SEVERAL APPROACHES TO SIGNAL RECONSTRUCTION FROM SPECTRUM MAGNITUDES

Adrian Burian, Jukka Saarinen, Pauli Kuosmanen

Tampere University of Technology
P.O. BOX 553, FIN-33101, Finland

Corneliu Rusu

Technical University of Cluj-Napoca
Baritiu 26-28, Romania

ABSTRACT

The problem of reconstructing a one-dimensional (1-D) signal from only the magnitude of its Fourier transform emerges when the phase of a signal is apparently lost or impractical to measure. Previous solutions usually employed an Iterative Fourier Transform (IFT) algorithm applied on a discrete approximation of a signal. The utilization of these algorithms is seriously limited by the unpredictability of their convergence. We propose several solutions to the phase retrieval problem. The first two proposed solutions uses relationships between the phase and the gain differences (GD), or gain samples (GS), in nepers. The last proposed solution uses a neural network (NN) for solving the problem. The NN incorporates a combination of the maximum entropy estimation algorithm with some additional nonlinear constraints. We compare our solutions by using some numerical examples. The performances under noisy conditions are also considered.

1. INTRODUCTION

The phase retrieval problem is associated with applications in which the wave phase is apparently lost or impractical to measure, and only intensity data are available [1][2]. These applications include antenna design, filter design, image reconstruction, wavefront sensing, and electron microscopy. Solutions of the phase retrieval problem are of two types: solutions depending on analytical properties and solutions depending on numerical procedures. The analytical solutions are usually related with the logarithmic Hilbert transform. Customary solution techniques approximate the solution by discretizing the continuous problem.

The 1-D phase retrieval problem is to reconstruct the phase given the modulus of its Fourier transform. Equivalently, we synthesize the signal by applying the inverse Fourier transform. The difficulty of the problem consists in the fact that if the magnitude of the Fourier transform of a function of finite support is known, then you cannot find the function if the function and the Fourier transform are 1-D. This is due to the non-uniqueness of the function. More information about the signal is needed to be known (for example that it is of minimum phase).

The discrete phase retrieval problem is to reconstruct a discrete time signal with known and compact support (N) from the magnitude of its discrete Fourier transform. There are 2^N solutions to the problem since the zeros of the z transform of the autocorrelation function occur in reciprocal conjugate quadruples [1]. The pair inside the unit circle or the one outside it may be chosen. The supplementary information, necessary to derive a unique solution, can consist of some time samples of the signal. For example, a single endpoint specifies a unique solution, the set of exceptions having the Lebesque measure zero.

There are two types of the ambiguities for the problem: trivial and non-trivial. The trivial ones include constant scale factors and translations: if $x(n)$ is a solution, then $x^*(n)$, $cx(n)$, $x(n-b)$ are also solutions for any integer b and any complex c with $|c| = 1$. Beyond these trivial factors, the ambiguity of the problem remains, because the complex zeros of $A^2(s)$ occur in conjugate pairs. For example, if $A^2(s)$ has N complex conjugate pairs of zeros, then there exist 2^N non-trivial solutions, because there are two ways of choosing one zero from each of the N conjugate pairs.

2. MINIMUM AND NON-MINIMUM PHASE RETRIEVAL

If $R(\omega)$ and $I(\omega)$ are the real and respectively the imaginary parts of an analytic function (its Fourier transform vanishes for $\omega < 0$), then $I(\omega)$ can be uniquely determined from $R(\omega)$ [3]. The phase is also the imaginary part of the natural logarithm of $X(\omega)$: $\ln X(\omega) = \ln A(\omega) + j \cdot \phi(\omega)$, but $\ln X(p)$ is not, in general, analytic for Re $p \geq 0$. $X(p)$ might have zeros with positive real parts, and those zeros are singularity points of $\ln X(p)$. If we assume that $X(p)$ is analytic and has no zeros for Re $p \geq 0$, then $\ln X(p)$ will be also analytic in the right half plane and the phase retrieval problem has a unique solution. The class of functions with this property is called minimum phase. A minimum phase system $H(s)$ has a causal impulse response $h(t)$. For these systems, the log-magnitude and phase functions of the frequency response form a Hilbert transform pair.

If $X(s)$ is not minimum phase, more information is needed in order to produce a unique solution. Non-negativity is usually not enough, so values of $x(t)$ or $X(s)$ are needed to be known. We will use the fact that $X(s)$ is completely characterized by its complex poles and zeros. If $X_{\text{MIN}}(s)$ is the minimum phase solution, then the non-minimum phase solution $X(s)$ can be written as a product of $X_{\text{MIN}}(s)$ and an all-pass term $X_a(s)$:

$$X(s) = X_{\text{MIN}}(s) X_a(s), \tag{1}$$

where $|X(s)| = |X_{\text{MIN}}(s)| = A(s)$ and $|X_a(s)| = 1$. This can be done by selecting for $X_a(s)$, a function whose zeros are the zeros of $X(s)$ in the right half plane and whose poles are symmetrical to the zeros with respect to the imaginary axis. So, for $s = j\omega$, $X_a(s)$ will have the form:

$$X_a(s) = \frac{\prod_{i=1}^{k} (s + s_i)}{\prod_{i=1}^{k} (s - s_i)}, \tag{2}$$

where k represent the number of zeros of the $X_{\text{MIN}}(s)$ corresponding to those k right half plane zeros of $X(s)$ flipped into the open left half plane. The corresponding non-minimum phase function can be written as:

$$\phi(\omega) = \phi_{\text{MIN}}(\omega) + \phi_a(\omega), \tag{3}$$

where

$$\phi_a(\omega) = 2\sum_{i=1}^{k} \arctan\left(\frac{\omega + \text{Im}(s_i)}{\text{Re}(s_i)}\right). \tag{4}$$

From (1) and (2), by denoting the coefficients of s in the term $\prod_{i=1}^{k}(s+s_i)$ with x_F^i we obtain:

$$\left(\sum_{i=0}^{k}(-1)^{k-i}x_F^i s^i\right)X(s) = \left(\sum_{i=0}^{k}x_F^i s^i\right)X_{\text{MIN}}(s), \tag{5}$$

where $x_F^k = 1$ and $x_F^0 = \prod_{i=1}^{k} s_i$. Without loss of generality we can assume that the initial conditions in the time domain are zero. The following differential equation is obtained:

$$\sum_{i=0}^{k}(-1)^{k-i}x_F^i \frac{\mathrm{d}^i x(t)}{\mathrm{d}t^i} = \sum_{i=0}^{k} x_F^i \frac{\mathrm{d}^i x_{\text{MIN}}(t)}{\mathrm{d}t^i}. \tag{6}$$

We used the following procedure to compute the non-minimum phase solution, from the minimum phase solution:

A. Given $|X(s)|$ compute the minimum phase solution ϕ_{MIN};

B. With the additional information of the number of zeros to flip and their type (real or complex) and the support of $x(t)$,, calculate the flip coefficients by using formula (6). Then calculate the phase of the all-pass term signal by using formula (4);

C. Compute $X(s) = |X(s)| \cdot \exp(j\cdot\phi)$;

D. Synthesize $x(t)$ by applying the inverse Fourier transform.

3. GAIN DIFFERENCES

The first method we developed uses of a relationship between the phase and the odd derivatives of the gain for computing the phase of the minimum phase solution [4]. We showed there that the phase at a given frequency is the series of the odd derivatives of the neperian gain evaluated at this frequency. In order to substitute the higher derivatives involved, we utilized finite differences.

The gain and the phase can be written as functions of variable u, without loss of generality: $a(u) = A(\omega_c e^u)$, $b(u) = \phi(\omega_c e^u)$. We proved in [4] that there exists the following relation between the phase and the gain differences for the phase of a minimum phase function:

$$b(u) = \sum_{n=0}^{\infty}\frac{2(2^{n+2}-1)\pi^{2n+1}|B_{2n+2}|}{(2n+2)!}a^{2n+1}(u), \tag{7}$$

where B_n are the Bernoulli numbers of order n.

The error of truncation is propagated through the frequency domain and the Gibbs phenomenon appears. In order to avoid it, we used the Feher kernel, i.e, we passed the weights of the derivatives through a triangular window. However, the practical problems give only the gain samples, and there is necessary to approximate the higher derivatives with differences. Using the Stirling numbers of the first kind, we have:

$$\frac{d^k y(a_0)}{dx^k} = \frac{k!}{h^k}\left(\sum_{z=k}^{n}\frac{S_z^{(k)}}{z!}\Delta^i f_0\right). \tag{8}$$

Let k be the number of terms in approximation. By using the Feher kernel formula, the following approach is obtained:

$$b_{F1} = \frac{\pi}{2}a'(0) \cong \frac{\pi}{2h}\Delta^1 f_0; \tag{9}$$

$$b_{F2} = \frac{\pi}{2h}\left(\Delta^1 f_0 - \frac{1}{2}\Delta^2 f_0 + \frac{1}{3}\Delta^3 f_0\right) + \frac{\pi^3}{48h^3}\Delta^3 f_0; \cdots \tag{10}$$

4. GAIN SAMPLING

In [5] we established a relationship for approximating the phase values from the gain samples in nepers for computing the phase of the minimum phase solution:

$$\begin{aligned} B(\omega) = {} & \frac{1}{\pi}[A(\omega\delta) - A(\omega/\delta)] + \\ & + \frac{2\ln\delta}{\pi}\int_1^k \frac{A(\omega\delta^z) - A(\omega\delta^{-z})}{\delta^z - \delta^{-z}}dz, \end{aligned} \tag{11}$$

where $\delta > 1$ and $k \in N, k \geq 1$ satisfy certain conditions.

For numerical computation support, it is of interest to develop a quadrature formula. The condition of equally spaced abscissas in the logarithmic frequency domain, leads to one of the Newton-Cotes or Simpson's quadrature formulae. By selecting the trapezoidal formula we have obtained the first approximation $B_T(\omega)$ for the phase:

$$B_T(\omega) = \sum_{p\in Z} T_p A(\omega\delta^p), \tag{12}$$

$$T_p = T_{-p} = \begin{cases} \frac{1}{\pi}(1 + \frac{\ln\delta}{\delta - 1/\delta}) & \text{if } p = 1 \\ \frac{2\ln\delta}{\pi(\delta^p - \delta^{-p})} & \text{if } p = 2,\ldots,k-1 \\ \frac{\ln\delta}{\pi(\delta^p - \delta^{-p})} & \text{if } p = k \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

By using the parabolic rule, for $k = 2m+1, m \geq 0$, a second quadrature approach $B_S(\omega)$ can be obtained:

$$B_S(\omega) = \sum_{p\in Z} S_p A(\omega\Delta^p), \tag{14}$$

$$S_p = S_{-p} = \begin{cases} \frac{1}{\pi}\left(1 + \frac{2/3\ln\delta}{\delta - 1/\delta}\right), & p = 1; \\ \frac{8\ln\delta}{3\pi\left(\delta^p - \delta^{-p}\right)}, & p = 2,4,\ldots,2m; \\ \frac{4\ln\delta}{3\pi\left(\delta^p - \delta^{-p}\right)}, & p = 3,5,\ldots,2m-1; \\ \frac{2\ln\delta}{3\pi\left(\delta^p - \delta^{-p}\right)}, & p = 2m+1; \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

It seems from the previous estimations that the two quadrature formulae are comparable in performance according to the number of samples. However, the multiplying constants and the level of the samples considered differ by much. By using the Simpson's formula the higher frequencies are enhanced, while the trapezoid rule reduces them. If we compare with the sums of GD, the convergence of the GS series is faster. Also, the numerical evaluations of the higher derivatives are likely to have sizeable errors. So, we expect that this approach outperform the previous one. The disadvantage consists in an increased computational effort.

5. NEURAL NETWORKS

We separate the real and imaginary parts of the spectrum. Let us denote by M_k the known spectral magnitudes, and by a_{nk} and b_{nk} the constants related to discrete Fourier transform: $a_{nk} = \cos(2\pi(n+k)/N)$, $b_{nk} = -\sin(2\pi(n+k)/N)$. The unknown phase is denoted by ϕ_k. The unknown nonnegative sequence $x(m)$ is denoted by x_m. The following equations are obtained:

$$M_k \cos\phi_k = \sum_m x_m a_m^k,\ M_k \sin\phi_k = \sum_m x_m b_m^k. \tag{16}$$

Let us consider the following nonnegative variables: $c_k = 1 + \cos\phi_k$, $s_k = 1 + \sin\phi_k$. By using the above definitions, the phase retrieval problem can be written as a nonlinear optimization problem, by the way of the maximum entropy restoration [6][7]:

$$\text{find max}\left\{-\sum_m x_m \ln(x_m) - \sum_k [c_k \ln(c_k) + s_k \ln(s_k)]\right\}, \tag{17}$$

subject to

$$\text{a) } x_m, c_k, s_k \geq 0; \tag{18}$$

$$\text{b) } \sum_m x_m a_m^k - M_k c_k = -M_k; \tag{19}$$

$$\text{c) } \sum_m x_m b_m^k - M_k s_k = -M_k; \tag{20}$$

with normalization:

$$(c_k - 1)^2 + (s_k - 1)^2 = 1. \tag{21}$$

We denote $f(\mathbf{y}) = -\sum_m x_m \ln(x_m) - \sum_k [c_k \ln(c_k) + s_k \ln(s_k)]$.

The maximum entropy method is well described in the literature, but the standard algorithm has the disadvantage of computational inefficiency. Since an artificial neural network has a strong computational capability, it can be used to solve these problems [8]. If we consider a neural network with fully interconnected neurons, an energy function *U* and an entropy function *S* can be associated with it [9]. For a positive convex entropy function, and for any non-decreasing temperature sequence *T*, the neural network admits a Lyapunov function, which is the Helmholtz free energy of the system: $F = U - TS$. The neural network will evolve in time until it reaches an equilibrium state that corresponds to a minimum of the free energy function *F*, which simultaneously minimizes the energy and maximizes the entropy.

The nonlinear, sigmoidal transfer function that determines the relation between an input $v_i = M_i$ and an output y_i (which includes all x_i, c_i and s_i) is given by:

$$y_i = 1 + \tanh\left(\frac{v_i - v_0}{g}\right). \tag{22}$$

The slope of the transfer function at the inflection point $v_i = v_0$ (where v_0 is the offset) constitutes the maximum gain of the amplifier in the practical realization, and is given by:

$$\lambda = \frac{\mathrm{d}y_i}{\mathrm{d}v_i} = \frac{1}{g}. \tag{23}$$

In order to obtain the phase normalization, the relation (16) is modified as follows:

$$y_{ic,S} = 1 + \frac{\tanh\left(\frac{v_{iC,S} - v_0}{g}\right)}{\sqrt{\tanh^2\left(\frac{v_{iC} - v_0}{g}\right) + \tanh^2\left(\frac{v_{iS} - v_0}{g}\right)}}. \tag{24}$$

We encoded the constraints into function U in the following way:

$$U(\mathbf{v}) = \sum_i \mu_i |[\mathbf{a}\ \mathbf{b}]^T \mathbf{v} + v_i|^2, \tag{25}$$

where μ_i is a scaling parameter. We obtained:

$$U(\mathbf{v}) = \sum_i \mu_i |\sum_j [a_{ij}\ b_{ij}]^T v_j + v_i|^2. \tag{26}$$

The time evolution of the neural network is described by:

$$\frac{\mathrm{d}y_i}{\mathrm{d}t} = -\lambda\left[[r_i\ q_i]^T + \frac{1}{T}\frac{\partial U}{\partial v_i}\right], \tag{27}$$

where $r_i = \mathbf{x}\,\mathbf{a}^i - v_i c_i + v_i$ and $q_i = \mathbf{x}\,\mathbf{b}^i - v_i s_i + v_i$ are the residuals of the neural network. The steady state of the neural network with the nonlinear constraints satisfied, will provide the solution for the phase retrieval problem.

6. SIMULATION RESULTS

The proposed solutions to the phase retrieval problem will be compared for several discrete noisy situations. We examined a signal similar with the one used in [2]:

$$x_{\text{MIN}}(t) = \frac{\omega_m + \omega_n}{\omega_m - \omega_n} \cdot [\exp(-\omega_n t) - \exp(-\omega_m t)] \cdot u(t), \tag{28}$$

with $\omega_m = s_2 > \omega_n = s_1 > 0$, which has two poles in the complex plane:

$$X_{\text{MIN}}(s) = \frac{\omega_n + \omega_m}{(s + \omega_n)(s + \omega_m)} = \frac{s_1 + s_2}{(s + s_1)(s + s_2)}. \tag{29}$$

Three non-minimum phase signals can be obtained by flipping either one or both of these poles. When the pole s_2 is flipped, the following non-minimum phase signal is obtained:

$$x(t) = \exp(\omega_m t) \cdot u(-t) + \exp(-\omega_n t) \cdot u(t). \tag{30}$$

By using a direct calculus it can be shown that for these 4 signals, the entropy of this solution is the biggest one.

The signal (30) has an infinite support on the entire real line, but there exist $t_0 > 0$ such that this signal is essentially zero for $|t| > t_0$. So, we can treat this signal as a compact one. The same is true for the corresponding minimum phase signal on the positive real line, which goes to zero slower than the non-minimum phase signal. The GD and GS methods will give the minimum phase solution. In order to find the maximum entropy solution, we applied the steps presented in Section 2. The non-minimum phase solution which has the maximum entropy was obtained by flipping the pole s_2 of the minimum phase solution. The flip coefficient was computed by succesively applying formula (6) for $k = 1$ and $k = 2$ in the corresponding 'smooth' subinterval $[1.3, 1.5]$. In this subinterval, $x(t)$ was considered to be equal to zero while $x_{\text{MIN}}(t)$ was not. In this situation, we have:

$$s_1 x_{\text{MIN}} + \frac{\mathrm{d}x_{\text{MIN}}}{\mathrm{d}t} = 0. \text{ So, } s_1 = \frac{\omega_n - \omega_m \exp[(\omega_n - \omega_m)t]}{1 - \exp[(\omega_n - \omega_m)t]} \approx \omega_n, \tag{31}$$

$$s_1 s_2 x_{\text{MIN}} + (s_1 + s_2)\frac{\mathrm{d}x_{\text{MIN}}}{\mathrm{d}t} + \frac{\mathrm{d}^2 x_{\text{MIN}}}{\mathrm{d}t^2} = 0 => s_2 = \omega_m. \tag{32}$$

In order to avoid the Gibbs phenomenon in the case of GD method, the Feher kernel was used. The only first two terms were

Table 1. Signal reconstruction performances.

SNR	GD MSE	$\times 10^{-4}$ MAE	GS MSE	$\times 10^{-4}$ MAE	NN MSE	$\times 10^{-4}$ MAE
Noiseless	114	8010	178	623	83	450
70 dB	523	1630	196	599	51	355
60 dB	427	1598	199	762	63	407
50 dB	758	2425	655	1291	69	357
40 dB	1551	1762	910	2141	73	433
30 dB	2148	3822	1294	2746	327	798

used in approximation. For GS method the phase of the minimum phase signal was approximated by using the Newton-Cotes approach for $\Delta = \sqrt{2}$ and $k = 5$. In [4] we gave a detailed analysis of the effect of distance between the gain samples and of the quality of the approximated phase. For neural networks, a special attention required the temperature parameter. If this parameter is too small, the time evolution of the neural network can be chaotic (from equation (27)), and if it is too big, a very large number of iterations is necessary. The starting temperature was $T = 100$. The raising of the temperature in steps of 100 was decided by a very small decreasing value of the error ($\leq 10^{-3}$ during 10 iterations). The iterative process was stopped when this situation appeared for the maximum allowed temperature ($T_{max} = 1000$). The error was computed as follows: Err $= \sum_i |r_i| + \sum_i |q_i|$. The entire neural network's output data were scaled between 0 and 2. This was done in order to encourage equitable distribution of importance.

The computed MSE and MAE for all the proposed approaches in different noisy situations are presented in Table 1, for $\omega_m = 20$ and $\omega_n = 4$. For an easier comparison, all the results were multiplied by 10^4. Plots of the actual signals (continuous line) versus the reconstructed ones (dashed line) are presented in Figure 1 for noiseless and SNR=30dB situations. The GD solutions are presented in Figure 1a)b), the GS solutions in c)d) and NN solutions in e)f). The length of the signals was 512. The obtained results were similar when the poles were close together ($\omega_m - \omega_n < 5$), or when they were far away ($\omega_m - \omega_n > 20$).

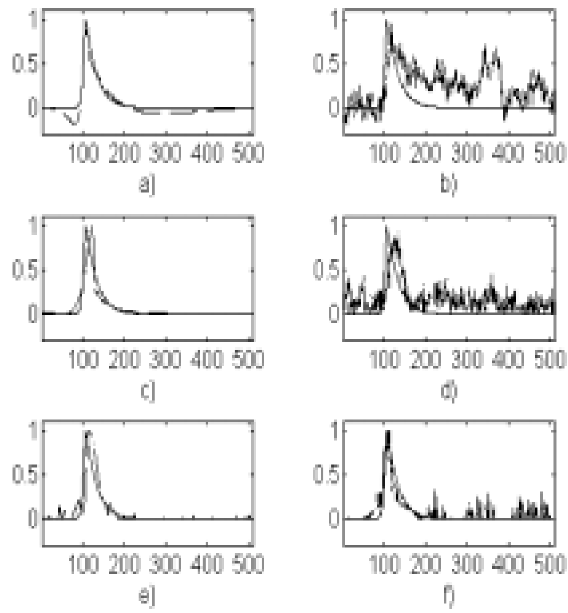

Fig. 1. Reconstruction solutions for noiseless a)c)e) and for SNR=30dB b)d)f).

7. CONCLUSIONS

We have proposed three solutions to the problem of phase retrieval. The first two solutions are not iterative, very simple and stable, so the convergence problem does not exist anymore. Better results are obtained by the GS method, but the use of GD has less computational complexity. The computer simulation results indicate that both approaches gives good results in noisy conditions with Medium SNRs($>$ 30 dB). If the number of samples is small (≤ 64), or when the SNR is low, the higher derivatives could have sizeable errors and the use of GD method is unsuitable. The third solution is iterative and uses a recurrent neural network for positive sequence reconstruction. This is the most complex solution, but it has the best performances.

8. REFERENCES

[1] L.S. Taylor, "The Phase Retrieval Problem", *IEEE Trans. on Antennas and Propagation*, March 1981, pp. 386-391.

[2] A.E. Bell, A.E. Yagle, "1-D Continuous Non-Minimum Phase Retrieval using the Wavelet Transform", *IEEE International Conference on Acoustic, Speech and Signal Processing*, Seattle, Washington, May 1998, pp. III 1641-1644.

[3] A. Papoulis, *The Fourier integral and its applications*, McGraw-Hill, 1962.

[4] A. Burian, P. Kuosmanen, C. Rusu, "1-D Non-Minimum Phase Retrieval by Gain Differences", *IEEE International Conference on Electronics, Circuits and System*, Paphos, Cyprus, September 1999, pp. 1001-1005.

[5] C. Rusu, P. Kuosmanen, "Logarithmic Sampling of Gain and Phase Approximation", *IEEE International Symposium on Circuits and Systems*, Orlando, Florida, June 1999, pp. VI 41-44.

[6] A.T. Bajkova, "Applications of the new Generalized form of Maximum Entropy Method to Solving Inverse Problems", *IEEE International Symposium on Antennas and Propagation*, vol.1, 1992, pp. 1695-1698.

[7] A. Burian, P. Kuosmanen, J. Saarinen, C. Rusu, "Two Dimensional Phase Retrieval using Neural Networks", *IEEE International Workshop on Neural Networks for Signal Processing*, Sydney, December 2000, pp. 652-662.

[8] A. Cichocki, R. Unbehauen, *Neural Networks for Optimization and Signal Processing*, John Wiley & Sons, 1993.

[9] D. Ingman, Y. Merlis, "Local minimum escape using thermodynamic properties of neural network", *Neural Networks*, vol.4, no. 3, 1991, pp. 395-404.