kronig relations (5), we compute Hilbert transforms and form complex functions that we know to be in $\boldsymbol{H}^2$

$$\begin{aligned} f^{(A)}(\omega) &= f_{\mathrm{r}}^{\mathrm{ext}}(\omega) - \mathrm{i}\frac{2\omega}{\pi} \int_0^{\infty} \frac{f_{\mathrm{r}}^{\mathrm{ext}}(s)}{s^2 - \omega^2}\, \mathrm{d}s, \\ f^{(B)}(\omega) &= \frac{2}{\pi} \int_0^{\infty} \frac{s f_{\mathrm{i}}^{\mathrm{ext}}(s)}{s^2 - \omega^2}\, \mathrm{d}s + \mathrm{i} f_{\mathrm{i}}^{\mathrm{ext}}(\omega). \end{aligned}$$

Both $f^{(A)}(\omega)$ and $f^{(B)}(\omega)$ can be viewed as initial approximations of the measured function $f(\omega)$.

Definition 4. *The residual $R(\omega)$ for $\omega \in (0, 1)$ is given by the difference between the measurement and the average of these two causal extensions:*

$$R(\omega) = f(\omega) - \left(\frac{f^{(A)}(\omega) + f^{(B)}(\omega)}{2} \right). \tag{8}$$

Recalling that $f(\omega)$ satisfies the Kramers–Kronig relations, and performing some algebra, we may write

$$\frac{2}{\pi} \int_1^\infty \frac{s\, \Delta f_{\mathrm{i}}(s)}{s^2 - \omega^2}\, \mathrm{d}s - \mathrm{i}\frac{2\omega}{\pi} \int_1^\infty \frac{\Delta f_{\mathrm{r}}(s)}{s^2 - \omega^2}\, \mathrm{d}s = R(\omega) \tag{9}$$

for $\omega \in (0, 1)$, where

$$\Delta f_{\mathrm{r}}(s) = \tfrac{1}{2}(f_{\mathrm{r}}(s) - f_{\mathrm{r}}^{\mathrm{ext}}(s)),$$
$$\Delta f_{\mathrm{i}}(s) = \tfrac{1}{2}(f_{\mathrm{i}}(s) - f_{\mathrm{i}}^{\mathrm{ext}}(s)),$$

for $s \in (1, \infty)$.

Our central observation is that the smoothness of the residual $R(\omega)$ for $\omega \in (0, 1)$ is largely dependent on whether the original data is 'locally consistent', and is independent of the quantity of structure exhibited by the measurements. Consider, for example, the functions given by

$$\begin{aligned} f_1(\omega) &= 1/(0.18 - \omega^2 - 0.6\omega\mathrm{i}), \\ f_2(\omega) &= 1/(0.18 - \omega^2 - 0.6\omega\mathrm{i}) + 1/(0.49 - \omega^2 - 0.06\omega\mathrm{i}), \\ f_{\mathrm{nc}}(\omega) &= \mathrm{Re}\,(f_1(\omega)). \end{aligned} \tag{10}$$

The first two belong to $\boldsymbol{H}^2$ and have poles at $\{\pm 0.3 - 0.3\mathrm{i}\}$ and approximately $\{\pm 0.3 - 0.3\mathrm{i}, \pm 0.7 - 0.03\mathrm{i}\}$ respectively. The third is clearly non-causal. Restricting the above functions to the data window $[0, 1]$ and extending them smoothly to zero, we obtain the residual functions $R(\omega)$ shown in figure 1. We observe that the residuals derived from the two causal functions are equally featureless despite the additional pole in f_2. This is not so for the non-causal function. Also in that figure, we show the magnitudes of the projections of the residuals onto the left singular functions of the appropriate operators (see (12) below). The decaying versus non-decaying magnitudes of these projection coefficients are indicative of the causal versus non-causal character of the above data functions. This will hold true generically.

We draw two obvious conclusions: (1) the problem is clearly ill-posed and (2) if the residual has significant local structure, enormous amounts of L_2 energy must be required outside the interval to create it. Analytically, this may be seen by inspection of (9), which is a first-kind Fredholm integral equation. The right-hand side $R(\omega)$ is available explicitly and the functions Δf_{r}, Δf_{i} appearing under the integrals are unknowns defined for $s \in (1, \infty)$. These densities contribute to the residual function on $D = (0, 1)$ as smooth, 'far-field' contributions.

It remains only to invert (9) in some systematic fashion. All suitable inversion procedures can be interpreted in this context and require regularization. Options include the following.

(1) After computing the residual by means of the Hilbert transform (as above), one can switch to a time-domain representation for the residual, solving

$$R(\omega) = \mathcal{F}[\sigma](\omega) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \mathrm{e}^{\mathrm{i}\omega t} \sigma(t)\, \mathrm{d}t \tag{11}$$

for the unknown time signal $\sigma(t)$. An immediate concern is that this operator is not compact. Smoothness of the residual, however, suggests that we look for short 'time

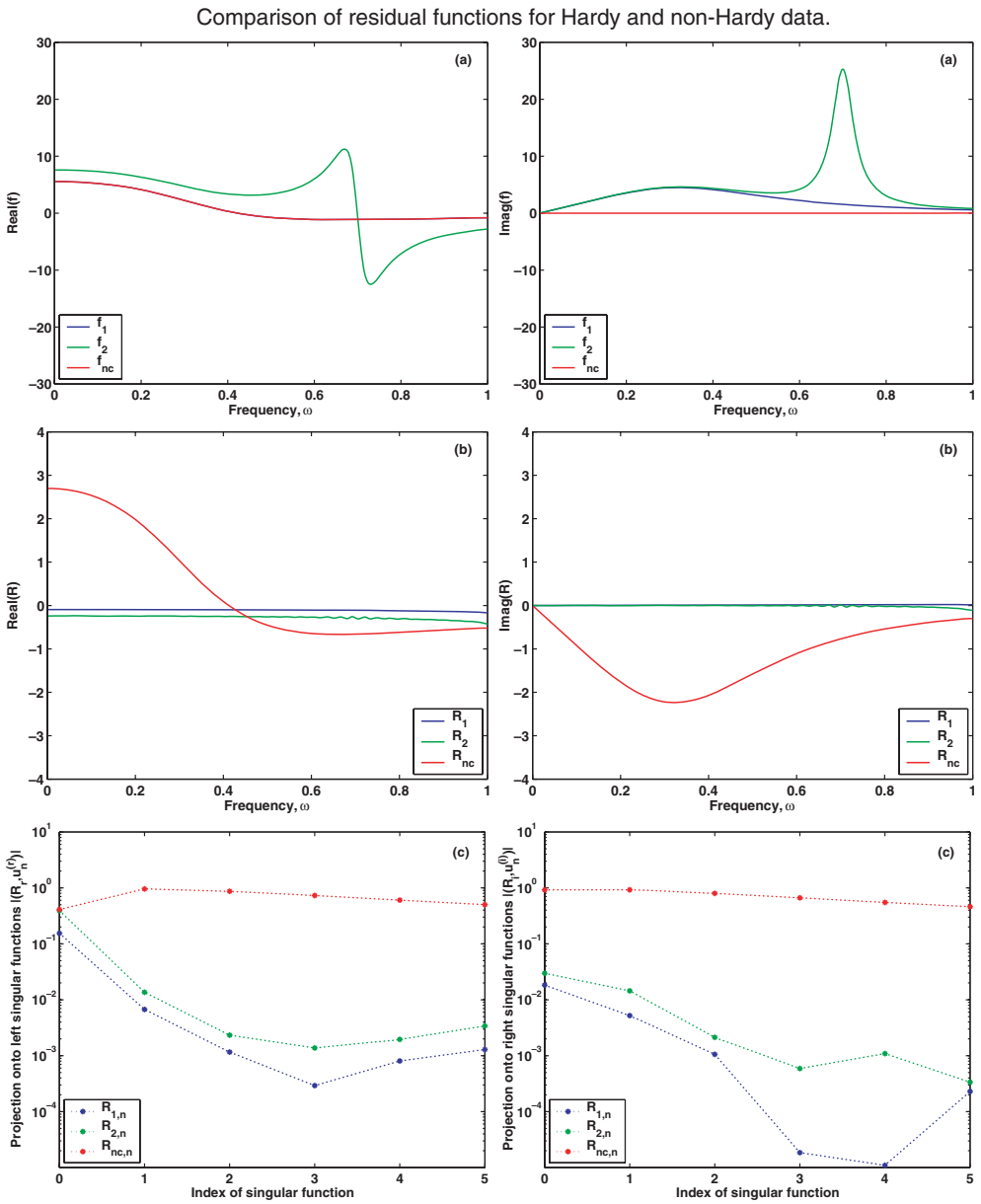

Figure 1. Comparison of residuals. (*a*) Input data (note that the real parts of f_1 and $f_{\rm nc}$ are superimposed in the graph), (*b*) residual, (*c*) projections onto the appropriate right singular functions. The functions f_1, f_2 and $f_{\rm nc}$ are given by (10). Real and imaginary parts are plotted on the left and right respectively.

signatures', leading to a consideration of weighted L^2 spaces. For this, we let L^2_P be the set of functions $\{f \mid \int_0^\infty |f(t)|^2 |P(t)|^{-2}\,{\rm d}t < \infty\}$. It is straightforward to show that L^2_P is a Hilbert space with inner product, $(f, g)_P = \int_0^\infty f\bar{g}|P(t)|^{-2}\,{\rm d}t$. If $P(t)$ is bounded and $\int_{R^+} t^n |P(t)|^2\,{\rm d}t < \infty$ for $n = 0, 1, \ldots, M$ for $M \geqslant 2$, then $\mathcal{F} : L^2_P \to L^2(D)$ as defined in (11) is compact. Suppose, for example, that $P(t)$ is simply the characteristic function

for the interval $[0, T]$. Then, as T increases, we increase the range of representable functions, allowing for the approximation of more and more structured residuals. This is closely connected to the notion of band-limited approximation for which an appropriate theory has been developed using prolate spheroidal wavefunctions [SlPol, LaPol, Rok]. The smallest value of T needed to achieve a specific approximation error is clearly a good measure of the difficulty of extension.

(2) Remaining in the frequency domain, one can employ a variety of compact versions of the integral operators in (9). One approach is to make use of the smoothness of the underlying functions, since we have constructed C^2 extensions. If we let

$$W_0([1,\infty)) = \left\{ f \,|\, f \text{ absolutely continuous, } f(1) = 0, \int_1^\infty |f|^2 + |f'|^2 < \infty \right\},$$

then it is straightforward to show that the integral operators

$$\begin{aligned} &K_1, K_2 : W_0([1,\infty), \mathrm{d}s) \to L^2((0,1), \mathrm{d}\omega), \\ &[K_1 f](\omega) = -\frac{2\,\mathrm{i}\,\omega}{\pi} \int_1^\infty \frac{1}{s^2 - \omega^2} f(s)\,\mathrm{d}s, \\ &[K_2 f](\omega) = \frac{2}{\pi} \int_1^\infty \frac{s}{s^2 - \omega^2} f(s)\,\mathrm{d}s \end{aligned} \tag{12}$$

are compact. We now need to invert equation (9), which takes the form

$$K_2 \Delta f_\mathrm{i} + K_1 \Delta f_\mathrm{r} = R. \tag{13}$$

3.2. Analytic interpolation

We next consider the interpolation problem. The application that we have in mind is the 'phase retrieval' problem in which the modulus of a Hardy function is measured over a finite range of frequencies, and one requires its phase over that same range. In order to make the connection with real and imaginary parts of analytic functions, one can write $f(\omega) = |f(\omega)| \exp(\mathrm{i}\phi(\omega))$ and take the logarithm,

$$\ln(f(\omega)) = \ln(|f(\omega)|) + \mathrm{i}\phi(\omega).$$

As the logarithm is analytic in any domain where its argument is analytic and non-zero, Kramers–Kronig analysis of this function can often be justified. Regardless, the generic problem is that one measures one-half of a Hardy function, e.g. the real part, and wishes to reconstruct its imaginary part. The algorithm is much the same as above, with the obvious exception that the symmetry between real and imaginary parts is broken. We denote the corresponding residual by R_ϕ so as to distinguish this case from the previous one.

Given a smooth data function $f_\mathrm{r}(\omega), \omega \in (0, 1)$, we form its extension, $f_\mathrm{r}^\mathrm{ext}$, as in (7). Next, we compute the appropriate Hilbert transform

$$f_\mathrm{i}^{(A)}(\omega) = -\frac{2\omega}{\pi} \int_0^\infty \frac{f_\mathrm{r}^\mathrm{ext}}{s^2 - \omega^2}\,\mathrm{d}s,$$

and define the residual as the difference

$$\begin{aligned} R_\phi(\omega) &\equiv f_\mathrm{i}(\omega) - f_\mathrm{i}^{(A)}(\omega), \qquad \omega \in (0,1) \\ &= \frac{2\omega}{\pi} \int_1^\infty \frac{1}{s^2 - \omega^2} \Delta f_\mathrm{r}(s)\,\mathrm{d}s, \end{aligned} \tag{14}$$

where

$$\Delta f_\mathrm{r}(s) = f_\mathrm{r}(s) - f_\mathrm{r}^\mathrm{ext}(s).$$

Again, we can invert (14) using either the Fourier time-domain representation (11) with the compact operator $\mathcal{F} : L^2_P \to L^2((0,1))$, or remain in the frequency domain and use the operator K_1.

We expect that R_ϕ is representable with a small number of terms:

$$R_{\phi,N}(\omega) = \sum_{n=1}^{N} \beta_n u_n(\omega),$$

where $\{u_n(\omega)\}$ are the left singular vectors of K_1. The distinction between this problem and the previous one is that, in the continuation problem, there was an experimental determination of $f_\mathrm{i}(\omega)$, $\omega \in D = (0,1)$. Therefore it was possible to project $\mathrm{Im}\,(R)$ onto the space $\mathrm{Span}\{u_n, n = 1 \ldots N\}$ by computing quadratures. Here, the function R_ϕ given by (14) is unknown. Therefore, we need additional information if we are to compute the coefficients β_n. This information may come in several forms. The simplest is the situation in which, via supplementary measurements, it is possible to determine the imaginary part of f at a small number of frequencies. One can, for example, assume that two point values $\{f_\mathrm{i}(\omega_1),\ f_\mathrm{i}(\omega_2)\}$ are available. (In optical experiments, such measurements can be made using a tunable laser and a procedure such as ellipsometry which returns the complex reflectance at fixed wavelengths.) In this case, one may compute the values of $R_\phi(\omega)$ at two frequencies and determine the values of $\{\beta_n\}$ by solving the linear system

$$\begin{aligned} R_{\phi,2}(\omega_1) &= \beta_1 u_1(\omega_1) + \beta_2 u_2(\omega_1) \\ R_{\phi,2}(\omega_2) &= \beta_1 u_1(\omega_1) + \beta_2 u_2(\omega_2). \end{aligned} \tag{15}$$

We note that if it is not possible to obtain any extra information, the problem may be weakened. Namely, we can determine the initial extension and assert that the imaginary part *is* its transform. Recognizing that the norm of a compact operator is given by its largest singular value, it is possible to compute an L^2-error bound with this assertion given only an energy bound on the real part of the function outside the measurement window.

3.3. Discretization

All the above methods involve a SVE of compact operators. For the sake of brevity, we restrict our attention to the continuation problem using the operators K_1 and K_2 from (12). To enforce boundary conditions and differentiability we employ a Galerkin discretization. For this, the functions Δf_i and the real part of the residual R_r are expanded in the following cosine series

$$\begin{aligned} \Delta f_\mathrm{i}(s) &= \sum_{k=0}^{\infty} \frac{\cos\left(\frac{2k+1}{2}\frac{\pi}{s}\right)}{s\sqrt{1+\pi^2\left(\frac{2k+1}{2}\right)^2}} a_k, \\ R_\mathrm{r}(\omega) &= \tfrac{1}{2} b_0 + \sum_{k=1}^{\infty} \cos(k\pi\omega) b_k. \end{aligned} \tag{16}$$

The function $R_\mathrm{r}(\omega)$ belongs to the even half of $L^2(-1,1)$ so the usual cosine series is appropriate. A few words are necessary to justify the expansion for $\Delta f_\mathrm{i}(s)$. In lieu of constructing an orthonormal basis for $W_0([1,\infty))$ defined above, we analyse the operators (12) after right multiplication by the operator defined by

$$\begin{aligned} &T^* : L^2((0,1), \mathrm{d}t) \to L^2((1,\infty), \mathrm{d}s), \\ &[T^* g](s) = \frac{1}{s} g\left(\frac{1}{s}\right). \end{aligned}$$

It is a simple matter to compute the kernel of the product, $K_2 T^* : L^2((0,1), \mathrm{d}t) \to L^2((0,1), \mathrm{d}w)$. Furthermore, the restriction of this product to functions with square-integrable

derivatives that vanish at $t = 1$ (the analogue of $W_0([1, \infty))$) will again be compact. The point of this contortion is that a trigonometric basis for this new space on $(0, 1)$ is immediately available. We return to the expansion for $\Delta f_{\mathrm{i}}(s)$. Recall that $\Delta f_{\mathrm{i}}(s)$ must have odd parity, therefore its pre-image under T^* is even, necessitating the use of cosines. The half-integer fundamental frequencies are chosen to enforce the vanishing at $t = 1$, and the square-root weighting reflects the use of the Sobolev norm. Applying the operator T^* to such expansions gives the form for $\Delta f_{\mathrm{i}}(s)$ above. The heuristic for our specific choice of T^* to accomplish this change of basis is that, as is possible to check, T^* and its adjoint are both unitary transformations in the L^2-norms. Given the SVE of an operator, it is relatively straightforward to compute the corresponding SVEs of new operators related to the original through pre- or post-multiplication by unitary transformations. We recognize that T^* and its adjoint are not unitary in the desired Sobolev norms; however this will introduce a penalty more in efficiency than accuracy. Similar considerations yield parallel expansions in terms of appropriately weighted sine series for the functions Δf_{r} and the imaginary part of the residual R_{i}.

We then represent the operators K_1 and K_2 in these bases. Once the matrix entries are computed, we employ a standard singular-value factorization scheme (from LAPACK) to return the singular-values of the continuous operators K_1 and K_2, and the corresponding cosine (sine) expansions of its singular vectors. The left and right singular functions and singular-values of K_1 are plotted in figure 2.

We have said little about truncating the SVEs after a finite number of terms. One can truncate the expansions (16) at N terms and solve a sequence of problems for $N = 1, 2, \ldots, \infty$ until the least-squares error is less than some prescribed tolerance, tol_{SVD}. The algorithm can then return the parameter N as well as the resulting analytic continuation. As N increases, of course, the solution is more and more susceptible to corruption by noise since the inversion process (6) involves division by the singular values, which are rapidly convergent to zero. We view N as an informal measure of the difficulty of the continuation process.

4. Numerical results

We apply our procedure to a standard low-pass fourth-order Chebyshev filter whose magnitude response is given by

$$\begin{aligned} |\chi_4(\omega)|^2 &= \frac{1}{1 + \gamma^2 |T_4\left(\frac{\omega}{\omega_c}\right)|^2} \\ \gamma &= 0.025 \\ \omega_c &= 0.75, \end{aligned} \tag{17}$$

where T_4 is the standard fourth-order Chebyshev polynomial. One can find the poles of the transfer function by solving for the eight roots of

$$T_4(\omega) = \pm \mathrm{i} \frac{1}{\gamma}.$$

We define $\chi_4(\omega)$ as the inverse of the product of the four roots that lie in the lower half of the complex plane, thus enforcing the requisite analyticity. For details, see [Chen]. Our model causal transfer function is then defined in the Fourier domain by

$$\epsilon(\omega) = \chi_4(\omega) \mathrm{e}^{2\pi \mathrm{i} \omega \tau_0}, \tag{18}$$

where τ_0 represents a time delay that has been introduced by 'upstream' data handling or uncertainties in the experimental setup. We arbitrarily set $\tau_0 = 3.1$.

The data below are obtained by sampling the above function $\epsilon(\omega)$ at approximately 1000 regularly spaced frequencies in the interval $(0, 1)$. In figure 3, we plot the results

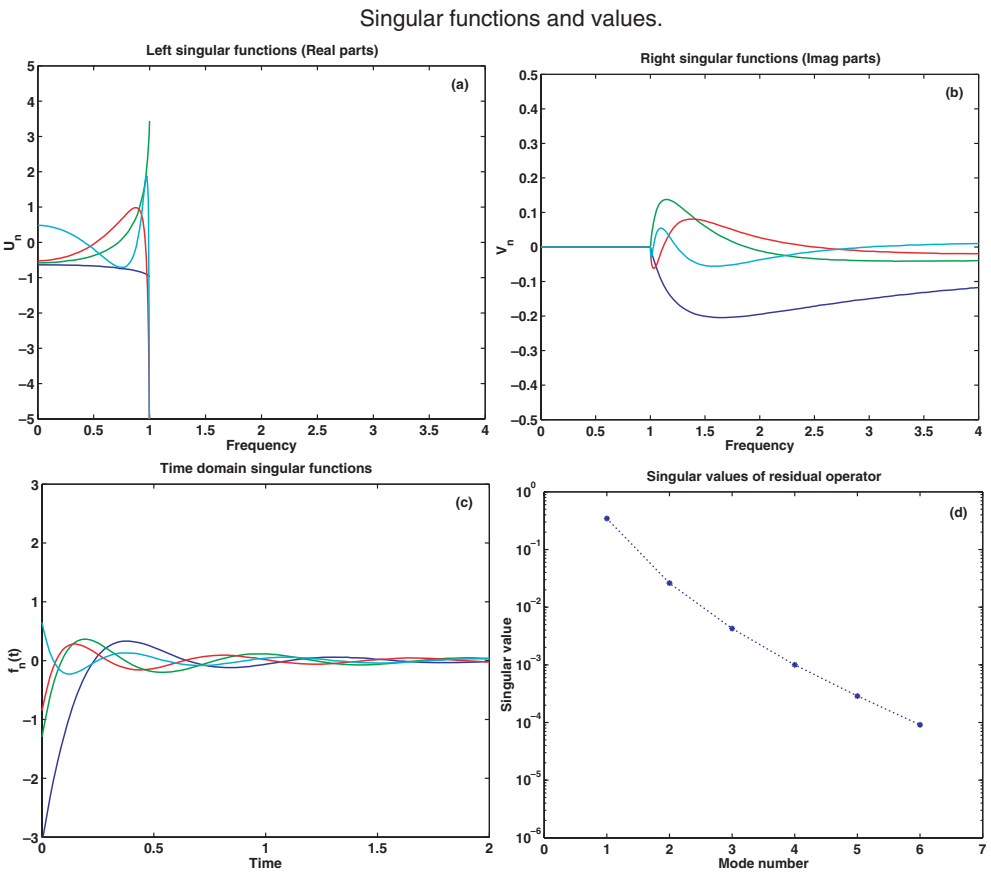

Figure 2. Singular functions. We plot the first four left, right and time-domain singular functions of K_1: $(u_n(\omega), v_n(s), \hat{v}_n(t))$ in (a), (b), (c) respectively. The first six singular values of the same operator are shown in (d).

of the continuation procedure applied to the sampling of $\epsilon(\omega)$ with no noise. The real and imaginary parts along with the extensions $\epsilon_{\mathrm{r}}^{\mathrm{ext}}, \epsilon_{\mathrm{i}}^{\mathrm{ext}}$ from (7) are shown in the top two plots (solid and dashed curves respectively). Note that the magnitude of the original data is of order one and that it contains a number of oscillations over the measurement window. We then compute the Hilbert transform of the extensions and form the residual (8). We plot $R(\omega)$ in figure 3(b) (solid curve). Note that the residual is roughly two orders of magnitude smaller than the original $\epsilon(\omega)$ and is relatively featureless. The expansion of R in the basis of left singular vectors converges with an absolute error

$$\|R - R_{\mathrm{SVE}}\| < 10^{-4}$$

using only two of the left singular functions in both real and imaginary parts. The SVE is inverted using either 1, 3, or 6 singular functions to give a continuation of $R(\omega)$ to all frequencies (figure 3(b), dashed curve). This result is added to the original extensions to form the continuation $\epsilon_C(\omega)$. We plot the real and imaginary parts of ϵ_C (solid curve) and compare with the exact function (dotted curves) in figure 3(c). The agreement is very good within the data window. This is to be expected as the

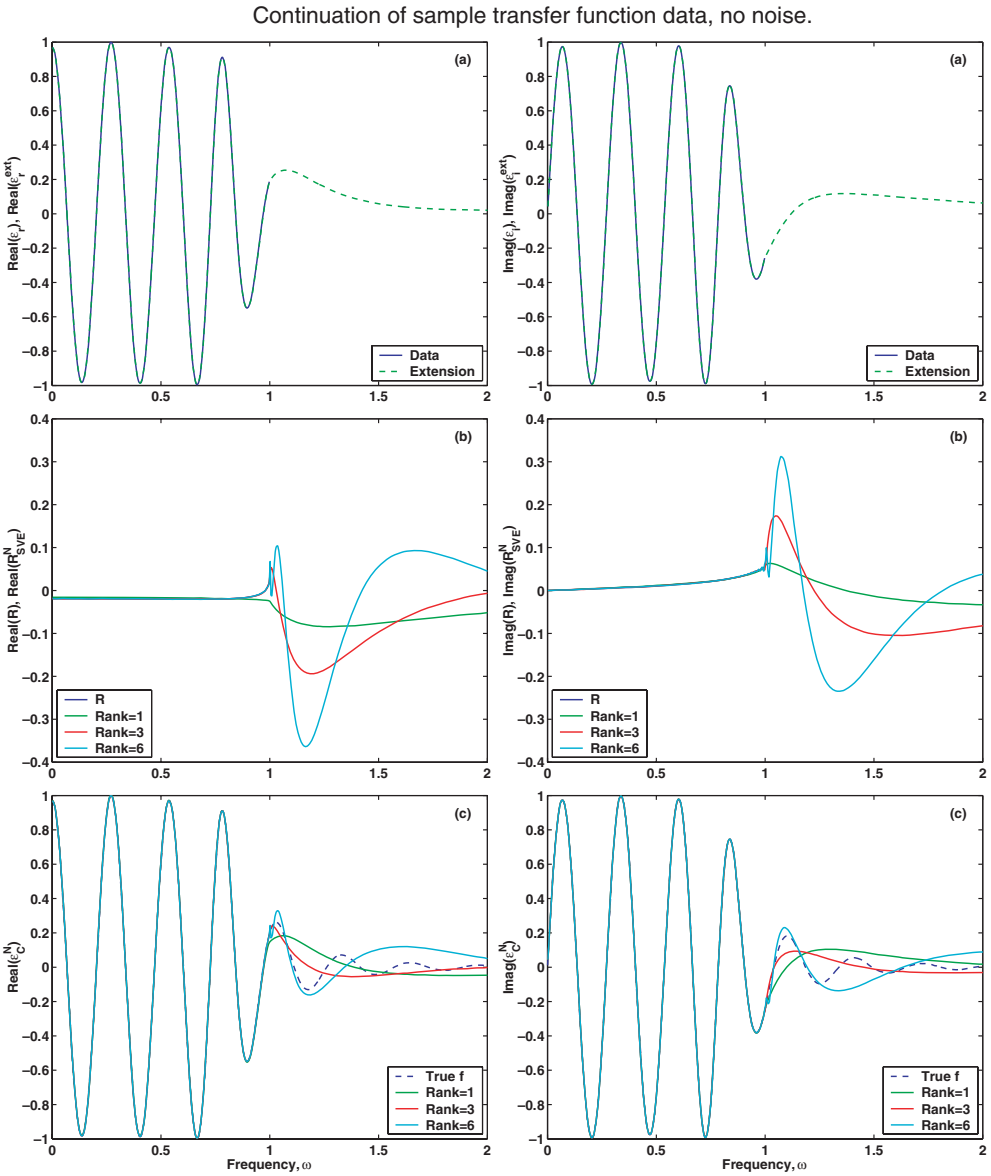

Figure 3. Continuation of causal data. (*a*) Data and initial extensions (solid, dashed), (*b*) residual and its expansion in singular functions, (*c*) true function and results of continuation (dashed, coloured). Real (imaginary) parts are plotted on the left (right).

small magnitude of the residual suggests that the original extensions are also good approximations to the true causal function. The determination of ϵ_C outside the data window is less impressive. While the extensions using six singular functions are more accurate than those using one (the extrapolation manages to 'capture' one additional oscillation of the true causal function), the inclusion of more terms fails to improve matters significantly.

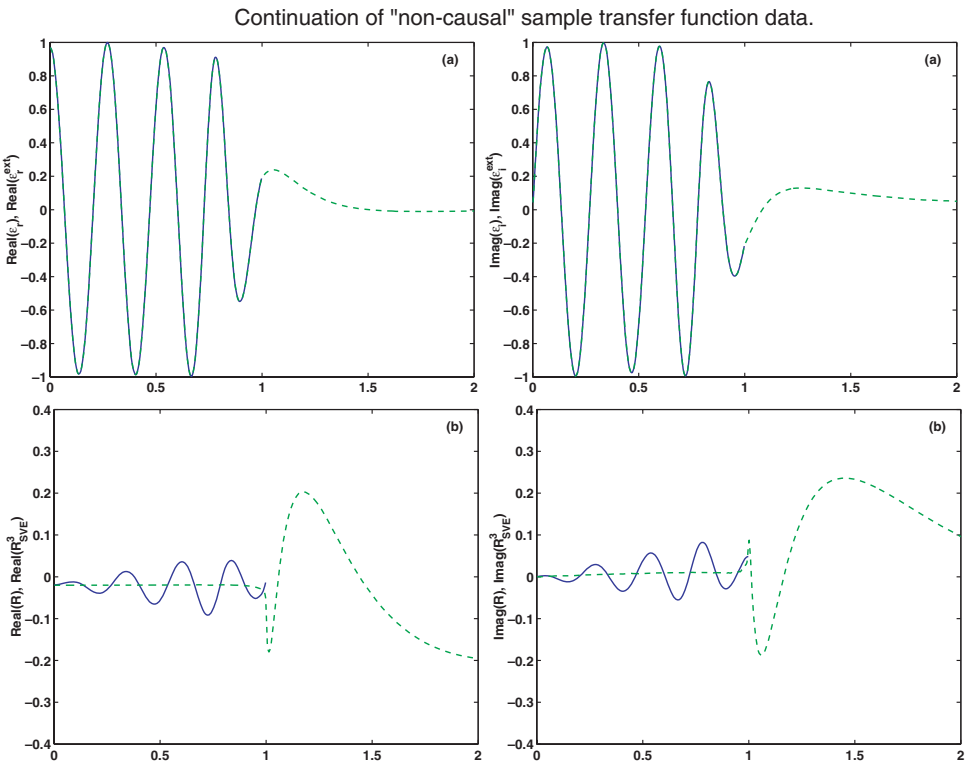

Figure 4. Continuation of non-causal data. (*a*) Data and initial extensions (solid, dashed), (*b*) residual and expansion of residual using three singular functions (solid, dashed). Note the large degree of structure in the residual compared with figure 3. Real (imaginary) parts are plotted on the left (right).

4.1. Non-causal data and the residual

As an illustration of the utility of passing to the residual problem, we repeat the above steps introducing a slight error in the measured signal in the form of a decorrelation between the real and imaginary parts. Specifically, we sample a new transfer function defined by

$$\epsilon(\omega) = \mathrm{Re}\left\{\chi_4(\omega)\mathrm{e}^{2\pi \mathrm{i}\omega\tau_0}\right\} + \mathrm{iIm}\left\{\chi_4(\omega)\mathrm{e}^{2\pi \mathrm{i}\omega\tau_0(1+\delta)}\right\},$$

where $\delta = 0.01$. As before, we plot the real and imaginary parts in figure 4(*a*). To the eye there is no difference between these curves and the previous example. Furthermore, the density theorem of Riesz implies that there exists a causal continuation of these two functions to the entire real line that agrees with the given data to any precision on (0, 1). Performing the initial extension and then computing the residual, however, the 'inconsistency' of the measurement becomes apparent. There is now significant visible structure in the residual function within the data window (figure 4(*b*)). Clearly one anticipates that such a function would have a nontrivial projection onto a large number of singular functions. As noted earlier, the continuation then requires division by a sequence of singular values which is rapidly decaying. In physical terms, this implies that the continuation is introducing considerable power outside the measurement region. We plot the inversion after projecting onto three singular functions as dashed curves in figure 4(*b*).

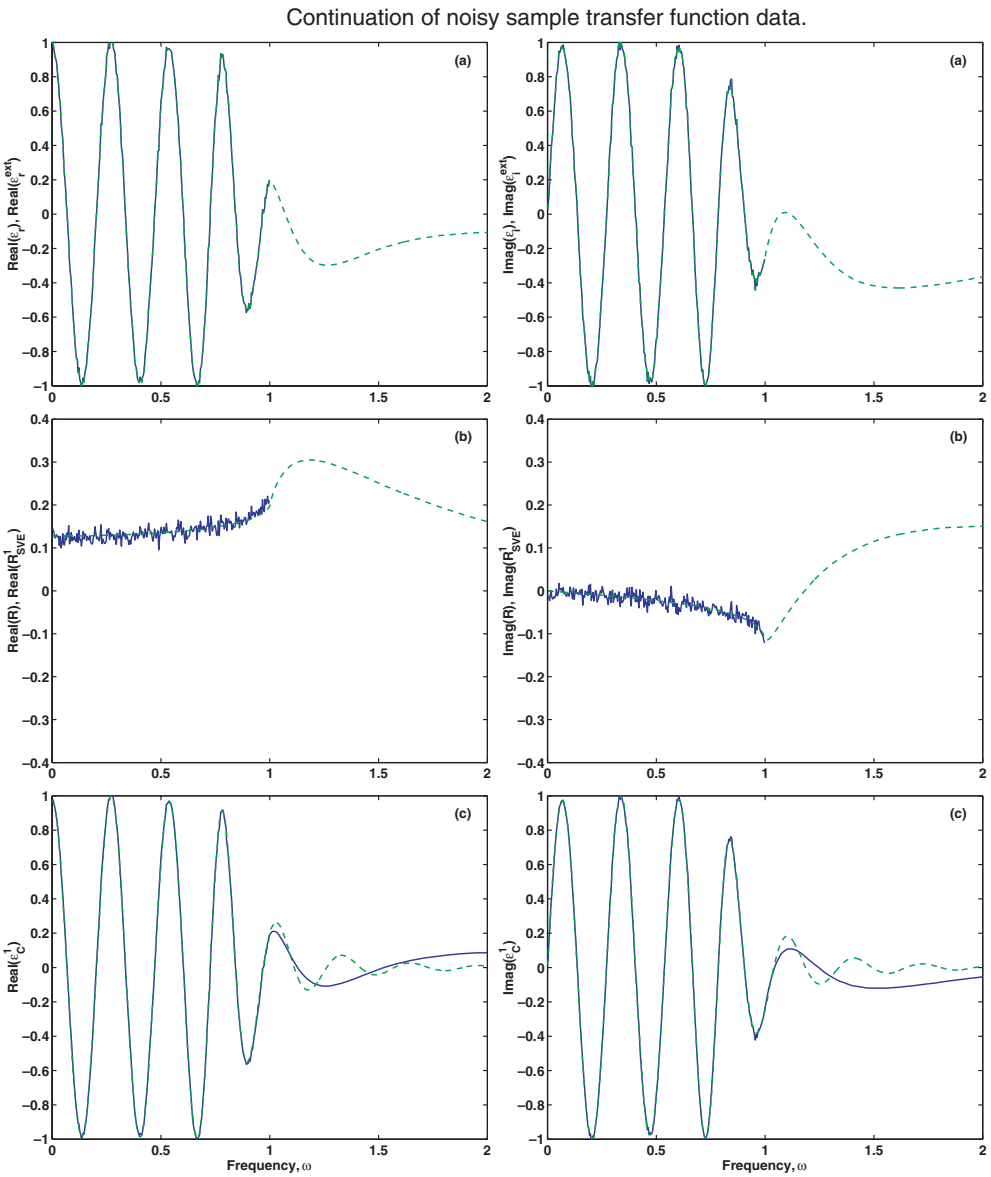

Figure 5. Continuation of causal data with experimental noise. (*a*) data and initial extensions (solid, dashed), (*b*) residual and continuation using one singular function (solid, dashed), (*c*) final continuation and noise-free function (solid, dashed). Real (imaginary) parts are plotted on the left (right).

4.2. Continuation of noisy data

We next demonstrate that the above analytic continuation procedure is robust in the presence of noise. Experimental error is simulated by discrete white noise, i.e. the input to the continuation program is

$$\epsilon_{\text{noise}}(w_{\text{i}}) = \epsilon(\omega_{\text{i}}) + n_{\text{i}},$$

where $\boldsymbol{n}$ is a vector of Gaussian random variables with mean 0 and standard deviation 0.025. The threshold parameter for truncating the projection onto the singular functions will be set equal to this standard deviation, $tol_{\mathrm{SVD}} = 0.025$.

For the underlying, causal, $\epsilon(\omega)$ we use the same low-pass filter as before (18). The results are plotted in figure 5. Again we plot the real and imaginary parts of the noisy data signal at the top. In the middle, we plot the corresponding residual. The high-frequency noise is apparent in this curve. The residual has an expansion in the left singular functions of the operators K_1, K_2 which converges to the noise level using just one singular vector. The lower plots show the result of the continuation along with the error-free $\epsilon(\omega)$. These should be compared with figure 3—approximately the same quality of reconstruction has been achieved on the data window $\omega \in (0, 1)$ while simultaneously controlling the extrapolation energy of the continued signal.

5. Conclusion

We have described a simple framework for analytic continuation problems using a 'deferred correction' point of view. We first compute an explicit guess (involving simple smooth extensions of real-valued functions and Hilbert transforms). We then form a residual problem and view all subsequent methods in terms of the residual equation (9) and SVEs of operators.

Since the procedure is inherently ill-posed, one cannot compute such continuations accurately. We believe the most valuable aspect of our approach is that the residual function itself contains useful information. With little structure present, the continuation process is easily carried out (but not guaranteed). More important is that, with a lot of structure present, we can be certain that the original function contains significant power outside the data window. This may be of diagnostic use. This notion is similar in spirit to that presented by Milton *et al* [MEM]. In some sense, the present paper can be viewed as the L_2 analogue of their L_∞ analysis.

Acknowledgments

The authors thank Yu Chen and Bob Kohn of the Courant Institute and Paul Hale and Brad Alpert at NIST/Boulder for many valuable discussions and insights relating to the problems addressed above. This work was funded by NSF and DARPA through the project Virtual Integrated Processing of YBCO Thin Films.

References

[Con] Conway J B 1990 *A Course in Functional Analysis* (New York: Springer)

[Chen] Chen W K 1986 *Passive and Active Filters: Theory and Implementations* (New York: Wiley)

[DM] Dym H and McKean H P 1972 *Fourier Series and Integrals* (New York: Academic)

[GRV] Gilliam D S, Lund J R and Vogel C R 1990 Quantifying information content for ill-posed problems *Inverse Problems* **6** 725–36

[Kra] Kronig R de L 1926 On the theory of dispersion of x-rays *J. Opt. Soc. Am.* **12** 547–57

[LaPol] Landau H J and Pollak H O 1961 Prolate spheroidal wave functions, Fourier analysis, and uncertainty—II *Bell Syst. Tech. J.* **40** (January) 65–84

[MEM] Milton G W, Eyre D J and Mantese J V 1997 Finite frequency range Kramers–Kronig relations: bounds on the dispersion *Phys. Rev. Lett.* **79** 3062–5

[Par] Partington and Jonathan R 1997 *Interpolation, Identification, and Sampling (London Mathematical Society Monographs, New Series #17)* (Oxford Science Publications)

[Rok] Rokhlin V 2000 A procedure for the design of apparata for the measurement and generation of band-limited signals *Research Report* 1196 Dept. of Computer Science, Yale University

[SlPol] Slepian D and Pollak H O 1961 Prolate spheroidal wave functions, Fourier analysis, and uncertainty *Bell Syst. Tech. J.* **40** (January) 43–63