Res. Lett. Inf. Math. Sci., 2004, Vol. **6**, pp 59-69
Available online at http://iims.massey.ac.nz/research/letters/

Analysis of a non-minimum phase acoustic beamformer

T.J.MOIR

Institute of Information & Mathematical Sciences
Massey University at Albany, Auckland, New Zealand

t.j.moir@massey.ac.nz

The two input Griffiths-Jim acoustic beamformer is analysed in the frequency domain using a Wiener type formulation. Unlike previous solutions the approach here is to look at the problem of non-minimum phase acoustic transfer functions which are encountered in many real filtering problems. The polynomial transfer function approach gives an elegant way of obtaining the frequency response of the beamformer and gives new insight to the problem in general.

1 Introduction

The problem of suppression of noise from speech has applications in such areas as speech recognition, tele-conferencing, hearing aids and hands-free telephony in automobiles. The basic idea of canceling noise from speech has been around for decades[1] and normally involves the use of at least two microphones. One microphone is placed near the noise source and the second microphone is placed such as to pick up the speech and the noise. An adaptive filter based on the least-mean-squares (LMS) method is then used to minimise the mean-squared-error and arrive at an adaptive tracking Wiener filter. It is well known however that in order for such a system to work effectively requires good coherence between the two microphones and this necessitates the microphones not being too far apart. Conversely if the microphones are too close together then the speech as well as the noise will be cancelled. Although there are certain applications in special environments which are suitable to such a solution it is generally accepted than modifications of this fundamental idea is necessary in many real environments.
In fact many of the above real-world problems require the microphones to be close together rather than far apart. Such a situation arises naturally in beamforming. There are many different types of beamformer, for example the delay and sum beamformer which steers the directivity towards a sound source[2]. The approach used here wishes to simplify the procedure as much as possible for future real-time implementation and so two microphones are used as an adaptive Griffiths-Jim beamformer [3] as shown in Figure 1 below. Such an approach (the generalized sidelobe canceller) steers a null towards the source of the noise and has found applications in hearing aid and speech recognition research [4,5,6]. Referring to Figure 1,the two signals from the microphones are fed to the (simplified) beam-steerer which consists of two delays τ_1, τ_1 and two scaling factors a_1, a_2 which compensate for any path difference of the speech and make the speech appear as if it is directly in front of the two microphones. However, for the purposes of this analysis we consider the speech to be already in front of the two microphones so that $a_1 = a_2 =1$ and $\tau_1 = \tau_2 =1$. Speaking directly in front of the microphones would be the case for a hearing aid where the wearer will turn and face the speaker. There are other possibilities, for example in distant talker speech recognition [7] where this will not be the case but for such problems it will be assumed that steering has already taken place and that the speech is again assumed to be directly in front of the microphones at least in a virtual sense. For such a problem the delays τ_1, τ_1 can no longer be taken to be unity as they will compensate for a real acoustic delay known as the time-difference of arrival (TDOA). However, after the beamformer has compensated for the TDOA, the net effect will be that the

speech will again be directly in front of the two microphones as the overall delay from the speech source will be the same in both the upper and lower arm of the beamformer.

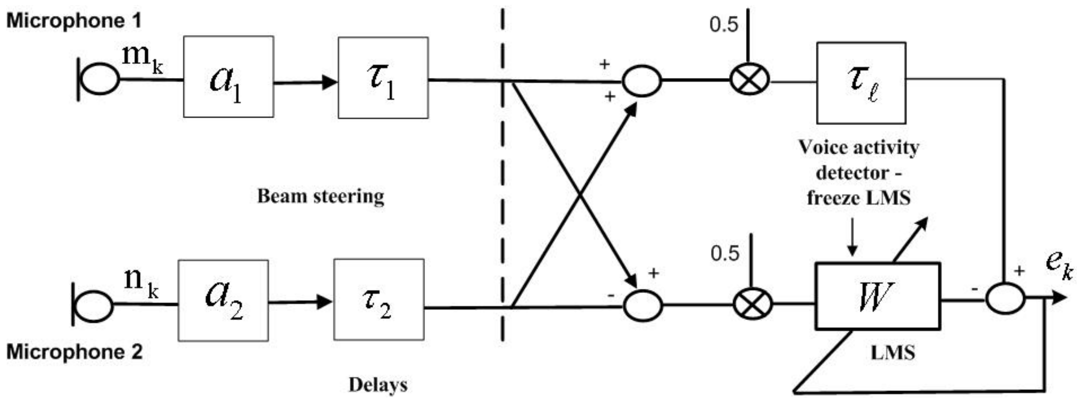

Figure 1 Two microphone Griffiths-Jim beamformer

In Figure 1 there is another delay τ_ℓ which is included to improve the performance of the beamformer. Its purpose becomes clearer later in this paper. The commonly used least-mean squared (LMS) algorithm [8] is used to minimize the mean-squared error using a finite impulse response (FIR). The top path of the beamformer has a summation term which forms the primary input whilst the bottom path has a difference term which forms the reference input. No blocking matrix is required as the subtraction itself is enough to get rid of the speech so that noise alone is fed into the reference. However, in real problems where there may not be exact steering (that is the signal does not appear exactly in phase after steering) then there may well be a small undesired signal component appearing in the reference. To avoid the LMS algorithm adapting during this time it is usual to include a voice activity detector (VAD) which only updates the LMS algorithm during periods of noise. Fortunately such a VAD has been reported in [9] and found to be robust in realistic environments.

2 Optimal beamformer

Mathematical preliminaries:

If a polynomial defined as $X(z^{-1}) = x_0 + x_1 z^{-1} + x_2 z^{-2} + \ldots + x_n z^{-n}$ of degree n has all its roots within the unit circle in the z plane, then it is termed minimum phase. No zeros are assumed to be on the unit circle. For simplicity $X(z^{-1})$ is often written as X, omitting the complex argument z^{-1}. The conjugate polynomial $X^*(z) = x_0 + x_1 z + x_2 z^2 + \ldots + x_n z^n$ is non-minumum phase having all of its roots outside the unit circle on the z plane. The reciprocal polynomial is defined as $\tilde{X}(z^{-1}) = x_0 z^{-n} + x_1 z^{-(n-1)} + x_2 z^{-(n-2)} + \ldots + x_n = z^{-n} X^*(z)$ which has all its roots outside the unit circle provided $X(z^{-1})$ is minimum phase. The zeros of $\tilde{X}$ are the zeros of X reflected in the unit circle. Similarly $\tilde{X}^*(z^{-1}) = z^n X(z^{-1})$ has all its roots within the unit circle. For polynomials which are non-minimum phase, we can factorise $X(z^{-1}) = X_1(z^{-1}) X_2(z^{-1})$ where X_1 is minimum phase and X_2 is non-minimum phase.

Beamformer equations:

For the purposes of the analysis all signals are assumed to be stationary but practically this will not be the case as the LMS algorithm will track any time-variation in the parameters. Consider a stationary signal s_k and a noise source η_k. Without loss of generality it can be assumed that the noise source is zero-mean and white, with variance σ_η^2. The white noise source will be coloured by two finite-impulse response (FIR) acoustic transfer functions $H_1(z^{-1})$ and $H_2(z^{-1})$ so that the received signals at the two microphones are

$$\mathrm{m_k} = \mathrm{s_k} + \mathrm{H_1}\eta_\mathrm{k} \quad (1)$$

and

$$\mathrm{n_k} = \mathrm{s_k} + \mathrm{H_2}\eta_\mathrm{k} \quad (2)$$

For a beamformer filter W, the error signal becomes

$$e_k = z^{-\ell}[s_k + 0.5(H_1 + H_2)\eta_k] - 0.5W(H_1 - H_2)\eta_k \quad (3)$$

where $z^{-\ell}$ is an artificial delay introduced to provide causality for non-minimum phase acoustic transfer functions. The z-transform spectral density of the error becomes:

$$\begin{aligned}\Phi_{ee} &= \Phi_{ss} + 0.25(H_1 + H_2)(H_1 + H_2)^* \sigma_\eta^2 - 0.25W^*(H_1 - H_2)(H_1 + H_2)^* z^\ell \sigma_\eta^2 \\ &-0.25W^*(H_1 + H_2)(H_1 - H_2)^* z^{-\ell} \sigma_\eta^2 + 0.25WW^*(H_1 - H_2)(H_1 - H_2)^* \sigma_\eta^2 \end{aligned} \quad (4)$$

and it has been assumed that the signal s_k and noise η_k are uncorrelated.
Parseval's formula gives the mean-square error for an ergodic system as

$$E[e_k^2] = \frac{1}{2\pi j} \oint_{|z|=1} \Phi_{ee}(z^{-1}) \frac{dz}{z} \quad (5)$$

from which the optimal filter W can be obtained by completing the squares[10]. This results in

$$E[e_k^2] = \frac{1}{2\pi j} \oint_{|z|=1} \frac{1}{4}\{[W(H_1 - H_2) - z^{-\ell}(H_1 + H_2)][W(H_1 - H_2) - z^{-\ell}(H_1 + H_2)]^* \sigma_\eta^2 + \Phi_{ss}\} \frac{dz}{z} \quad (6)$$

The minimum phase beamformer transfer function then follows [4]

$$W = \frac{z^{-\ell}(H_1 + H_2)}{H_1 - H_2} \quad (7)$$

Substituting (7) into (4) gives an error signal as noise-free delayed speech

$$e_k = s_{k-\ell} \quad (8)$$

with minimum mean-squared error given by

$$\sigma_e^2 = \frac{1}{2\pi j} \oint_{|z|=1} \Phi_{ss} \frac{dz}{z} \quad (9)$$

However, it is well established that the acoustic impulse response of real environments have non-minimum phase terms [11,12] and hence (7) could very well turn out to unstable. This is because $H_1 - H_2$ may well have both zeros inside and outside of the unit circle. In [12] it has been shown that for small rooms, the impulse response is minimum phase only for reflection coefficients below about 0.37. However, the LMS algorithm will normally still converge provided the artificial delay ℓ is large enough but it will not converge to the above transfer function (7) (its power series FIR expanded form) but rather to the transfer function discussed in the next section. It can be seen that ℓ does not contribute anything to the performance of the beamformer for the minimum phase case other than introduce a time-delay.

Non-minimum phase solution:

For the general non-minimum phase beamformer problem it is necessary to further refine the definitions of the acoustic transfer functions. Suppose H_1 and H_2 are defined in terms of two further FIR transfer functions with pure delays

$$H_1 = Lz^{-d1} \tag{10}$$

$$H_2 = Mz^{-d2} \tag{11}$$

where d2>d1. Then $H_1 + H_2 = z^{-d1}[L + Mz^{-(d2-d1)}]$ and $H_1 - H_2 = z^{-d1}[L - Mz^{-(d2-d1)}]$

The delay d2-d1 is the TDOA. In fact d2-d1 may well be the first of many TDOA's due to reverberations in a real environment. However, it can be seen that a delay z^{-d1} is now common to both the upper and lower arms of the beamformer and via (3) and the error spectrum (4), can be ignored in the proceeding analyses. Define the polynomials

$$C = C_1C_2 = [L + Mz^{-(d2-d1)}] \tag{12}$$

where C_1 has all its roots inside the unit circle and C_2 has all its roots outside the unit circle. Similarly define

$$B = B_1B_2 = [L - Mz^{-(d2-d1)}] \tag{13}$$

where B_1 has all its roots inside the unit circle and B_2 has all its roots outside the unit circle.

Now (6) can be written as

$$E[e_k^2] = \frac{1}{2\pi j}\oint_{|z|=1}\{\frac{1}{4}[WB_1B_2 - z^{-\ell}C_1C_2][WB_1B_2 - z^{-\ell}C_1C_2]^*\sigma_\eta^2 + \Phi_{ss}\}\frac{dz}{z} \tag{14}$$

The above includes terms in $B_2B_2^*$. Writing $B_2 = z^{-n}\tilde{B}_2^*$ and $B_2^* = z^n\tilde{B}_2$ then

$$B_2B_2^* = \tilde{B}_2\tilde{B}_2^* \tag{15}$$

Assuming B_2 is non-minimum phase, then the reciprocal polynomial $\tilde{B}_2$ will be minimum phase. Similarly B_2^* will be minimum phase whilst $\tilde{B}_2^*$ will be non-minimum phase. A similar method is used in control theory [13] and in deconvolution filters[14].

The mean-squared error (14) will now be written in an equivalent form by using (15) thus

$$E[e_k^2] = \frac{1}{2\pi j}\oint_{|z|=1}\{\frac{1}{4}[WB_1\tilde{B}_2 - z^{-\ell}C_1C_2\frac{B_2^*}{\tilde{B}_2^*}][WB_1\tilde{B}_2 - z^{-\ell}C_1C_2\frac{B_2^*}{\tilde{B}_2^*}]^*\sigma_\eta^2 + \Phi_{ss}\}\frac{dz}{z} \tag{16}$$

from which the optimal Wiener beamformer becomes (Appendix 1)

$$W = \{z^{-\ell}C_1C_2\frac{B_2^*}{\tilde{B}_2^*}\}_+\frac{1}{B_1\tilde{B}_2} \tag{17}$$

where the brackets $\{.\}_+$ denotes the inclusion of causal terms in negative powers of z (including z^0 terms).

Polynomial form:

The polynomial solution is covered in some detail in Appendix 2. It is shown that the optimal stable beamformer is given by

$$W = \frac{C_1C_2\tilde{D}_\ell + F}{B_1\tilde{B}_2} \tag{18}$$

where the reciprocal polynomial $\tilde{D}_\ell = z^{-\ell}D_\ell^*$

$$\tilde{D}_\ell = z^{-\ell}D_\ell^* = d_0z^{-\ell} + d_1z^{-\ell+1} + d_2z^{-\ell+2} + \ldots + d_\ell \tag{19}$$

and D_ℓ^* is the expansion to ℓ terms of the power series

$$\frac{B_2^*}{\tilde{B}_2^*} = d_0 + d_1z + d_2z^2 + \ldots \tag{20}$$

The polynomial F of degree r is found by the convolution (A17) making use of the polynomial E^* defined by (A14) and (A13).

For a non-minimum phase polynomial B_2 of degree nb2 the expansion (20) is easily found by a simple recursion. For example if

$$B_2 = \alpha_0 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \ldots + \alpha_{nb2} z^{-nb2} \tag{21}$$

then $d_0 = \dfrac{\alpha_0}{\alpha_{nb2}}$

$$d_i = \frac{(\alpha_i - \sum_{j=0}^{nb2-1} \alpha_{nb2-i+j} d_j)}{\alpha_{nb2}} \quad i = 1,2 \ldots nb2 \tag{22}$$

and

$$d_i = -\frac{1}{\alpha_{nb2}} \sum_{j=0}^{nb2-1} \alpha_j d_{i+j-nb2} \text{ for } i > nb2 \tag{23}$$

The enhanced speech signal is given by the error (3) which becomes

$$e_k = z^{-\ell} s_k + z^{-(\ell+d_1)} 0.5 C_1 C_2 \eta_k - z^{-d1} 0.5 W B_1 B_2 \eta_k \tag{24}$$

and using (18)

$$e_k = z^{-\ell} s_k + z^{-(\ell+d_1)} 0.5 C_1 C_2 \eta_k - z^{-d1} 0.5 \frac{B_2}{\tilde{B}_2} [C_1 C_2 \tilde{D}_\ell + F] \eta_k \tag{25}$$

For the minimum phase case we have $B_2 = 1$, $\tilde{D}_\ell = z^{-\ell}$ and $E^* = 0, F = G^* = 0$ and the error degenerates to that of (8).

3 Illustrative example

This simple example will illustrate the power of the method in showing the frequency response of the optimal beamformer. Suppose $C = C_1 C_2 = 1.5 - 4.1z^{-1}$ which is non-minimum phase with $B_1 = 1$ and $B_2 = 0.5 + 3.9z^{-1}$ (also non-minimum phase) which makes both the upper and lower arms (primary and reference inputs) of the beamformer non-minimum phase. Take the delay ℓ to be initially 2.

Form the reciprocal polynomial $\tilde{B}_2 = 0.5z^{-1} + 3.9$ and hence $\tilde{B}_2^* = 3.9 + 0.5z$. Now form the polynomial expansion

$$\frac{B_2^*}{\tilde{B}_2^*} z^{-\ell} = \frac{0.5 + 3.9z}{3.9 + 0.5z} z^{-\ell}$$

$$= (0.1282 + 0.9835z - 0.12609z^2 + 0.016166z^3 + 0.0020726z^4 \ldots) z^{-2}$$

Which results in the two polynomials

$$\tilde{D}_\ell = z^{-\ell} D_\ell^* = -0.12609 + 0.9835z^{-1} + 0.1282z^{-2}$$

and

$$E^* = 0.016166z + 0.0020726z^2 + \ldots \quad (e_0 = 0)$$

The coefficients of the polynomial $\mathrm{F} = \mathrm{f}_0 + \mathrm{f}_1 \mathrm{z}^{-1} + \mathrm{f}_2 \mathrm{z}^{-2} + \ldots + \mathrm{f}_\mathrm{r} \mathrm{z}^{-\mathrm{r}}$ are found from

$$\mathrm{f}_\mathrm{k} = \sum_{\mathrm{i}=0}^{\mathrm{ne}} \mathrm{e}_\mathrm{i}\, \mathrm{c}_{\mathrm{k}+\mathrm{i}}, \quad \mathrm{k} = 0,1,2,..\mathrm{r}$$

The above convolution was truncated to $ne = 10$ terms giving for r=0 (since the other higher order terms are negligible) $\mathrm{F} = -0.06628 + \ldots$

Hence the transfer function of the optimal beamformer is

$$W = \frac{C_1 C_2 \tilde{D}_\ell + F}{B_1 \tilde{B}_2}$$

$$= \frac{-0.25543 + 1.99235z^{-1} - 3.84z^{-2} - 0.52564z^{-3}}{3.9 + 0.5z^{-1}}$$

The LMS algorithm will converge to the FIR equivalent of the above, which is by long division

$$W \approx -0.065487 + 0.51922z^{-1} - 1.0564z^{-2} + \dots$$

This was confirmed by LMS simulation.

To further illustrate the mechanism by which the beamformer operates, consider the frequency response of the primary path $C_1 C_2 z^{-\ell}$ shown in Figure 2.

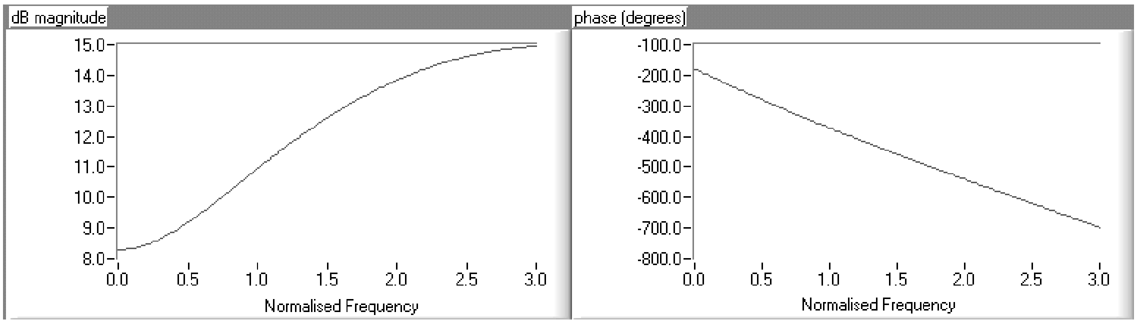

Figure 2. Frequency response of upper arm of beamformer $C_1 C_2 z^{-\ell}$

The phase is shown ‘unwrapped’ and normalized frequency is taken from dc to half sampling (3.14 radians).

For the error to be as close to zero as possible, the lower arm of the beamformer must match the frequency response in both amplitude and phase. The frequency response of the filter on its own is shown in Figure 3.

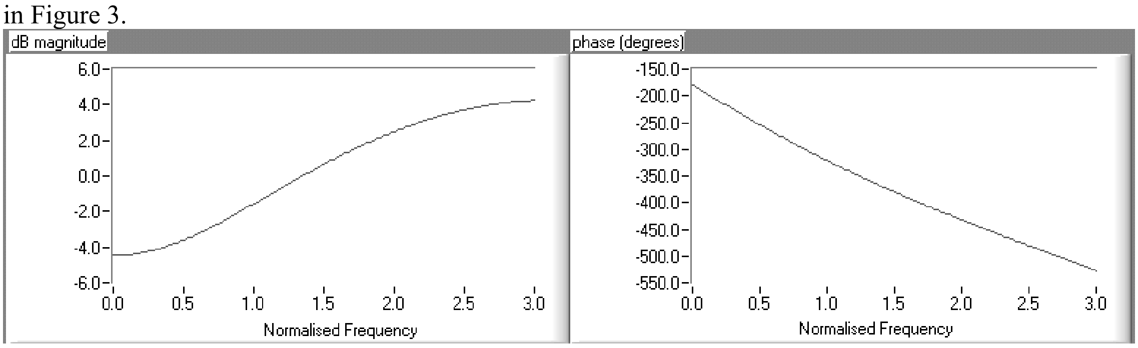

Figure 3. Frequency response of filter W

The filter convolved with the polynomial B (that is the lower arm) is shown in Figure 3.

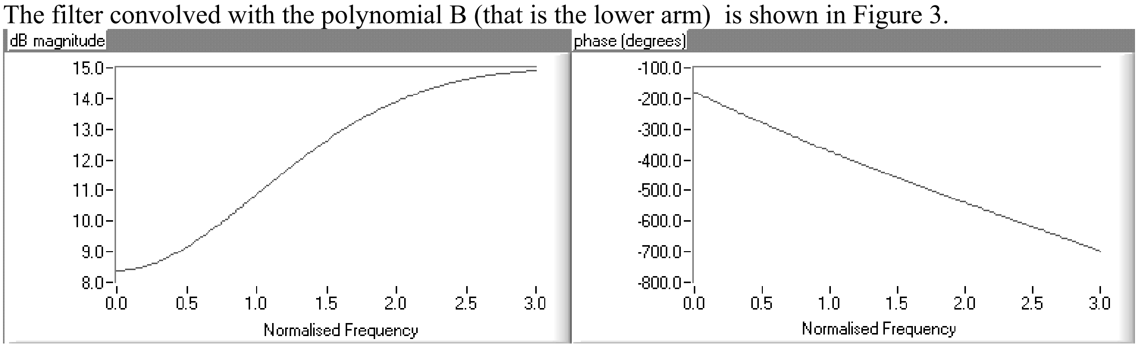

Figure 4. Frequency response of convolution $B_1 B_2 * W$

which can be seen to closely match the upper arm of the beamformer (Figure 2). In fact the error between the two graphs is shown in Figure 5.

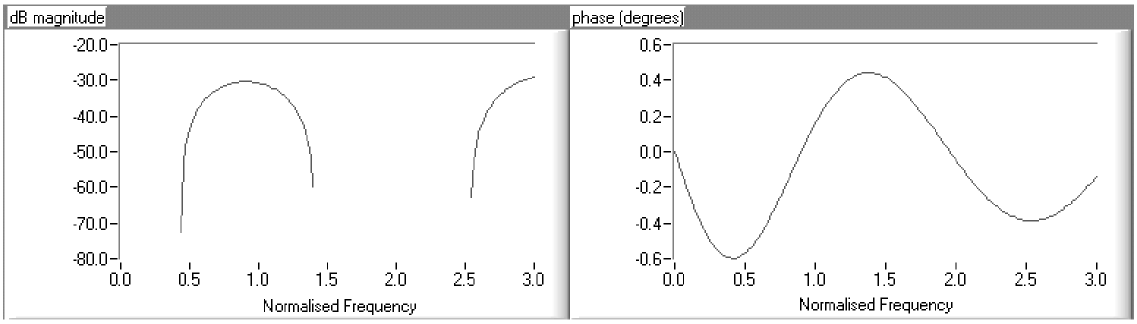

Figure 5. Frequency response of error.

The amplitude error is down -30dB whilst the phase error is less than 0.6 degrees. Finally, the effect of increasing the added delay ℓ from 2 to 4 is seen in Figure 6.

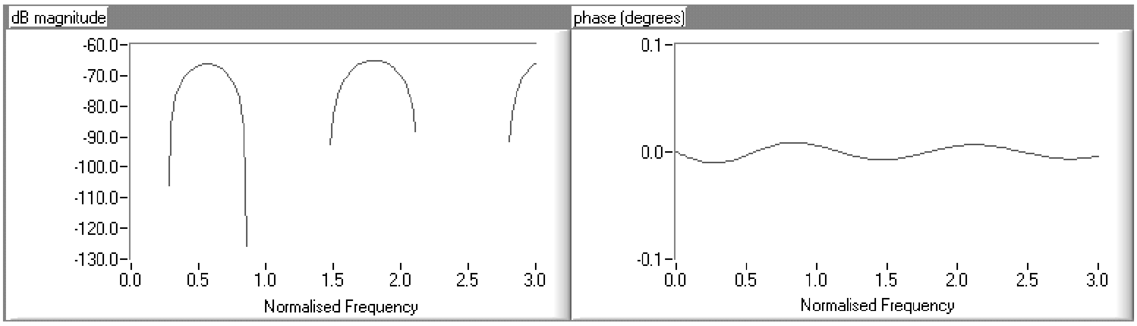

Figure 6. Frequency response of error for $\ell = 4$.

The effect of doubling the delay is to reduce the phase error to less than 0.05 degrees and the magnitude of the error reduces by 35dB to -65dB.

4 Conclusions

The Griffiths-Jim two-input beamformer has been analysed in the frequency domain using a Wiener polynomial based formulation. The particular problem of non-minimum phase acoustic impulse responses has been investigated. The LMS algorithm gives a converged solution which is identical to these results and the frequency-domain characteristics help to aid the understanding of the subject. In future it will be possible to apply the same theory to the multivariable case and with modifications the H_∞ beamformer problem.

Appendix 1

Non-minimum phase beamformer

This appendix will derive the beamformer given in (17) using the method of 'completing the square' [10] Starting from (16)

$$E[e_k^2] = \frac{1}{2\pi j}\oint_{|z|=1} \{\frac{1}{4}[WB_1\tilde{B}_2 - z^{-\ell}C_1C_2\frac{B_2^*}{\tilde{B}_2^*}][WB_1\tilde{B}_2 - z^{-\ell}C_1C_2\frac{B_2^*}{\tilde{B}_2^*}]^*\sigma_\eta^2 + \Phi_{ss}\}\frac{dz}{z} \qquad \text{(A1)}$$

write this in the form

$$E[e_k^2] = \frac{1}{2\pi j} \oint_{|z|=1} \{\frac{1}{4}[A^+ + A^-][A^+ + A^-]^* \sigma_\eta^2 + \Phi_{ss}\} \frac{dz}{z} \tag{A2}$$

where

$A^+ = WB_1\tilde{B}_2 - \{z^{-\ell} C_1 C_2 \frac{B_2^*}{\tilde{B}_2^*}\}_+$ will be in the form $v_0 + v_1 z^{-1} + v_2 z^{-2} + v_3 z^{-3} + \ldots$ (A3)

and

$A^- = -\{z^{-\ell} C_1 C_2 \frac{B_2^*}{\tilde{B}_2^*}\}_-$ will be in the form $u_1 z + u_2 z^2 + u_3 z^3 + \ldots$ (A4)

Clearly $A^+ A^{-*}$ and $A^- A^{+*}$ will have no terms in z^0 and by Cauchys residue theorem their integral must be zero leaving

$$E[e_k^2] = \frac{1}{2\pi j} \oint_{|z|=1} \frac{1}{4}[A^+ A^{+*} + A^- A^{-*}]\sigma_\eta^2 + \Phi_{ss} \frac{dz}{z} \tag{A5}$$

Since only A^+ includes the causal beamformer filter W then (A5) must be minimized when $A^+ = 0$ giving

$$A^+ = WB_1\tilde{B}_2 - \{z^{-\ell} C_1 C_2 \frac{B_2^*}{\tilde{B}_2^*}\}_+ = 0 \tag{A6}$$

or

$$W = \{z^{-\ell} C_1 C_2 \frac{B_2^*}{\tilde{B}_2^*}\}_+ \frac{1}{B_1 \tilde{B}_2} \tag{A7}$$

with the minimal mean-squared error being

$$\sigma_e^2 = \frac{1}{2\pi j} \oint_{|z|=1} \frac{1}{4}[A^- A^{-*}]\sigma_\eta^2 + \Phi_{ss} \frac{dz}{z} \tag{A8}$$

Appendix 2

Polynomial simplification:

Appendix 1 gives the Wiener solution to the optimal non-minimum phase beamformer problem. This appendix simplifies the computation of the transfer function solution (A7).

Consider the term within the {.} brackets in (A7). Taking $\frac{B_2^*}{\tilde{B}_2^*}$ alone, this can be expanded as a convergent power series in positive powers of z.

$$\frac{B_2^*}{\tilde{B}_2^*} = d_0 + d_1 z + d_2 z^2 + \ldots \tag{A9}$$

When combined with the pure delay $z^{-\ell}$ this can be written as

$$\frac{B_2^*}{\tilde{B}_2^*} z^{-\ell} = d_0 z^{-\ell} + d_1 z^{-\ell+1} + d_2 z^{-\ell+2} + \ldots + d_\ell + d_{\ell+1} z + d_{\ell+1} z^2 + \ldots \tag{A10}$$

$$= \tilde{D}_\ell + E^* \tag{A11}$$

where

$$\tilde{D}_\ell = z^{-\ell} D_\ell^* = d_0 z^{-\ell} + d_1 z^{-\ell+1} + d_2 z^{-\ell+2} + ... + d_\ell \quad \text{(A12)}$$

and is E^* an infinite series

$$E^* = d_{\ell+1} z + d_{\ell+1} z^2 + ... \quad \text{(A13)}$$

re-indexing we can write (A13) as

$$E^* = e_0 + e_1 z + e_2 z^2 + ... \quad \text{(A14)}$$

where $e_0 = 0,\ e_i = d_{\ell+i}, i = 1,2...$
Using (A11) in (A7) results in

$$\{z^{-\ell} C_1 C_2 \frac{B_2^*}{\tilde{B}_2^*}\}_+ = C_1 C_2 \tilde{D}_\ell + \{E^* C_1 C_2\}_+ \quad \text{(A15)}$$

The term $\{E^* C_1 C_2\}_+$ has an infinite series E^* which is an uncausal sequence and a finite polynomial $C = C_1 C_2 = c_0 + c_1 z^{-1} + c_2 z^{-2} + ... + c_{nc} z^{-nc}$ which is causal (though possibly still with some non-minimum phase zeros). Hence we have the Laurent series

$$\mathrm{E}^* \mathrm{C}_1 \mathrm{C}_2 = \sum_{k=-\infty}^{\infty} f_k z^{-k} \quad \text{(A16)}$$

Which can be truncated to some suitable length r thus,

$$\mathrm{E}^* \mathrm{C}_1 \mathrm{C}_2 = \sum_{k=-r}^{r} f_k z^{-k} \quad \text{(A17)}$$

where the coefficients f_k may be found for positive values of k as [15]

$$f_k = \sum_{i=0}^{ne} e_i c_{k+i}, \quad k = 0,1,2,..r \quad \text{(A18)}$$

and for negative values

$$f_k = \sum_{j=0}^{ne} e_{-k+j} c_j, \quad k = -1,-2,..-r \quad \text{(A19)}$$

(A18) and (A19) have been convolved with an upper limit $ne >> nc$ which is the truncated limit of the polynomial E^*.
Giving two polynomials.

$$\mathrm{E}^* \mathrm{C}_1 \mathrm{C}_2 = \mathrm{F} + \mathrm{G}^* \quad \text{(A20)}$$

A causal polynomial F

$$\mathrm{F} = f_0 + f_1 z^{-1} + f_2 z^{-2} + ... + f_r z^{-r} \quad \text{(A21)}$$

and an uncausal polynomial G^*

$$\mathrm{G}^* = g_1 z + g_2 z^2 + ... + g_r z^r \quad \text{(A22)}$$

where $g_i = f_{-i}, i = 1,2...$
The polynomial form for the optimal beamformer then follows from (A7)

$$W = \frac{C_1 C_2 \tilde{D}_\ell + F}{B_1 \tilde{B}_2} \quad \text{(A23)}$$

The minimum mean squared error is found from (A8)

$$\sigma_e^2 = \frac{1}{2\pi j} \oint_{|z|=1} \frac{1}{4}[GG^*]\sigma_\eta^2 + \Phi_{ss} \frac{dz}{z} \quad \text{(A24)}$$

Indicating an increase in mean-squared error over the minimum phase case (9).
Sub-optimal beamformer:
An approximate sub-optimal filter can be found by assuming that the higher order terms in the power series expansion $\frac{B_2^*}{\tilde{B}_2^*}$ will be negligible making the coefficients of E^* negligible and hence via (A20) both F and E^* can be neglected. This will only be the case if the delay ℓ is chosen to be large enough resulting in a sub-optimal filter $W^'$ where

$$W^' \approx \frac{C_1 C_2 \tilde{D}_\ell}{B_1 \tilde{B}_2} \quad \text{(A25)}$$

5 References

[1] B.Widrow & S.D Stearns,. 1985 *Adaptive Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ

[2] J.L.Flanagan,J.D.Johnston,R.,Zahn, and G.W.Elko, Computer-steered microphone arrays for sound transduction in large rooms, *J.Acoust.Soc.Am*,78,No 5,pp 1508-1518,1985

[3]L.J.Griffiths, and C.W.Jim, An alternative approach to linearly constrained adaptive beamforming,*IEEE Trans AP*,AP-30,No 1,pp 27-34,1982.

[4]D.R.Campbell and P.W.Shields,, Speech enhancement using sub-band adaptive Griffiths-Jim signal processing, *Speech communication*, 39, pp 97-110, 2003

[5] I.A.McCowan,D.C.Moore and S.Sridharan,. 2002 Near-field adaptive beamformer for robust speech recognition, *Digital Signal Processing* 12,1,pp87-106

[6]J.Vanden Berghe and J.Wouters, An adaptive noise canceller for hearing aids using two nearby microphones,*J Acoust.Soc.Am*,103,No 6,pp 3621-3626,June 1998

[7] T.Nishiura,R.Gruhn, and S.Nakamura, Automatic steering of microphone array and video camera toward multi-lingual tele-conference through speech-to-speech translation,*Proc IEEE Inter Conf Multimedia and Expo,ICME-2001*, pp 569-573,2001

[8]S.Haykin,*Adaptive signal processing,*Prentice Hall,1985

[9]H.Agaiby, &T.J. Moir,. 1997 Knowing the wheat from the weeds in noisy speech, *Proc. 5th European Conf. Speech Comm. & Technology*, Rhodes, Greece, Sept 22-25.,1979
[10]J.F.Barrett and T.J.Moir, A unified approach to multivariable discrete-time filtering based on the Wiener theory,*Kybernetica*,23,No 3,pp 177-196,1987

[11]S.T.Neely and J.B.Allen,Invertibility of room impulse reponse,*J.Acoust.Soc.Amer*,66,No 1 pp 165-169,1979

[12]M.Miyoshi, and Y.Kaneda, Inverse filtering of room acoustics,*IEEE Trans. ASSP*,36,No 2,pp 145-152,1988

[13]M.J.Grimble, A control weighted minimum-variance controller for non-minimum phase systems,*Int.J.Cont*,33, No 4, pp 751-762,1981

[14]A.Ahlen and M.Sternad, Optimal deconvolution based on polynomial methods,*IEEE Trans.ASSP*,37,No 2,pp 217-226,1989

[15] T.J.Moir, and J.F.Barrett, A Kepstrum approach to Filtering, Smoothing, and Prediction with Application to Speech Enhancement,*Proc Royal Soc.London,A*, vol. 459, no. 2040, pp 2957-2976,Dec 2003.