Ensuring Good Behaviour of the Discrete Hilbert Transform

PHILIP J. PARKER and BRIAN D. O. ANDERSON
Department of Systems Engineering, Research School Physical Sciences, Australian National University, G.P.O. Box 4, Canberra, A.C.T. Australia

ABSTRACT: This paper studies the generation of stable transfer functions for which the real or imaginary part takes prescribed values at discrete uniformly spaced points on the unit circle. Formulas bounding the error between a particular interpolating function and any function consistent with the data are presented, which have the desirable property that the error goes to zero exponentially fast with the number of interpolating points. The paper also examines generation of stable minimum phase transfer functions for which the magnitude takes prescribed values at uniformly spaced points on the unit circle, and presents error bounds for this problem. Connection with the discrete Hilbert transform is made. The effect of uncertainty in the original data is also examined.

1. INTRODUCTION

This paper's concern is with the following style of problem. Suppose we have the real part of a (discrete-time) transfer function at n discrete points on the unit circle; moreover, suppose that the function is stable. How can one reconstruct the function, both real and imaginary parts, on the whole of the unit circle? A related problem is how to reconstruct, from values of the transfer function magnitude given at only n discrete points on the unit circle, the complete transfer function — magnitude and phase — on the entire unit circle. In addressing this problem we combine two ideas: Interpolation and the Hilbert transform.

Apart from theoretical interest this problem has also practical motivation. Measurement data on a transfer function, either of the real part or of the magnitude, is unlikely to exist for every frequency. Much more probable is that measurements yield the real part or magnitude at only discrete frequencies and one must then recover the complete transfer function at not only the frequencies given but at all points on the unit circle.

Given real part data on the entire unit circle, recovery of the imaginary part (and vice-versa) is a more-or-less standard problem, tackled via the Hilbert transform, see e.g. [1]. Under the assumption that a transfer function is stable (i.e. free of singularities in $e^{j\omega}$) and that the real part is known at all frequencies (i.e. on $|z|=1$), the Hilbert transform constructs the imaginary part to within an additive constant. If the transfer function corresponds to a real stable function, it must be real at $z=1$, and this will identify the constant. Similarly, if the transfer function is stable and its imaginary part is known at all frequencies, one can construct the real part to within an additive constant.

Recovery from phase or magnitude data given on the entire unit circle is almost as simple. Under the assumption that a transfer function is stable and minimum phase (all singularities of the function and its inverse lie in $|z|<1$), so that its logarithm is stable, the logarithm of the gain and the phase, being the real and imaginary part of the logarithm of the function, are related by the Hilbert transform, and so can be determined from each other, to within an arbitrary constant. Alternatively, given the magnitude squared, spectral factorization yields the whole function.

In considering use of the conventional Hilbert transform, the question of how errors in the data translate into errors in the result is an important one, the solution of which is not as widely known as it should be. Consider for example constructing the imaginary part from the real part of a stable transfer function. Then an arbitrarily small perturbation in the real part can induce an arbitrarily large change in the imaginary part. For suppose $G(z) = R(e^{j\omega}) + jI(e^{j\omega})$ with $G$ stable. Then we have, see [1, p. 344],

$$1(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{R(e^{j\theta})\cot \frac{\theta - \omega}{2}}{e^{j\theta}} d\theta$$

(1.1)

(The Cauchy principal value of the integral is to be calculated). Suppose there is a perturbation in $R(e^{j\omega})$ by a constant amount $\varepsilon$ over $(-\omega, \omega)$. Then the corresponding perturbation in $I(e^{j\omega})$ is

$$\delta I(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\delta R(e^{j\theta})\cot \frac{\theta - \omega}{2}}{e^{j\theta}} d\theta$$

Take $\omega = -\omega_0$, then the integral diverges.

The same sort of discontinuity can arise in going between magnitude and phase of a minimum phase stable transfer function, see e.g. [2]. It is this discontinuity which underpins the comparative difficulty of doing Hilbert transforms with data at only discrete points. Suppose that one is given the values of the real part of a stable function at frequencies $\omega_k = \exp(2\pi i k/n)$, $k=0,1,\ldots,n-1$. A naive approach to reconstruction of the function, or an approximation to it, might involve fitting a curve through the interpolation data, and then Hilbert transforming as in (1.1). Because of the lack of continuity of the Hilbert transform, one could not expect, at least a priori that the reconstructed object is close to the true object.

The approach we follow exploits the "Hilbert transform for the discrete Fourier transform" [1, p.335]. Suppose we have data $Re[G(e^{j2\pi k/n})]$, $k=0,1,\ldots,n-1$, and we wish to find $G(e^{j2\pi k/n})$, $k=0,1,\ldots,n-1$. There is a unique finite impulse response $g(t)$, $t=0,1,\ldots,n-1$ with $g(0)=0$ for $|n/2|<1$ satisfying that its discrete Fourier transform $G(e^{j2\pi k/n})$ has a real part that coincides with the given data. Uniquely determined, therefore, is the transfer function

$$P(z) = \sum_{l=0}^{n-1} g(l)z^{-l}$$

While this guarantees interpolation in that

$$Re[P(e^{j2\pi k/n})] = Re[G(e^{j2\pi k/n})] \quad k=0,1,\ldots,n-1$$

it is not part of our problem specification that the given data corresponds to an FIR transfer function: in fact the "true" transfer function $G(z)$ may be any member of the infinite class of functions which take the prescribed real values at certain discrete points on the unit circle. We hope that $P(z)$ is "close" to the true function $G(z)$: it is important to quantify the approximation's accuracy. Section 2 of this paper gives the construction $P(z)$ and bounds the error between $P(z)$ and $G(z)$. 

The situation is less clear again when gain/phase data is involved. In [1, p.257], the following idea is advanced. Suppose $G(z)$ is the $z^{-1}$ transform of a finite length sequence of $g(t)$, and let $G(z)=\log G(z)$. Choose $n$ at least equal to twice the length of $g(t)$, and suppose that $G(\exp[2j\pi k/n])$ is known for $k=0,1,\ldots,n-1$. Using the Hilbert transform of the discrete Fourier transform, a collection of values of argument is obtained. These are not exactly the values of $\arg G(\exp[2j\pi k/n])$, basically because $G(z)$ does not correspond to a finite impulse response. Further, when $G(z)$ does not correspond to a finite length sequence, the validity (even in an approximate sense) is yet more difficult to pin down. Error formulas presented in Section 3 do however serve to display the validity. Also in Section 3 is an alternative approach, where we find an FIR, non-causal system which interpolates $G(\exp[2j\pi k/n])$, $k=0,1,\ldots,n-1$; then spectral factorization yields an FIR approximation to $G(z)$.

Our approach to all these problems rests on two broad ideas. First, we describe an algorithm for interpolating the given data to come up with a stable, or stable and minimum phase transfer function. Then we establish an error bound between the function we construct and the class of functions which could have generated the data. The error becomes vanishingly small as the number of interpolating points becomes larger.

Never is one given perfectly accurate data. In Section 4, we analyse the effect of uncertainty in the original (i.e. discrete frequency) data. (This is quite different to defining the maximum error between the interpolating function we construct, and any function which could have generated error free initial data). Our analysis includes two different models for the data uncertainty, and focusses upon the effect of increasing the number of interpolation points.

Illustrative examples follow in Section 5; the paper provides concluding remarks in Section 6.

2. FUNCTION CONSTRUCTION FROM REAL PART
2.1 Problem Statement
For an unknown transfer function $G(z)$, we are given the following information:
(i) $G(z)$ is analytic in $1z>\rho^{-1}$ for some $\rho>1$
(ii) Values of $\Re\{G(\exp[2j\pi k/n])\}$, $k=0,1,\ldots,n-1$
(iii) $G(z)$ corresponds to a real impulse response, so that $\Re\{G(\exp[j\theta])\}$ and $\Im\{G(\exp[j\theta])\}$ are respectively even and odd in $\omega$.

From this data, we seek to find a $v$
$$P(z) = \sum b_i z^{-i},$$
with $n=2v$ or $n=2v+1$ such that
$$\Re\{G(\exp[2j\pi k/n])\} = \Re\{G(\exp[2j\pi k/n])\} \quad k=0,1,\ldots,n-1$$
and we seek a value of $K$ such that
$$\left|G(z)\right| = \max_{\omega} \left|G(\exp[j\omega])\right| \in K$$
with the property that $K>0$ as $n\to\infty$ in a predictable fashion.

2.2 Construction of interpolating function $P(z)$
We proceed in two steps. First, we construct a two-sided (non-causal) FIR $Q(z)$ with values interpolating $\Re\{G(\exp[2j\pi k/n])\}$. Then from $Q(z)$ we find a (causal) FIR $P(z)$ satisfying (2.1). We treat the cases of $n$ odd and $n$ even separately.

(a) $n$ odd, $n=2v+1$.
By DFT methods, it is straightforward to construct $Q(z)$, of the form
$$Q(z) = b_0 z^{-1} + \ldots + b_{2v} z^{-1} + \ldots + b_{2v-1} z^{-1} + b_{2v} z^{-v}$$
$$+ \ldots + b_{2v} z^{-v} + \ldots + b_{2v} z^{-v} + b_{2v} z^{-v}$$

From this data, we seek to find a $v$
$$P(z) = \sum b_i z^{-i},$$
with $n=2v$ or $n=2v+1$ such that
$$\Re\{G(\exp[2j\pi k/n])\} = \Re\{G(\exp[2j\pi k/n])\} \quad k=0,1,\ldots,n-1$$
and we seek a value of $K$ such that
$$\left|G(z)\right| = \max_{\omega} \left|G(\exp[j\omega])\right| \in K$$
with the property that $K>0$ as $n\to\infty$ in a predictable fashion.

2.3 Error in the Real Part
In this subsection, we shall examine the error in $\Re\{P(z)\} = \Re\{G(\exp[j\theta])\}$. To derive an error bound, we need some preliminary results. First we recall the following result, proved in [4]:

Lemma 2.1: Let $H(z)$ be a transfer function analytic in $|z|<\rho$ for some $\rho>1$.

$$|Q_{2v}(z)| = \sum b_i z^{i},$$

interpolates $H(z)$ in the $(2v)$th roots of unity, then for any $R_{1}$ with $1<R_{1}<\rho$

$$\begin{align*}
|H(z)| &< \left(\frac{2\rho}{(R_{2}^{2v+1})-R_{1}^{-1}}\right) \\
&\quad \times \max_{|\theta|\leq\pi} |H(R_{1} \exp[2j\pi \theta])| \\
&\quad + \max_{|\theta|\leq\pi} |H(R_{2} \exp[2j\pi \theta])| \quad (2.8a)
\end{align*}$$

1. If $Q_{2v+1}(z) = \sum b_i z^{i}$
interpolates $H(z)$ in the $(2v+1)$th roots of unity, then for any $R_{1}$ with $1<R_{1}<\rho$

$$\begin{align*}
|H(z)| &< \left(\frac{2\rho}{(R_{2}^{2v+1})-R_{1}^{-1}}\right) \\
&\quad \times \max_{|\theta|\leq\pi} |H(R_{1} \exp[2j\pi \theta])| \\
&\quad + \max_{|\theta|\leq\pi} |H(R_{2} \exp[2j\pi \theta])| \quad (2.8b)
\end{align*}$$

Actually, only the result for even $n$ is proved in [4]. However, the result for odd $n$ follows by virtually the same argument.

We shall also need a further result on the behaviour of the last coefficient $b_{-v}$ in a $(2v)$th roots of unity interpolation.

Lemma 2.2: Assume the same hypotheses as Lemma 2.1, with $n=2v$. Then
between fmt zero at the interpolating points, but in general only at these opposed to its real part. The error in the real part is bounded in a manner described in the following theorem:

**Theorem 2.1** Assume G(z) satisfies the assumptions listed in Section 2.1. Let P(z) be defined as described in Section 2.2. If n = 2v + 1, there holds for any R, with 1 < R, < ρ,

\[ \text{Re}P(e^{j\theta}) - \text{Re}P(e^{j\varphi}) \leq \frac{2K_{RV}^{2v+1}}{R^{2v+1}} \left[ \max_{\theta \in (-\pi, \pi]} |C(R, e^{j\theta})| + \max_{\theta \in (-\pi, \pi]} |G(R, e^{j\theta})| \right] \]

while if n = 2v, there holds

\[ \text{Re}P(e^{j\theta}) - \text{Re}P(e^{j\varphi}) \leq \frac{2K_{RV}^{2v+1}}{R^{2v+1}} \left[ \max_{\theta \in (-\pi, \pi]} |C(R, e^{j\theta})| + \max_{\theta \in (-\pi, \pi]} |G(R, e^{j\theta})| \right] \]

**Remark:** The error is effectively of the form KR^{2v+1}/2.

**Proof:** Omitted due to space limitations.

With these lemmas in hand, we can now bound the error between ReP(e^{jθ}) and ReG(e^{jθ}). This error is of course zero at the interpolating points, but in general only at these interpolating points. A first step for obtaining a bound on |P(e^{jθ}) - G(e^{jθ})|, as opposed to its real part. The determination of this error is a key step for obtaining a bound on |P(e^{jθ}) - G(e^{jθ})|, as opposed to its real part. The key is the following Lemma.

**Lemma 2.2** Let \( M(z) = \text{Re}P(e^{j\theta}) + iM(e^{j\varphi}) \), with \( M(e^{j\theta}) \) real and even in \( \omega \), \( M(e^{j\theta}) \) real and odd in \( \omega \), and \( M(e^{j\theta}) \) the Fourier transform of a real, causal impulse response, such that \( ||M(e^{j\theta})||_1 \) and \( ||1/M(e^{j\theta})||_1 < 1 \) are both finite. Then

\[ M(e^{j\theta}) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{f}(\omega) \left[ \frac{d}{d\omega} M(e^{j\omega}) \right] 1_{<\omega<\pi} \frac{d\omega}{2} \]

and

\[ ||M(e^{j\theta})||_1 < 2 \int_{-\pi}^{\pi} \left| \frac{d}{d\omega} M(e^{j\omega}) \right| 1_{<\omega<\pi} \frac{d\omega}{2} \]

**Proof:** Omitted due to space limitations.

To return to the interpolation problem of interest to us, we simply identify M(z) in the Lemma with G(z) - P(z). Putting together Theorem 2.1, Theorem 2.2 and the Lemma, we have

**Main Theorem 2.3** Suppose G(z) is analytic in \( \text{Re} z > \rho^{-1} \) for some \( \rho > 1 \), that values of \( \text{Re}P(e^{j\omega}) \), \( \kappa = 0,1, \ldots, n-1 \) are prescribed, and that G(z) corresponds to a real impulse response. Then we can find P(z) a polynomial of degree \( \nu = n/2 \) or \( \nu = (n-1)/2 \) in \( z^{-1} \) such that,

\[ \text{Re}P(e^{j\omega}) = \text{Re}G(e^{j\omega}) \]

and such that for any \( R_1 \), with \( 1 < R_1 < \rho \),

\[ ||G(e^{j\theta}) - P(e^{j\theta})||_1 < \left( K_1(R_1) + K_2(R_1) \right) \max_{\theta \in (-\pi, \pi]} |C(R_1, e^{j\theta})| + \max_{\theta \in (-\pi, \pi]} |G(R_1, e^{j\theta})| \]

with \( K_1(R_1), K_2(R_1) \) going to zero as \( R_1 \to \infty \).

**Remark:** A very similar theorem naturally holds for the problem of passing from values of \( \text{Im}G(e^{j\omega}) \), \( \kappa = 0,1, \ldots, n-1 \) plus the value of \( \text{Re}G \) at one particular point to an estimate of \( G(e^{j\omega}) \) using a polynomial of degree \( n \).

**2.6 Rapprochement with the Discrete Hilbert Transform**

The discrete Hilbert transform (DHT) [1, p.353], [5] arises in one of two ways. Let h(n) denote the evaluation of \( T_0g(n) \) at \( \chi = \exp(j2\pi n) \); then the discrete Hilbert transform constructs \( \text{Im}h(k) \), \( k = 0,1, \ldots, n-1 \), from \( \text{Re}h(k) \), \( k = 0,1, \ldots, n-1 \), or vice versa. Second, let h(n) be a complex periodic sequence of period \( n \), equal to zero for \( n < 1/2 \). Let H(k) denote the coefficients of a Fourier series for h(n). Then the Hilbert transform links \( R(k) \) to \( \text{Im}h(k) \).

It is commonly suggested that if one has \( \text{Im}G(e^{j\omega}) \) for a stable G, then \( \text{Im}G(e^{j\omega}) \) should be estimated as if G were associated with a finite impulse response of length \( n/2+1 \), i.e. via the discrete Hilbert transform. Left unexplained is how the result of this computation might relate to the true value of \( \text{Im}G(e^{j\omega}) \) and how the \( z^{-1} \)-transform of the finite impulse response might relate to \( G(z) \). As it turns out, we now have the wherewithal to clarify these questions.

In [1], the DHT construction is performed for \( n \) even. Comparison of Section 2.2 with [1, p.354] shows that the quantity \( \text{P}(2) \) of the \( z^{-1} \) -transform function computed by the DHT, corresponding to an FIR with support over [0, n/2]. Hence our main result provides a theoretical underpinning to the use of the DHT when the underlying transfer functions may not come from a finite impulse response: the underpinning rests on the error formulas of the Main Theorem 2.3.
2. INTERPOLATION OF GAIN-PHASE DATA

In this section, we shall concentrate on the problem where the given data is magnitude data. Of course, very similar results apply when the data is phase data.

There are two broad approaches, summed up in Figures 3.1 and 3.2, and we will discuss them in turn.

3.1 Problem Statement

For an unknown transfer function \( G(z) \), we are given the following information:

(i) \( G(z) \) and \( G^{-1}(z) \) are analytic in \( |z| > \rho^{-1} \) for some \( \rho > 1 \).

(ii) \( G(z) \) corresponds to a real impulse response, so that \( \arg G(\exp j2\pi k/n) \) is even and \( \arg G(\exp j\omega) \) is odd in \( \omega \).

From this data, we seek to find an interpolant of \( G(z) \) of the form

\[ P(z) = \sum_{k=0}^{n} b_k z^{-k} \quad \text{or} \quad P(z) = \exp \left( \sum_{k=0}^{n} b_k z^{-k} \right) \quad \text{for} \quad n = 2v \text{ or } n = 2v+1, \]

such that \( P(\exp j2\pi k/n) = G(\exp j2\pi k/n) \). Arguing as in Section 2, it follows that \( b_k = b_{-k} \) and \( z \). We shall argue later that for large enough \( v \), \( Q(e^{j\omega}) > 0 \) for all \( \omega \).

It follows that we can factor \( Q(z) \) as

\[ Q(z) = P(z)P(z^{-1}) \]

with

\[ P(z) = c_0 + c_1 z^{-1} + \ldots + c_{v-1} z^{-v} \quad \text{and} \quad c_0 = -c_{v-1} \quad \text{real} \]

and all zeros of \( P(z) \) inside \( |z| < 1 \). \( P(z) \) is then the solution of the interpolation problem.

In case \( n = 2v \), define

\[ \hat{Q}(z) = 2 \sum_{k=0}^{v-1} b_k z^{k} \]

such that \( \hat{Q}(\exp j2\pi k/n) = \hat{G}(\exp j2\pi k/n) \), and then set

\[ Q(z) = \sum_{k=0}^{v-1} b_k z^{k} + \sum_{k=0}^{v-1} b_{v-k} z^{v-k} \]

At the points \( \exp j2\pi k/n \), \( Q \) takes the same values as \( Q \). As before, \( b_k = b_{-k} \) and is real. Again, for large enough \( v \), \( Q(e^{j\omega}) > 0 \) for all \( \omega \). We define \( P(z) \) as in (3.4), (3.5).

In both cases \( n = 2v \) and \( n = 2v+1 \), we have

\[ G(\exp j2\pi k/n) = P(\exp j2\pi k/n) \quad \text{for} \quad k = 0,1,\ldots,n-1 \quad \text{(3.8)} \]

The determination of \( P(z) \) from \( Q(z) \) in (3.4) is not necessarily an easy task. It should be noted however that procedures exist which do not rely on determination of all the zeros of \( Q(z) \), see e.g. [6,7] for a linearly convergent algorithm and [8] for a quadratically convergent algorithm.

3.2 Error Bounds

We shall set out the arguments justifying the error bounds in a much briefer way than in the last section.

**Lemma 3.1** Assume \( G(z) \) satisfies the assumptions listed in Section 3.1 and \( H(z) = G(z)G^{-1}(z) \). Let \( Q(z) \) be formed as described in Section 3.2. Then for any \( R \), with \( |R| < \rho \),

\[ |H(e^{j\omega}) - Q(e^{j\omega})|_\infty < R |\alpha(R)|_{\infty} \max_{\theta \in (-\pi, \pi)} |H(R, e^{j\theta})| \quad \text{(3.9)} \]

\[ \| \frac{d}{d\omega} [H(e^{j\omega}) - Q(e^{j\omega})] \|_\infty \leq R^2 [\alpha(R) + \gamma(R)] \max_{\theta \in (-\pi, \pi)} |H(R, e^{j\theta})| \quad \text{(3.10)} \]

Here \( \alpha(R) \), \( \beta(R) \), and \( \gamma(R) \) are functions of \( R \), which are independent of \( \omega \), but will vary according as \( n \) is even or odd.

**Proof** Recognize that \( Q \) interpolates \( H \) at \( \exp(j2\pi k/n) \) in case \( n \) is odd, and \( Q + j\beta \exp(j2\pi k/n) \) interpolates \( H \) in case \( n \) is even. Then the same arguments as yielded Lemma 2.1 and Theorem 2.2 yield (3.9) and (3.10).

Remark At this stage, we can justify our earlier claim that \( Q(e^{j\omega}) > 0 \) for suitably large \( n \). Because \( G \) has all its zeros inside \( |z| < \rho^{-1} \), \( G \) is nonzero on \( |z| = 1 \). Hence \( |H| \) is nonzero on \( |z| = 1 \). By (3.9), for suitably large \( n \), we can ensure that

\[ |H(e^{j\omega}) - Q(e^{j\omega})|_\infty < \max_{\theta \in (-\pi, \pi)} |H(R, e^{j\theta})| \]

for all \( \omega \).

Then \( Q(e^{j\omega}) > 0 \) for all \( \omega \).

We need to deduce several consequences of this Lemma. The first two relate to the behaviour of \( |G|^2 - 1 \) and its derivative.

**Corollary 3.1** Assume the same hypotheses as for Lemma 3.1, and assume that \( P(z) \) is defined as described in Section 3.2. Suppose that

\[ |G(e^{j\omega})| > c \quad \text{for all} \quad \omega \quad \text{(3.11)} \]

Then

\[ |I! - \|P\|_\infty < c^{-2} |\alpha(R)|_{\infty} M \quad \text{(3.12)} \]

**Proof** Omitted due to space limitations.

**Corollary 3.2** Assume the same hypotheses as for Lemma 3.1, and assume that \( P(z) \) is defined as in Section 3.2. With \( c, M \) as in Corollary 3.1, and

\[ \frac{d}{d\omega} \|P\|_\infty \leq \frac{\frac{c}{2} \tilde{G}(R, e^{j\omega})}{\tilde{G}(R, e^{j\omega})} \quad \text{(3.14)} \]

there holds, for sufficiently large \( n \), say \( nN \),

\[ \|G(e^{j\omega})\|_\infty \leq \frac{\frac{c}{2} \tilde{G}(R, e^{j\omega})}{\tilde{G}(R, e^{j\omega})} \quad \text{(3.15)} \]

**Corollaries 3.1 and 3.2 allow us to construct a bound on the argument of \( G \). The tool is to appeal to much the same argument as used in Section 2, where we had bounds on the derivative of the real part of a function, and obtained a bound on its imaginary part.**

**Lemma 3.2** Assume the same hypotheses as Lemma 3.1 and suppose \( P(z) \) is determined as in Section 3.2. Then for all \( n \) exceeding some \( N \), there exist constants \( K, L \) depending on

\[ \inf_{\omega} |G(e^{j\omega})| \geq K \quad \text{and} \quad \max_{\theta \in (-\pi, \pi)} |H(R, e^{j\theta})| \]

such that

\[ \arg G \leq (K+L) \quad \text{(3.16)} \]

**Proof** Omitted due to space limitations.

Finally this brings us to the main result of Section 3.2.

**Main Theorem 3.1** Let \( G(z) \) be an unknown transfer function with \( G(z) \) analytic in \( |z| > \rho^{-1} \) for some \( \rho > 1 \), and with values of \( G(\exp j2\pi k/n) \) prescribed for \( k = 0,1,\ldots,n-1 \). Suppose \( Q(z) \) corresponds to a real impulse response, so that \( 1/G(\exp j\omega) \) and \( \arg G(\exp j\omega) \) are even and odd in \( \omega \). Let \( P(z) \) be a polynomial in \( z^{-1} \) of degree \( v \), where \( n = 2v \) or \( 2v+1 \), with all zeros (regarded as a function of \( z \)) in \( |z| < 1 \),
such that $P_1$ interpolates $G_1$ at $\exp(2\pi k/n)$, with $P(z)$ determined according to the procedures of Section 3.2. Then
$$\Pi 1 - G^{-1} P_1 |_{\infty} \leq (K_1 + vL_1) R_Y^2$$
(3.17)
for constants $K_1$, $L_1$ depending on
$$\inf_R \| G(e^{j\omega}) \|_2, \| G'(e^{j\omega}) \|_2, \| R \|_2,$$ and max $\| H(e^{j\omega}) \|_2, \| H'(e^{j\omega}) \|_2$, \(\omega \)
where $H(z) = G(z)G(z^{-1})$.

The above procedure generates a rational approximation of $\log G$. There is a minor variation that will generate a rational approximation of $G$: we find a
$$n-1 \quad \Pi(z) = \sum_{m=0}^{n-1} P_m z^{-m}$$
such that $P(z)$ interpolates $\exp T(z)$ at $z = \exp(2\pi k/n)$. Note that this means that $P(z)$ interpolates $G(z)$ at the same points. The theory of [3] ensures that
$$\| P(z) - P(e^{j\omega}) \|_2 \leq \| P(e^{j\omega}) - \exp T(e^{j\omega}) \|_2 \leq K R_Y$$
(3.24)
where $K$ depends on
$$\max \\inf \| T(R, e^{j\omega}) \|_2 = \epsilon(e^{j\omega}) \|_2.$$ (4.1)

This bound can be combined with that for $1 - P_1 G^{-1} |_{\infty}$ to obtain a bound for $1 - G^{-1} P_1 |_{\infty}$.

We remark that in case $n$ is even, the construction of $T(z)$ is just that of [1, p.357], where the discrete Hilbert transform is suggested as a procedure for recovering phase data. The inequality (3.22) allows us to pin down the error that may result in its use, without any prior assumption that $G(z)$ or $\log G(z)$ corresponds to a finite impulse response.

4. ERRORS IN THE DATA

To this point, we have considered the construction of a transfer function from data (i.e. real part, or magnitude at discrete points on the unit circle) which has been assumed error free. In this section, we postulate the existence of errors in this data, and consider the resulting (extra) errors which now arise between the true function and that constructed from the error—contaminated data.

We shall be particularly interested in the behaviour of the error with increasing $n$, the number of interpolating points. Most error models will permit a discrete increase in value at one point, and discrete decrease at an adjacent point. This means that as $n \rightarrow \infty$, i.e. as the spacing between adjacent points decreases, the derivative (with respect to angle around the unit circle) can approach infinity. In the light of remarks earlier in this paper and in [2], we must expect that the error in the constructed function could in some way diverge as $n \rightarrow \infty$.

Such behaviour is indeed just what we discover when we postulate arbitrary variation of (up to) $\pm$ in the values, say real part, from which an interpolation function is constructed. However, if we postulate random fluctuations which are gaussian, we do not get such unpleasant behaviour, at least when we examine the variance in the error of the constructed function. Roughly, the averaging process somehow smoothes out the difficulties caused by the derivative of the error possibly becoming unbounded as $n \rightarrow \infty$.

We shall consider this statistical approach first.

4.1 Statistical Approach to Errors

To fix ideas, let us consider the construction of a function from discrete values of its real part (a similar discussion would apply for magnitude data). To pursue a statistical approach, we might suppose that the values of $ReG(z)$ at $z = \exp(2\pi k/n)$, $k = 0, \ldots, n-1$ are perturbed by independent, zero mean gaussian random variables of variance $\sigma^2$. However, we must reflect the fact that $ReG(z)$ is necessarily even. So this suggests that we assume $G(1)$ is known to within standard deviation $\sigma$, and $ReG(\exp(2\pi k/n)$ for $k < (n/2)$ to within standard deviation $\sigma/2$ (where we exploit the fact that $ReG(\exp(2\pi k/n)$ and $ReG(\exp(2\pi n-k/n)$ take the same value.)

Let $P$ be the interpolating FIR function that would be associated with noiseless data, and let $\Pi$ be the interpolating FIR function that is associated with noisy data. Since
$$nG - \Pi |_{\infty} \leq \| G - \Pi |_{\infty} + \| \Pi - \Pi |_{\infty}$$
(4.1)

the damage done by the noisy data is reflected in $\Pi - \Pi |_{\infty}$, which is actually a random variable.

Space limitations prevent us giving the derivation, but it is not difficult to show that
$$E[\| (G - \Pi) \|_2^2] \leq \sigma^2 \| \Pi \|_2.$$ (4.2)

We see that then at any one frequency $\Pi - \Pi$ differs from

14
\( \text{G-P} \text{I} \) by a random variable of variance \( \sigma^2 \). In particular, at that frequency for which \( \text{G-P} \text{I} \) is maximum, i.e. \\
\( \text{G-P} \text{I} = \text{G-P} \text{I} \), we have \( \text{G-P} \text{I} \) differing from \\
\( \text{G-P} \text{I} \) by a random variable of variance \( \sigma^2 \). However, \\
we cannot make a firm statement about \( \text{G-P} \text{I} \).

4.2 Deterministic Approach to Error

We shall now show, with the aid of an example, that if \\
errors of up to \( \Delta \) are postulated in the value of the real \\
part of \( \text{Re}(\text{e}^{j\omega}) \) at interpolating points, \( \text{Re}(\text{i} \text{e}^{j\omega}) \) can grow \\
as large as \( \Delta \). This shows that in the presence of deterministic \\
errors, the benefit obtained from finer and finer sampling \\
(convergence in the sense of Section 2 or 3) is eventually \\
outweighed by the effect of the errors in the sampled values.

Take \( n \) as a multiple of 4, viz. \( n = 2 \times 4^n \). Consider \\
errors in \( \text{Re}(\text{e}^{j\omega}) \) with the property that \\
\[ \text{Re}(\text{e}^{j\omega}) = \Delta \text{ when } \text{Re}(\text{e}^{j\omega}) > 0 = 0 \text{ when } \text{Re}(\text{e}^{j\omega}) = 0 = -\Delta \text{ when } \text{Re}(\text{e}^{j\omega}) < 0 \]

Again space limitations prevent us presenting the 

artificial. However, with some manipulation one can show that 

\[ \text{Im}(\text{e}^{j\omega}) = \frac{4\Delta}{n} \sum_{n=1}^{n} \cot(\frac{m\pi}{4}) \tag{4.4} \]

It is not hard then to prove that 

\[ \Delta^2 \left[ \ln \left( n - 2 \right) \right] > \left| \text{Re}(\text{e}^{j\omega}) \right| > \Delta^2 \left[ \ln \left( n - 2 \right) \right] \tag{4.5} \]

Hence it is clear that \( \text{Im}(\text{e}^{j\omega}) \) grows as \( \ln n \). In 

absolute terms, the values stay modest for large values of \( n \), 

e.g. \( n = 4096 \) yields \( \text{Im}(\text{e}^{j\omega}) = 5.3752 \).

As foresawed, this example shows that the error can 
certainly diverge as fast as \( \ln n \). Whether or not there exist 
other constructions [i.e. choice of errors of which (4.3) is one example] 
such that divergence occurs faster than \( \ln n \) is an open problem. 
However, it is true that divergence occurs no faster than \( \ln n \). We know that 
\( \text{Re}(\text{Re}(\text{e}^{j\omega}) < \Delta \) 
gives, from [9, Corollary 2.1], 

\[ \text{Im}(\text{Re}(\text{e}^{j\omega})) \leq \Delta (1 + 2 \ln n) \tag{4.6} \]

Now for \( n \) odd 

\[ \text{Re}(\text{e}^{j\omega}) = b_0 + \sum_{k=1}^{n} b_k \text{e}^{j\omega} + b_k \text{e}^{-j\omega} \]

\[ = b_0 + \sum_{k=1}^{n} 2b_k \cos \omega \tag{4.7} \]

Using [10, 3.5.42, p.265], if 
\[ \text{Re}(\text{Re}(\text{e}^{j\omega}) < M, \text{then} \]

\[ \text{Re}(\text{e}^{j\omega}) = b_0 + \sum_{k=1}^{n} 2b_k \text{e}^{j\omega} \]

\[ = b_0 + \sum_{k=1}^{n} 2b_k \cos \omega \tag{4.7} \]

where \( A \) is a constant. Identifying \( M \) with the right hand 
side of (4.6), then combining with (4.8) gives an \( n \log n \)

increase. A similar calculation applies for a even.

5. EXAMPLES

In this section, we present two examples of generating 
an approximate frequency response from interpolation data: 
the first example uses real part data; the second example 
operates on magnitude data. With frequency response plots 
we illustrate the salient features of each method, including 
the reduction in approximation error as the number of 
interpolation points increases.

The first example uses the stable transfer function 
\[ \text{G}(z) = \frac{1}{z^{2}+0.64} + \frac{1}{z^{2}+0.5} \]

(poles are at \( z = \pm 0.8, 0.5 \)) and we evaluate for \( n = 8 \) 
and \( b = 16 \).

\[ \text{Re}(\text{e}^{j\omega}) = 0.1, \ldots , n-1 \]

From this real part data, we construct \( P(z) \), causal and FIR 
according to equations (2.3)-(2.5). Figure 5.1 charts the 
frequency responses \( \text{Re}(\text{e}^{j\omega}) \) and \( \text{Re}(\text{e}^{j\omega}) \) of \( \text{Im}(\text{e}^{j\omega}) \) and 
\( \text{Im}(\text{e}^{j\omega}) \). Comparison of Figure 5.1 
\( (a) \) \( n = 8 \) and \( b = 8 \) shows the effect of increasing 
the number of interpolation points. For \( n = 8 \) the error 
\( \text{G}(z) = 0.8 \), for \( n = 16 \), \( \text{G}(z) = 0.96 \).

These plots illustrate other points of interest. Note that 
the real parts of \( \text{G}(z) \) and \( \text{P}(z) \) coincide at 
the equispaced interpolation points, but, as expected, the 
imaginary parts do not act likewise. Observe carefully 
the frequencies where the error \( \text{G}(z) = P(z) \) is large. 
These are not coincident with frequencies where \( \text{Re}(\text{e}^{j\omega}) \) 
is large, but rather at those frequencies where \( \text{Im}(\text{e}^{j\omega}) \) 
is large, consistent with the importance we placed on the 
error bounds of Section 3.5.

The second example illustrates the gain phase 
relationships. For this we need a stable, minimum phase 
transfer function \( G(z) \). Note that any strictly proper, rational 
transfer function \( G(z) \) has one or more zeros at infinity; 
therefore, it is non—minimum phase. For this example we 
use the stable minimum phase

\[ \text{G}(z) = (z-0.4) (z-0.8) \]

and evaluate 
\[ \text{G}(z) = \frac{1}{z^{2}+0.64} + \frac{1}{z^{2}+0.5} \]

From this magnitude data we construct \( P(z) \) according to 
either "method 1" or "method 2" outlined in Figures 3.1 and 
3.2 respectively. Figure 5.2 plots the results for \( n = 8 \); 
Figure 5.3 for \( n = 10 \).

Similar observations to the real—imaginary case apply here: 
the regularly spaced interpolation of magnitude plots; 
decreasing error \( \text{G-P} \text{I} = \text{I} \) as \( n \) increases from \( 8 \) to \( 10 \); the 
error \( \text{G}(z) = \text{P}(z) \) being large at frequencies where 
the magnitude has large derivative. Contrasting method 1, which 
relies on spectral factorization, and method 2, based on the 
logarithm, shows that the former gives a smaller error. 
Whether this is true for general \( G(z) \) is not apparent from 
the error bound calculations of Sections 3.3 and 3.4. 
Moreover, Figures 5.2 and 5.3 show the additive error 
\( \text{G-P} \text{I} \); for multiplicative error \( (1-P/G) \), the two methods 
may rank in the reverse order.

The two examples of this section are in qualitative 
agreement with the theoretical calculations of preceding 
sections. (Very lengthy calculations are required to check 
quantitative agreement and hardly seem justified). The 
examples illustrate that a good approximation to a frequency 
response given only real part interpolation data or magnitude 
interpolation data is indeed possible, provided the underlying 
transfer function is stable, or stable and minimum phase, and 
sufficient interpolation points are taken. The error 
phenomena associated with imprecise data and described 
in the previous section generally only become of concern 
for larger values of \( n \), far larger in fact than the values 
required in these examples to secure good approximation.

6. CONCLUSIONS

Despite the discontinuity questions associated with use of 
the Hilbert transform, we have been able to clarify the 
circumstances under which Hilbert transform type calculations 
can be used to recover a transfer function from discrete 
frequency data for its real part, imaginary part, magnitude, 
or phase.

With a finite number of points, it is never possible to 
recover the transfer function exactly, unless the class of 
considered transfer functions is heavily restricted a priori 
e.g. to being FIR. Instead, we can hope to recover an FIR 
function which is close to the desired transfer function, with 
the error going to zero as the number of data points 
becomes infinite.

There however a practical limit to this result imposed 
by data errors. For the effect of data errors on the 
determined FIR transfer function becomes greater as the 
number of data points increases, at least for certain types 
of error model.
There are a number of further questions which future work could examine. One could seek to extend the results of the paper to matrix transfer functions. The connection between real and imaginary part has nothing new in it; the connection between spectrum matrix and minimum phase stable spectral factors (the matrix analog of the magnitude to transfer function calculation) is somewhat different, since one lacks the ability to work with Hilbert transforms of the real and imaginary part of the logarithm of the spectrum.

Another direction of generalization would involve changing the postulates about the spacing of discrete frequencies. One could be interested in working with information over only a limited part of the unit circle, or with frequencies which were the bilinear transforms of (continuous time) frequencies uniformly spaced in a logarithmic sense, or with frequencies which are spaced more densely in the neighborhood of high derivatives of the data.

A third direction of generalization would involve using discrete data to define a transfer function which was not FIR, for example IIR with prescribed poles, or all-pole. Besides explaining how to construct such a function from the data, one would have to establish a bound between the constructed transfer function and the class of all transfer functions consistent with the data and establish that the bound went to zero as the number of data points became infinitely dense.

Finally, we note that we have not established the worst possible behavior of the error in a Hilbert transform induced by (deterministic) bounded errors in the data.

REFERENCES


Figure 5.2 Frequency responses for the magnitude data example, \( n=8 \). Solid line \( G \), broken line \( P \), in the magnitude and phase plots.

Figure 5.3 Frequency responses for the magnitude data example, \( n=10 \). Solid line \( G \), broken line \( P \), in the magnitude and phase plots.