BEAMFORMERS AND SPATIAL TRANSFORMS

David R. Bergman

1Exact Solution Scientific Consulting LLC
3159 Schrader Road, 07801, Dover, USA

ABSTRACT

In this paper, the mathematical equivalence between beamformer algorithms and a
spatial transform of the field data collected by the array elements is explored. A proof is
presented which illustrates the equivalence. When applied to the classic delay and sum
beamformer, or a pure phase in the frequency domain, the result relates this beamformer
to the two-dimensional Fourier transform of the array data. This places the connection
between beamformer and imaging on firm theoretical ground. The equivalence also
suggests algorithmic enhancements that can help in producing better images, and produce
optimized approaches to signal processing. These enhancements are dependent on the
additional preprocessing of raw data that is required to capture pulsed data, and isolate
specific frequency bins. Specific cases applied to simulated data along with a cost
analysis for estimating the processing load are presented.

1 INTRODUCTION

The connection between a far field, or DAS, beamformer and a discrete Fourier transform
(DFT) has been noted and used in sonar signal processing algorithms. The realization is fairly
straight forward, having been used by Weber and Heisler to develop high performance
algorithms used in towed linear arrays of equally spaced elements [1]. This presentation is
unique in that it presents a derivation that takes a slightly different approach than identifying
DAS with DFT. The connection presented here does not assume a specific form of the
weights, uses a continuum approach, and in theory applies to a broader class of beamformer
algorithms. For the purpose of physical motivation, beamforming and signal processing
involved in locating multiple sources and targets in the far field of the array is considered.
The array may be active of passive, this has no bearing on the main result but may introduce
additional signal processing steps to single out a specific frequency of interest in a pulse echo.
The beamformer algorithm is expressed in the frequency domain and narrow band signals are
assumed as a starting point. The array is assumed to be a planar array of identical, ideal,
sensors. The beamformer weights are functions of a set of scan-grid parameters, see Fig. 1. The following variables are defined:

Indices \( n \) or \( m = 1, \ldots , N \) are used to label the sensors in the array.

Indices \((i, j)\) or \((i', j')\) are used to label coordinates in the array face.

- \( \tilde{r}' = (x', y') \) 2-D Cartesian coordinates in the plane of the array face.
- \( \tilde{r}_k = (x_k, y_k, z_k) \) Scan-grid coordinates location of the \( k \)-th grid point.
- \( \tilde{r}_{\perp k} = (x_k, y_k) \)
- \( r_k = |\tilde{r}_k| \) Distance to the \( k \)-th grid point.
- \( (\theta_x, \theta_y, r) \) Angular coordinates for a shell of radius \( r \) centered at the array.
- \( w_m(\tilde{r}) \) or \( w(\tilde{r}, \tilde{r}') \) Beamformer weights.
- \( u_m(t) \) Time series data in sensor \( m \).
- \( u(\tilde{r}', t) \) Time series data in sensor \( m \), expressed in array face coordinates.
- \( f_c \) Carrier frequency.
- \( k_c \) Wave number associated with the carrier frequency.

\image{2}{beamforming_conference.png}
2 THEORY

2.1 The generic beamformer

The starting point is the expression for a standard beamformer, Eq. (1). A set of sampled array data \{u_k(t)\} is projected onto a set of weights and the result summed. Weights are chosen to represent a source at bearing \{\theta_x, \theta_y\} [2]. The squared magnitude to the summation, often referred to as the beam power, can be thought of as representing the likelihood that the data measured by the array is due to a source at \{\theta_x, \theta_y\} when properly normalized. The data may already be passed through a filter to single out a specific frequency bin and one typically takes the time average over several samples over a time interval T.

\[
b(\vec{r}_k) = \sum_{n=1}^{N} \sum_{m=1}^{N} w_m^*(\vec{r}_k) \left\{ \frac{1}{T} \int u_m(t) u_n^*(t) dt \right\} w_n(\vec{r}_k) = \frac{1}{T} \int dt \left| \sum_{n=1}^{N} w_m^*(\vec{r}_k) u_m(t) \right|^2
\]  

(1)

In Eq. (1) m, n indices go from 1 to N, the number of array elements. This index scheme does not allude to any geometric placement of the array elements. The discrete sums inside the integral of the second equal sign may be defined such that the summation is over discrete coordinates in the face of the array.

\[
\sum_{n=1}^{N} w_m^*(\vec{r}_k) u_m(t) = \sum_{(i,j)} w_w^*(i,j) u(i,j)(t) = \sum_{(i,j)} w^*(x',y',\vec{r}_k) u(x',y',t)
\]  

(2)

The limits of the sum are suppressed in Eq. (2), their meaning is defined in Fig. 2.

A few points regarding the beamformer expression(s) in Eq. (1) are mentioned. First, the quantity in brackets in the first equality is the cross spectral density matrix, CSDM or CSM [2,3].

\[
A_{mn} = \frac{1}{T} \int u_m(t) u_n^*(t) dt
\]  

(3)

Thus, another expression for the beamformer, Eq. (1), is \langle w|Aw \rangle = w_m^* A_{mn} w_n, summation implied. An interpretation of this is the expectation value of the CSDM operator in the basis |w\rangle. The second point is that the beamformer weights are normalized such that \langle w|w \rangle = w_m^* w_n = 1. In the case of the delay and sum beamformer, DAS, this achieved by dividing the weights by \sqrt{N}. For the following this factor is carried explicitly, defining \langle W \rangle = |W|/\sqrt{N}. When |W\rangle is a pure phase, this is sufficient to normalize the weights. In more general cases where windowing functions are applied additional normalization factors may be required. These are assumed to be included in |W\rangle. The uniformly spaced sample points, i.e. sensor locations, are separated by \Delta_x, \Delta_y, which are assumed equal, \Delta_x = \Delta_y = \Delta.

Taking the limit as \Delta \to 0 while N \to \infty, keeping \sqrt{N}\Delta = L constant, the double sum goes to an integral.
\[
\sum_{(i,j)} w(x_i', y_j', \tilde{r}_k) u(x_i', y_j', t) \to \frac{1}{A} \int dx' dy' W^*(\tilde{r}', \tilde{r}_k) u(\tilde{r}', t)
\] 

(4)

In Eq. (4) the area of the array aperture is introduced, \( A = (N \Delta)^2 \). The integral in Eq. (4) is a generalized transform written in compact form in Eq. (5).

\[
\psi(u)(\tilde{r}_k, t) = \int dx' dy' W^*(\tilde{r}', \tilde{r}_k) u(\tilde{r}', t)
\] 

(5)

The transformation in Eq. (5) takes \( u \) from array face variables \((x_i', y_j')\) into scan-grid variables \((\theta_x, \theta_y)\), for now \( t \) goes along for the ride. Equation (1) may now be expressed in terms of this transform.

\[
b(\tilde{r}_k) = \frac{1}{T} \int dt |\psi(u)(\tilde{r}_k, t)|^2
\] 

(6)

The beam former algorithm may now be described by the following steps.

1. The data across the array face, \( \tilde{r}' \), is transformed into a new function in \( \tilde{r}_k \) for each time, \( t \), \( \psi(u)(\tilde{r}, t) \).
2. The squared magnitude of each of these transforms is constructed, $b(\vec{r}_k, t) = \psi^*(u)\psi(u)$.

3. The quantity in step 2 is integrated over the block of time $T$, forming a time averaged transform, $b(\vec{r}_k) = T^{-1} \int b(\vec{r}_k, t) dt$.

From these steps it is clear that the output of the beamformer is equivalent to a time average of the square amplitude of the spatial transform from array face coordinates to scan-grid coordinates.

It should be mentioned that the integral in Eq. (5) has finite limits, e.g. $[-L/2, L/2]$. In passing to the continuum limit the data sample can be considered as a function defined on a finite region of $R^2$, and the domain of the weights taken to be all of $R^2$.

### 2.2 Far-field Green’s function and Fourier transform

To make contact with the DAS beamformer one can simply choose the weights to be pure phase. The same result can be derived using the Green’s function as the weight and taking the far field limit. This approach demonstrates a direct connection between the choice of weights and the propagator from source to array sensor. The standard beam former uses the propagator or Green’s function for a point source located at grid point $\vec{r}_k$ for the weights.

$$w(\vec{r}', \vec{r}) = A_0 \frac{\exp ik_c |\vec{r} - \vec{r}'|}{|\vec{r} - \vec{r}'|}$$ \hspace{1cm} (7)

For brevity the amplitude and phase functions are defined, $A(\vec{r}', \vec{r}) \equiv A_0 / |\vec{r} - \vec{r}'|$ and $f(\vec{r}', \vec{r}) \equiv |\vec{r} - \vec{r}'|$. For this choice Eq. (5) becomes,

$$\psi(u)(\vec{r}, t) \sim \int dx' dy' u(\vec{r}', t) A(\vec{r}', \vec{r}) e^{-ik_c f(\vec{r}', \vec{r})}$$ \hspace{1cm} (8)

Taking the far-field limit, the phase and amplitude functions are approximated, $f(\vec{r}', \vec{r}) \approx r_k \{1 - r_k^{-2} \vec{r}' \cdot r_{\perp k}\}$, and $(\vec{r}', \vec{r}) \approx A_0 / r_k$. The transform may be written as follows.

$$\psi(u) \sim A_0 \frac{e^{ik_c r_k}}{r_k} \int dx' dy' u(\vec{r}', t) e^{ik_c \vec{r}' \cdot \vec{r}_{\perp k}/r_k}$$ \hspace{1cm} (9)

Equation (9) is recognized as the Fraunhofer diffraction integral (or its inverse depending on convention) with $u(\vec{r}', t)$ as a source term. A pure phase weight in the frequency domain equivalent to the delay and sum beamformer, hence the transform may be identified with the 2-dim Fourier transform of the data at instant $t$.

This section closes with some comments on the DAS beamformer in the continuum limit. The previous section closed with the statement that the weights could be extended to all of $R^2$ and the data defined on a finite patch. In the case of the pure phase weights this allows one to identify the weights with a complete orthonormal set of functions and make use of the completeness relation.
3 PRACTICAL CONSIDERATIONS

3.1 FT versus FFT

Recognizing the pure phase beamformer as a spatial Fourier transform (FT) motivates one to consider using the FFT to produce beamformer results for each instant of collected data, or each time for which filtered data is procured. While the FFT is fast, there are limitations imposed by the structure of this algorithm that may compete with other interests. This needs to be addressed on a case by case basis.

3.2 Resolution

Identifying Eq. (9) with a 2dim FFT the transform variables are as follows.

\[(x', y') \rightarrow \lambda^{-1}(\sin \theta_x, \sin \theta_y) \quad (10)\]

Limits and sampling resolution of \(\sin \theta_x, \sin \theta_y\) are determined by the sampling resolution and limits of \((x', y')\). The beamformer can be evaluated for any choice of \(\sin \theta_x, \sin \theta_y\) and is not necessarily limited by the sampling of the field by the array face. However, the use of the FFT algorithm to evaluate \(\psi(u)\) will only give \(N\) transformed quantities for \(N\) data points. This does not mean that the output is restricted to only \(N\) values. For a discrete set of input data, the FT still provides a continuum of values. Given the discrete FT of a data sample \(\{u_k\}, k = 1, \ldots, N\), with sampling period \(\Delta\) the FT of \(\{u_k\}\) is,

\[U(f_n) = \Delta \sum_{k=0}^{N-1} u_k e^{i2\pi kn/N} \quad (11)\]

Now consider \(U(f_n + \delta)\).

\[U(f_n + \delta) = \Delta \sum_{k=0}^{N-1} [u_k e^{i2\pi k\Delta \delta}] e^{i2\pi kn/N} \quad (12)\]

The final equality is recognized as the FT of a phase modulated data set \(\{u_k e^{i2\pi k\Delta \delta}\}\). It should be noted that this procedure will generate any frequency bin of the discretely sampled data, which may be thought of as the convolution of the continuous field with a comb function. This will not generate the complete spectra of the original field. Once this is identified as an FT of a new data set the FFT can be applied to the phase modulated data to generate as much of the spectrum of the discrete function \(\{u_k\}\) as desired. This also holds for the spatial FFT applied to array face data. Since the beamformer does not presume a scan-grid resolution one is free to make the grid as dense or sparse as desired. The take away from this is that by application of phase shifts to data one can increase the scan grid sampling by use of the FFT approach on multiple copies of the array face data. This will not improve target resolution. Equations 11 and 12 illustrate interpolation in the time – frequency variable pair. To apply this to the array data requires a spatial phase shift applied to the data.
\[ U(\tilde{x} + \delta) = \text{FFT}(e^{ik\cdot\delta}u(\tilde{x})) \] (13)

One does not need to perform this transform once for each desired point. For example, to increase the scan grid resolution in each direction by a factor of 2 the shift only needs to be performed 3 times, once in each direction and once along a 45 degree diagonal. Figure 3 illustrates this. The large black crosses mark the location of the grid points without shift. The light blue crosses mark grid points obtained from the output for a shift in the \(x\) and \(y\) directions by \(1/2\) the original grid spacing, the diagonal shift is omitted to keep the figure legible.

![Figure 3 - Increasing grid density by phased FFT](image)

The desire to increase the grid density by phase shifted data is motivated by the desire to make use of the FFT in developing the instantaneous images. This is a matter of choice. Zero padding is another approach to effectively increasing the scan grid resolution. In this paradigm, apply spatial windowing functions and zero padding can be applied to the array data to reduce side lobe levels and increase resolution of the scan grid. The effect of these signal processing considerations is illustrated in Figs. 4 and 5 below. Figure 4 shows the output of the standard DAS beam former for an array designed for a maximum frequency of 2200Hz acoustic signal in air \((c = 343\text{m/s})\). There are two targets (or sources) at \(\theta = \pm 15^0\). The array is 3m-by-3m with 40 elements along each direction, a total of 1600 elements with spacing of \(-7.7\text{cm}\).
The scan grid is 200-by-200 sampled from $\sin \theta \in [-0.5, 0.5]$ in each direction. In Fig. 5 two examples of the FFT method are illustrated. The first, on the left, is an FFT of the array data with no additional signal processing methods applied. One can see that the two targets are clearly defined at the correct bearing angles but the images is coarse. The right panel of Fig. 5 is the result of the FFT applied to the array data zero padded out to 256-by-256 data points and with a smoothing window applied, $\cos(\pi x / L) \cos(\pi y / L)$. Since the array is sampled to Nyquist for the frequency considered the FFT provides data out to $\sin \theta = \pm 1$. Sidelobe suppression relative to the main peak is clear in this example. The effect of smoothing the
data does not increase the target resolution, which is determined by the Rayleigh criterion, \( \sim \lambda/L \). The signal processing methods merely enhance the output.

### 3.3 Missing elements and irregular arrays

The relationship between the beamformer and spatial transform made use of regularly spaced array elements. In fact, the standard approach to FFT requires evenly spaced sample data. It is often the case that evenly sampled data has missing data or that application of a higher order algorithm produces irregularly spaced effective elements. Independently of this, optimal array design often leads to irregular element spacing or placement along geometries that do not fit into a rectangular grid [2]. In such cases it is still possible to make use of the FFT for imaging targets by use of non-uniform FFT algorithms [4,5].

### 3.4 Complexity

The complexity of the various approaches to evaluating the beamformer are compared. Three approaches are considered in the limit of large number of elements and scan grid points. Overall factors and cost of primitive operations is scaled to 1 in big \( O \) notation. The first implementation is the expectation value of the CSDM for each scan grid point, \( \langle w \mid Aw \rangle \). Building the CSDM, \( A_{nm} \), requires taking \( N^2 \) multiplications taken \( K \) times then averaged over the \( K \) time samples for a total of \( O(N^2K) \) operations. Next is the operation of a matrix times a vector, followed by the inner product of two vectors, performed for each scan grid point. This is total of \( O(N^2M) + O(NM) \) operations. Assuming large values for \( N \) and \( M \), \( \sim O(N^2(M + K)) \). Operations such as conjugation etc. have been ignored. Changing the order of operations consider taking the inner product of the weight and data vectors over the array face, i.e. \( \langle w \mid u \rangle \). This is an \( O(N) \) operation. This is done for each grid point and each time for a total of, \( O(NMK) \). Taking the magnitude of each scalar and summing over \( K \) for each \( M \) is \( \sim O(MK) \). Last, consider the 2dim FFT for evaluating the beamformer. This is an \( O(N \log_2 N) \) operation. Done \( K \) times and summed, \( \sim O(KN \log_2 N) + O(NK) \). For large \( N \) and \( M \), with \( M = N \), and a single time evaluation, the ratio of the three methods is approximately \( N^3 \cdot N^2 \cdot N \log_2 N \). The analysis implies that the time average of \( \langle |w \mid u \rangle \rangle^2 \) is a better option than evaluating \( \langle w \mid Aw \rangle \). This depend on other factors. Methods for reducing noise rely on either eliminating terms in the CSDM or an eigenvalue analysis of the CSDM requiring a fully developed matrix before beamforming. On the other hand, detection of narrow band signals in active sonar systems under relatively noise free environments are ideal candidates for the other two procedures.

### 4 EXTENSIONS

#### 4.1 T-A-P and co-array

Two common extensions of the standard beamforming algorithm are the time-averaged-product array, or T-A-P, and the co-array. The T-A-P makes use of time averaged product of the sensor data to generate array processing output. It has been shown that the effect of such algorithms is equivalent to a sum array for a larger number of elements [6,7]. Recent work in this area suggests applying the standard beamformer to the product array outputs [8]. Figure
4 is an example 5-by-5 array along with the product array or co-array generated by the original 5-by-5 array. For this case both the T-A-P and co-array contain the original array elements.

![Diagram of Original and Co-Array Elements](image)

Figure 6 – Sample physical array and equivalent array generated by T-A-P or co-array process

A co-array is generated by taking all differences in array element location. T-A-P array processing involves taking the time averaged product between pairs of elements and passing these quantities through a summation. Consequently, there are a wide variety of T-A-P processing choices for a given set of inputs. These are divided into classes; the reader being referred to reference [6] for details on class definitions. As illustrated in Fig. 6, the effect of such algorithms is to increase the effective array aperture size. For the case provided here, the larger array contains the original elements as a subset but such algorithms can result in irregular array element spacing. Figures 7 and 8 below show example output from a physical array, of size $L \times L$, and T-A-P processing of a smaller array, $L/2 \times L/2$. There are two identical sources present, each with $f = 2200$ Hz, placed at $\pm 5^\circ$. The element spacing is chosen to produce a critical angle of $\theta_c \pm 30^\circ$, or $\sin \theta_c = \pm 0.5$. In both cases the sampled data was zero padded to 128-by-128 and the images generated by 2-dim FFT, rather than direct application of the DAS beamformer.
The two-dimensional surface plot in Fig. 7 is of the amplitude of the image normalized by the largest value, e.g. colour scale maps to $[0, 1]$. Figure 8 is a cross section of the surface plot through the centre line on a log scale. A few features are worth mentioning. First is that the T-A-P processing maxima are at $\pm 5.2^\circ$, while the direct result has maxima at $\pm 4.3^\circ$. The T-A-P processing involves only 25% as much input data as the direct case, at the cost of a more complex processing flow. The T-A-P result contains slightly larger sidelobes, and a more
complex structure. There exists more than one approach to defining the inputs generated by T-A-P processing. In this case all products correlating to a given effective element location are summed. This approach produces a natural windowing function.

4.2 Non-linear beamformers

A wide class of high resolution beamformers, also known as inverse beamformers or nonlinear beamformers, exist that give rise to better resolution of targets for a fixed array geometry as compared to DAS [3, 9,10,11,12,13]. These beamformers involve the use of the inverse of the cross spectral density matrix, rather than the direct CSDM. Such beamformers require the complete CSDM for the array data before the weights are generated or pre-steering applied. Another way to think of these beamformers is that the weights are data dependent. The approach taken in section 2 assumes that the weights are independent of the data, forming a basis for a function space in which the data is projected. Consequently, these types of beamformers do not map to the FFT algorithm as the DAS. Consider the form of the MEM beamformer \[3\], \(\text{MEM}^{-1} = \langle w|A^{-1}|w \rangle\). Using a series expansion of the inverse matrix (provided it exists) allows this to be written,

\[
\sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} (-1)^k \langle w|A^k|w \rangle
\]

(13)

Provided that the expansion can be applied, the individual terms may be expressed in terms of the image projected on the array. This can be illustrated either in the original discrete sum or in the continuum limit. The time averaging in the definition of \(A\) will need to be considered carefully. As an explicit example consider the second order term.

\[
w_n^*A_n^2w_l = w_n^*A_{nm}A_{ml}w_l = \frac{1}{T^2} w_n^* \left\{ \int_0^T u_n(t)u_m^*(t)dt \int_0^T u_m(t')u_l^*(t')dt' \right\} w_l
\]

\[= \frac{1}{T^2} \int_0^T dt dt' \{w_n^*u_n(t)\}{u_m^*(t)u_m(t')}{u_l^*(t')}w_l \]

\[= \frac{1}{T^2} \int_0^T \int_0^T dt dt' T(u(t))\{u_m^*(t)u_m(t')\}T^*(u(t)) \]

Expressed in terms of a discrete time sampling,

\[
w_n^*A_n^2w_l = \frac{1}{T^2} T_i(u)M_{ij}T^*_j(u) \]

(14)

Since \(T\) is used as the sampling interval and the image operator, the argument \(u\) is retained to avoid confusion. The new matrix \(M_{ij}\) is constructed by taking sum over all array elements of every pair of time samples in the window \(T\). Higher order terms are fairly straightforward,
\[ w_n^* A_{nl}^k w_l = \frac{1}{T^k} T_{\alpha_1} (u) M_{\alpha_1 \alpha_2} M_{\alpha_2 \alpha_3} \cdots M_{\alpha_{(k-1)} \alpha_k} T^*_{\alpha_k} (u) \] (15)

The action of the operator, \( M_{ij} \), and comparison to the DAS beamformer, are illustrated in Fig. 9 and Fig. 10.

**Figure 9 – Pictoral representation of the DAS beamformer**

**Figure 10 – Pictorial representation of the quadratic term in the non-linear beamformer expansion**

The DAS beamformer takes the time average of the square image amplitude, Fig. 9. In contrast, each term in the expansion of the inverse beamformer couples the images from different times using a weighted average over powers of the time sequence array data, Fig. 10. This view of the action of the inverse beamformer may offer some theoretical insight but estimating the output of these approaches using a truncated series leads to suboptimal results.
5 SUMMARY

A connection between the classic beamformer method and spatial transform has been demonstrated and discussed. The approach makes use of the continuum representation of an array, later identifying a discretely sampled array with DAS weights with the Fourier Transform. Application of this approach to extensions of the linear beamformer were also considered. For irregular spaced arrays non-uniform FFT techniques offer a possible extension of this technique. Applied to non-linear and inverse beamformers show that these types of algorithms can be interpreted as an infinite series in the images produced at a sequence of times. While this is interesting in theory it is not likely to lead to fast high-resolution algorithms as with the linear beamformer.

REFERENCES