# COMMENTS ON UNIQUE RECONSTRUCTION FROM IRREGULAR SAMPLING

## 1. Introduction

In this memo we address the question of under what conditions a continuous signal be reconstructed from discrete samples, considering cases of non-regular sampling sequences. A similar problem involving an interleaved sampling sequence has been dealt with by Bracewell (1986) and references therein, here we consider the same problem from the Fourier transform domain perspective and calculate the increase in noise expected for linear reconstruction from non-regular sampling.

### 1.1 Nyquist Sampling for Regular Sequences

The Nyquist sampling theory *as taught to undergraduate students of electrical engineering or physics* usually only discusses the case of regular spaced sampling in time of a bandlimited analog signal, with an infinite number of samples. It is easily proved that in this case a perfect reconstruction of the analog signal is possible provided the samples are closer than the Nyquist sampling limit of $1/2B'$ (where $B'$ is the bandwidth defined as the range of positive frequencies over which the FT of the signal is non-zero; if both positive and negative frequencies are considered as defining the bandwidth $B = 2B'$ then the sampling criteria is $1/B$). Reconstruction of the analog signal between two samples is in principle easily achieved by convolving all delta function samples with a sinc-function which has its first nulls at +/-1/B.

### 1.2 Sampling theory for interferometry

In the case of interferometry we often observe sources which are to good approximation space limited, since they lie on a black (actually 3K) background (true of most small extragalactic sources for which Olbers principle applies). In other cases the image can be considered approximately space limited due to multiplication by the primary beam response. We can consider therefore that the uv visibility is to some approximation a bandlimited signal with bandwidth B, where B corresponds to the source size or primary beam size, whichever is smaller. The visibilities correspond to the sampled time signal in the Electrical engineering analogy (see Section 1.1) while the image corresponds to the FT of the visibilities (anlog spectrum) before sampling. For most arrays of interest the uv plane is irregularly sampled. Given these facts the question of unique linear reconstruction of bandlimited signals in the case of irregular sampling is of interest.

Amongst astronomers it is often stated that a space-limited source of maximum angular size size 'B' can be reconstructed uniquely provided each uv cell of size (1/B) is sampled at least once. Of course there are in fact always a limited number of uv samples and a maximum baseline for the array so all that can actually be achieved is complete sampling of the uv cells out to some maximum spatial frequency. This is what is often referred to as 'complete sampling' (a better name would be 'complete sampling out to a maximum spatial frequency' but this name is less catchy). In order to make a good estimate of the true image convolved with some restoring beam i.e. at limited resolution, without re-weighting the data a naturally tapered distribution of sampling density is also required. As in the 1D signal processing case such a tapered distribution also ensures that the residual sidelobes caused by the abrupt edge of the uv coverage are minimised. It is good to minimise these sidelobes as much as possible since they cannot in practice be removed by linear methods since, as is well known, extrapolation of the sampled signal using linear methods is extremely ill-conditioned with respect to noise. The only ways to reduce these 'residual sidelobes' is by data weighting (which loses sensitivity) or the application of non-linear algorithms which incorporate *a priori* knowledge.

### 1.3 UV sampling criteria

An important part of the sampling criteria described in section 1.2 is that each 'cell' of size (1/B) in the uv plane within the outer uv boundary should be sampled at least once. Now intuitively the criteria of sampling every cell of size (1/B) makes sense since the typical size over which the visibility will vary in the uv plane will be (1/B). In practice amongst those working on this problem **there seems to be no disagreement that if such sampling is achieved for every cell than a linear reconstruction of fixed resolution is possible which is unique within the constraints of the 'residual beam' and is well conditioned with respect to noise.**.

Despite the above some disagreement **does** seem to exist as to the exact status and definition of the sampling theory as applied to the case of irregular sampling. Specifically, does the sampling theory rigorously require that literally every cell out to the uv boundary must be sampled at least once to allow unique linear reconstruction? Should we even consider the cells themselves as fundamental or is instead the correct criteria that all holes in the uv coverage must have diameter of less than (1/B)? Alternatively is the sampling condition actually a much looser rule which if true guarantees a well conditioned reconstruction with respect to the noise, but allows linear reconstruction (with increased noise) even if it is not fulfilled? Support for the latter view comes from considering the problem of estimating an image from its visibilities as one of matrix algebra where the unknowns (values of the intensity in each beam area) are linked to the measurements (visibilities) by a matrix. In the absence of noise a unique solution to this problem requires that the matrix has an inverse. The primary criteria for this is that the number of visibilities exceeds the number of beams in the image, given this then for most uv distributions an inverse exists. This criteria implies a *mean* separation between uv samples of (1/B). Of course the conditioning with respect to added noise might be poor if points are not also evenly distributed in the uv plane.

This memo we will consider the above question. Possible candidate criteria for unique analog signal recovery are discussed in Section 2. In Sections 3 and 4 we consider specific cases of a sampled 1D signal in which the gaps between the samples are not constant and ask under which conditions we can reconstruct the continuous signal. In this memo we will consider only the case of an infinite sequence of samples. Although such sampled signals are not directly comparable to the interferometry case with its limited number of uv points, this idealised problem has the advantage that it is analysable using standard Fourier methods and quantitative results are easily derived. The more directly applicable case of a finite number of samples will be dealt with in a subsequent memo (Conway, in prep) using matrix-vector methods.

## 2. Unique Reconstruction Criteria for Irregular Sampling

Let us consider a 1D analog band-limited signal (of bandwidth B) which is sampled in an irregular fashion so that the interval between samples is not constant. We will assume an infinite sequence of samples. For some non-regular sampling distributions; specifically a regular distribution of sample separation $< 1/B$ plus randomly positioned samples (Boone, private communication), it is clear that the analog signal can be reconstructed. The proof depends on that used to prove reconstruction for a purely regular sampling sequence. Let us however consider a more general non-regular distribution. What is the criteria that a unique perfect reconstruction of the analog signal is possible. Possible criteria are given below;

(I) It can be guaranteed that one can reconstruct uniquely without use of a priori knowledge - IF AND ONLY IF - the signal is sampled such that the MAXIMUM gap between any two samples is not larger than (1/B).

(II) It can be guaranteed that one can reconstruct uniquely without use of a priori knowledge - IF AND ONLY IF -the signal is sampled such that the MEAN gap between any two samples is not larger than (1/B). One obvious ambiguity in this case is the question of over which scale the mean sampling separation should be computed. However for most classes of irregular sampling sequence the mean sampling interval will approach a converging limit as the number of samples considered goes to infinity.

(III) It is possible to uniquely reconstruct - IF AND ONLY IF - each individual 'cell' of width (1/B) is sampled. This is the criteria often used in evaluating arrays via computer. It is unlikely that this is a fundamental criteria, since it depends on what is taken to be the zero point of the cells i.e. consider the sequence 0.99, 2.01, 2.99, 4.01, 4.99, 6.01, 6.99.... (where the unit spacing is equal to 1/B). If the cells are defined for 0-1, 1-2, 2 - 3, etc then this sequence is only 50% sampled, but if the cells are taken to run from 0.5-1.5, 1.5-2.5, 2.5-3.5 the sequence is 100% sampled. Since we can always shift the sequence, process it to recover the analog signal and shift back this cannot be a fundamental criteria of whether we can reconstruct the continuous signal

We should bear in mind that the correct criteria applicable to allow reconstruction in the general case might in fact be none of the above, the criteria might be more complex. It is interesting to note that for a regular sampling all of the above definitions (I,II,III) are fulfilled simultaneously if the sampling rate is greater than Nyquist, but there are clear distinctions between them for non-regular sampling, in particular for case (I) even if one gap is greater than Nyquist the signal cannot be reconstructed uniquely i.e. the sampling sequence (in units of 1/B) of 1, 2.001, 3, 4, 5, 6... is non-unique according to (I) but is unique according to (II). Note also if the condition for (I) is true then so are the conditions for (II) and (III), the opposite however does not follow; if the condition for (II) is true this does not imply that (I) is true. Since (I) is the most stringent condition (and assuming the above list contains the true criteria) then we can be assured that if condition (I) is fulfilled that we can uniquely reconstruct. The question remains however of whether there are signals not fulfilling (I) which can also be uniquely reconstructed.

Proving any of the prepositions (I), (II) and (III) to be true in general is likely to be difficult. Not only are the number of possible sampling sequences infinite but we don't know which, if any, of the prepositions is true, and therefore which is one the one we must prove(!). On the other hand proving a given condition to be false is much easier, all we need do is find a sampling sequence for which the preposition predicts the analog signal to be unrecoverable and then prove that it is recoverable - this proves the preposition false. In the next sections we consider families of irregular sequences which prove that (I) cannot be the correct universal criteria for reconstruction, and strongly suggest (II) to be true.

## 3. Reconstruction from an interleaved sequence

Consider a complex valued band-limited signal as a function of u. The Fourier transform of this signal $S(x)$ is non-zero only between $x = -B/2$ and $x = B/2$, where B is defined as the full bandwidth of the signal. Since the signal is complex the spectrum of the signal $S(x)$ need not be Hermitian conjugate about x=0.

We wish to sample the continuous signal at discrete values of u. From what sequences of samples can we recover the original analog signal exactly? Possible criteria are suggested in Section 2, namely (I),(II),(III). Here we consider a particular sequence that proves (I) to be false, and we generalise this result in Section 4.

Consider a sampling sequence generated from a regular infinite sequence of samples made just at the Nyquist sampling interval $1/B$ by displacing every second sample by an amount $\Delta u$ (see Fig 1a). The resulting sampling sequence can be considered as a sequence of two samples of separation $(1/B - \Delta u)$ repeated infinitely at an interval $2/B$ or alternatively as two infinite interleaved sets of samples each having interval $2/B$. The problem of reconstruction from such an interleaved sequence has been dealt with by Bracewell (1986), considering the form of the interpolation functions needed, here we condier the same problem from the Fourier transform point of view.

The Fourier transform of this irregular sampling sequence (see Fig 1b) is an infinite sequence of delta functions having interval $B/2$ with the nth delta function measured from x=0 being multiplied by a complex number $M_n$

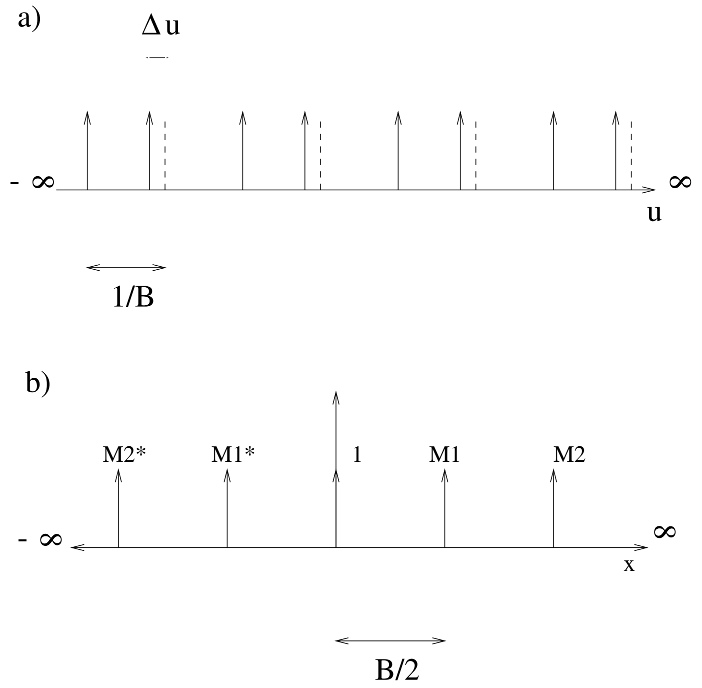


Figure 1: a) A repeating uneven sequence of samples, consisting of delta functions. This sequence has been generated from an infinite sequence of evenly spaced samples which is just Nyquist sampled, i.e. with interval $\Delta u = 1/B$ where $B$ is the bandwidth of the signal, by displacing each second sample by $\Delta u$. This infinite sampling sequence can also be considered as two regular interleaved sequences each having separation between samples of $2/B$ or alternatively as a sequence of two delta functions (with separation $(1/B - \Delta u)$) repeated at intervals of $2/B$. (b) The Fourier transform of the sampling function. This consists of an infinite set of delta functions each separated by (B/2) but multiplied by complex factors $M_1$,$M_2$ etc.

These complex multiplying factors are related to $\Delta u$ via

$$M_n = (1 + exp(-2\pi i(1/B - \Delta u)nB/2))/2$$

or

$$M_n = (1 + exp(-\pi in(1 - \Delta uB)))/2$$

This means that if $\Delta u = 0$, every $M_n$ where n is odd equals zero, and where n is even equals 1. The resulting sequence is, as expected, the FT of a regular set of samples of interval $1/B$. Likewise when $\Delta u = 1/B$ all $M_n = 1$ and we recover the Fourier transform of an infinite sequence sampled at interval $2/B$,

If a complex bandlimited signal is sampled by the sequence given in Figure 1 (with in general some non-zero $\Delta u$) then from the convolution theorem the Fourier transform of the sampled sequence is given by the FT of the sampling sequence convolved with the FT of the unsampled signal, S(x). This is shown pictorially in Figure 2.

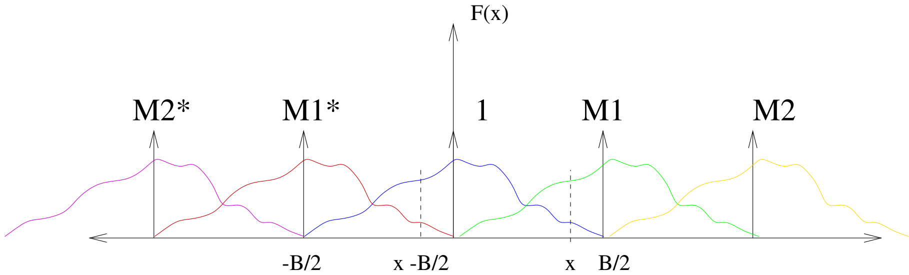


Figure 2: Fourier transform of the sampled signal. This equals the convolution of the Fourier transform of the unsampled signal S(x) with the Fourier transform of the sampling function (Fig 1b). The different colours indicate that each shifted S(x) is multiplied by a different complex number. The total Fourier transform F(x) is the sum of all the coloured functions. Indicated by the dashed vertical lines are the values of F(x) at x and x-B/2 which are related.

If $\Delta u = 0$ then all the odd n, $M_n$ are zero, hence there is no aliasing and the FT of the sampled signal in the range of x from -B/2 to B/2 equals the FT of the unsampled signal. It is therefore simple to recover the pre-sampled signal by multiplying by a top hat function which is non-zero only between -B/2 to B/2, and then doing the inverse Fourier transform.

At the other extreme is the case that $\Delta u = 1/B$, then all the factors $M_n = 1$ and there is aliasing from which we can never recover. Let us consider however the intermediate case $0 < \Delta u < 1/B$ what happens then? Consider the Fourier transform of the sampled sequence at x, $F(x)$ and at $x - B/2$, $F(x - B/2)$. Since the bandwidth of the unsampled signal S(x) extends only from $-B/2$ to $B/2$, the values of $F(x)$ and $F(x - B/2)$ for x in the range 0 to B/2 depend only on the FT of the unsampled signal convolved with the delta functions at n=0 and 1. Likewise for x in the range -B/2 to 0 the FT of the sampled signal only depends on the S(x) convolved with the delta functions at n=-1,0. We have then from Fig 2 that

$$F(x) = S(x) + M_1 S(x - B/2)$$

and

$$F(x - B/2) = M_1^* S(x) + S(x - B/2)$$

which in matrix notation

$$\begin{bmatrix} F(x) \\ F(x - B/2) \end{bmatrix} = \begin{bmatrix} 1 & M_1 \\ M_1^* & 1 \end{bmatrix} \begin{bmatrix} S(x) \\ S(x - B/2) \end{bmatrix}$$

or

$$\mathbf{F} = \mathbf{T}\mathbf{S}$$

Where **F** and **S** are vectors of length 2 and **T** is the 2x2 transformation matrix. Provided the 2x2 matrix $B$ has an inverse then we can therefore recover $S(x)$ and $S(x - B/2)$ from the FT of the sampled data. The condition for the existence of an inverse is that $det(T) = (1 - M_1 M_1^*) \neq 0$. The amplitude of the complex number $M_1$ is always equal or less than one, and it becomes exactly equal to 1 only in the case that $\Delta u = 1/B$, therefore unless this is the case an inverse exists. and hence;

$$\begin{bmatrix} S(x) \\ S(x-B/2) \end{bmatrix} = \frac{1}{detT} \begin{bmatrix} 1 & -M_1 \\ -M_1^* & 1 \end{bmatrix} \begin{bmatrix} F(x) \\ F(x-B/2) \end{bmatrix}$$

where $det(T) = (1 - M_1 M_1^*)$.

Graphically this means that we can recover S(x) for x between 0 and B/2 via the following steps. (1) Take the section of the FT of the sampled signal F(x) between x=0 and x=B/2 and multiply by a complex number determined by the above matrix (i.e. $1/det(T)$). (2) Take the section of F(x) for x between -B/2 and 0, shift this by +B/2 so it lies in the range 0 to B/2 and multiply by a complex number determined from the above reconstruction matrix (i.e. $-M_1^*/det(T)$). (3) Add the results of the first two steps to give F(x) over the range x =0 to B/2. Using a similar procedure we can recover S(x) between x=-B/2 and 0. Once the spectrum $S(x)$ is fully determined the original complex analog signal can then be exactly recovered. These steps can be shown equivalent to interpolation in the u domain via the interpolation functions calculated by Bracwell (2000).

## 4. Increase in Noise

If we assume that the sampled signal has noise added when we make our measurements then the FT of the sampled signal $F'(x) = F(x) + N(x)$ If we try to reconstruct S(x) by the above procedure then the effect in general will be to amplify the noise. The error in the estimate of $S(x)$, namely $N'(x)$ will obey

$$\begin{bmatrix} N'(x) \\ N'(x-B/2) \end{bmatrix} = \frac{1}{detT} \begin{bmatrix} 1 & -M_1 \\ -M_1^* & 1 \end{bmatrix} \begin{bmatrix} N(x) \\ N(x-B/2) \end{bmatrix}$$

the analysis is complicated by the fact that in general the noise at $N(x-B/2)$ is correlated with $N(x)$. The noise contribution $N(x)$ is created by the FT of the noise added to each of the measurements as a function of $u$ in Fig 1 (the complex number $\sigma_n$ added at each sample point]

$$N(x) = \sum_{k=-\infty,\infty} \sigma_{2k} exp(-2\pi i x(2k/B)) + \sum_{k=-\infty,\infty} \sigma_{2k+1} exp(-2\pi i x[(2k+1)/B) - \Delta u])$$

$$N(x) = N_{odd}(x) + N_{even}(x)$$

it follows that

$$N(x - B/2) = N_{odd}(x) + N_{even} exp(\pi i(1 - B\Delta u))$$

or

$$N(x - B/2) = N_{odd}(x) + KN_{even}$$

Where $K = exp(\pi i(1 - B\Delta u))$, so that $M_1 = (1 + K^*)/2$. From Equation ? it follows that

$$N'(x) = (1/det(T))[N(x) - M_1 N(x - B/2)]$$

$$= (1/det(T))[(1 - M_1)N_{odd} + (1 - KM_1)N_{even}]$$

$$= (1/det(T))[(1 - K)N_{odd} + (1 - K)N_{even}]/2$$

Hence after reconstruction, compared to the case of regular sampling, (which has $\Delta u = 0$ and $K = -1$) noise in the extimate of $S(x)$ is amplified by a factor

$$R = \frac{|1 - K|}{1 - M_1 M_1^*}$$

Which from the definition of $M_1$ and $K$

$$R = \sqrt{\frac{2}{1 - cos(\pi(1 - \Delta u B))}}$$

This function is plotted in Figure 3. This result has been verfified in numerical experiments reconstructing a 100 pixel image from 100 uv samples, with 50 interleaved elements of the sampling function offset to the other 50 by $(1/B - \Delta u)$. Examining Fig 3 we find that as expected for large $\Delta u$ noise becomes greatly amplified, tending to infinity as $\Delta u$ goes to $1/B$. On the other hand $R$ stays relatively small for small displacements from regularity, and only increases to 1.08 for displacements as large as $0.25/B$

## 5. Generalisation to recurring irregular sequences

We can generalise the reconstruction method described above (see Figure 4) to a repeating sequence of $N_{samp}$ samples repeating every $N_{int}/B$ where $N_{samp}$ and $N_{int}$ are integers. For the irregular set of samples the $N_{samp}$ values of u, $u_k$ are distributed over the range $x = 0, N_{int}/B$. In this case there is an $N_{int}$ times $N_{int}$ matrix linking $N_{int}$ values of $S(x)$ with $N_{int}$ values of $F(x)$. For instance for the case where $N_{int} = 4$ we have

$$\begin{bmatrix} F(x - B/N) \\ F(x) \\ F(x + B/N \\ F(x + 2B/N)) \end{bmatrix} = \begin{bmatrix} 1 & M_1 & M_2 & M_3 \\ M_1^* & 1 & M_1 & M_2 \\ M_2^* & M_1^* & 1 & M_1 \\ M_3^* & M_2^* & M_1 & 1 \end{bmatrix} \begin{bmatrix} S(x - B/N) \\ S(x) \\ S(x + B/N \\ S(x + 2B/N)) \end{bmatrix}$$

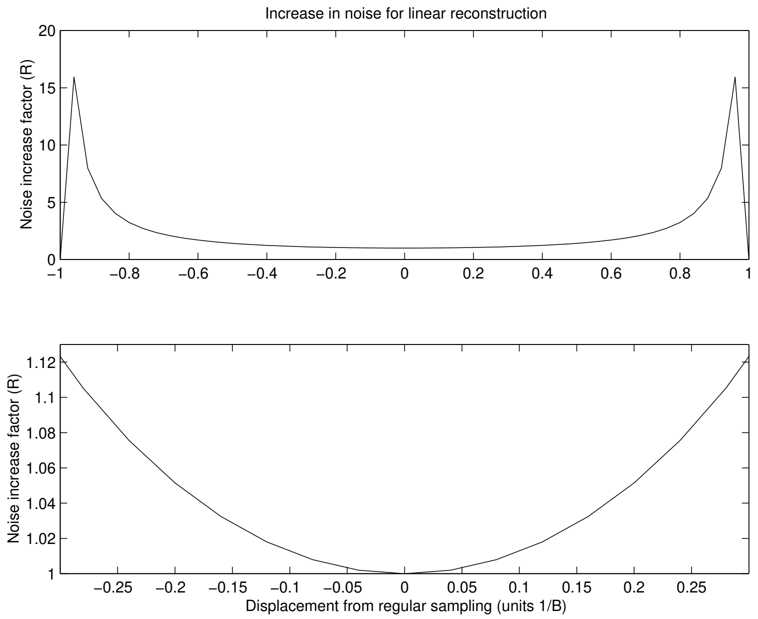


Figure 3: Increase in Noise from linear reconstruction of the signal as a function of the displacement from regularity $\Delta u$ for the sequence shown in Fig 1a

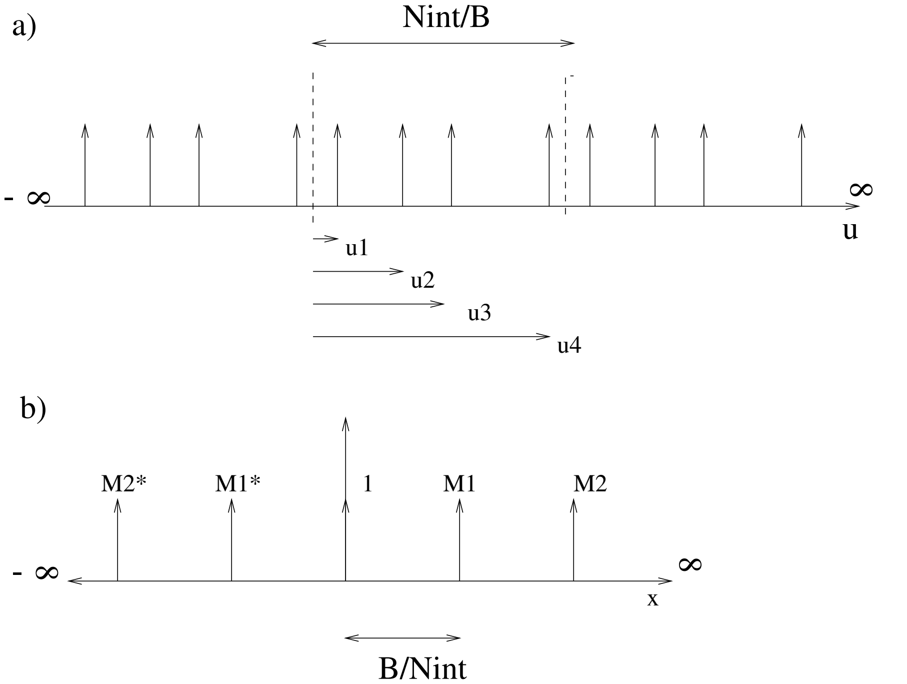


Figure 4: a) A repeating non even spaced sequence of samples, consisting of $N_{samp}$ delta functions, repeated every $N_{int}/B$, in the case shown $N_{samp} = N_{int} = 4$. (b) The Fourier transform of the sampling function. This consists of an infinite set of delta functions each separated by $(B/N_{int})$ but multiplied by complex factors $M_1$,$M_2$ etc.

or

$$\mathbf{F} = \mathbf{T_4 S}$$

where

$$M_n = \sum_{k=1,N_{samp}} exp(-2\pi u_k Bn/N_{int})$$

The values of $M_n$ are functions of the $N_{samp}$ variables ($u_k$), and the transformation matrix depends on these values of $M_n$. The number $N_{samp}$ therefore expresses the number of degrees of freedom in the transformation matrix. Given this when $N_{samp} < N_{int}$, the transformation matrix will not have full rank and therefore is not invertible, hence it is not possible to recover the un-sampled signal from the sampled signal. This condition has been verified by numerical simulation in MATLAB (this remains to be proved mathematically rigorously however). Bracewell (1986) notes that a solution is possible for repeating sequence provided the mean seperation is 1/B. When $N_{samp} = N_{int}$ the matrix is invertible and a unique linear reconstruction is possible. However, as for the case of $N_{samp} = N_{int} = 2$ considered in Section 3 the noise will be increased by a factor which depends on the transformation matrix $\mathbf{T_4}$ and its determinant.

The condition for unique reconstruction of $N_{samp} = N_{int}$ corresponds to having a mean separation between samples which is less than $1/B$. It therefore appears that for repeating irregular sequences it is possible to recover the bandlimited signal even if the largest gaps are larger than $1/B$ provided that the *mean* gap is $1/B$ (or presumably less). Since we have found types of sampling from which we can reconstruct uniquely despite condition (I) not being true (see section 2) this proves that condition (I) cannot be the universal condition for unique reconstruction. In contrast we find that condition (II) is the valid one for deciding if such repeating sequences are reconstructable. Note that for such a repeating sequence there is no ambiguity over the range over which the mean seperation should be calculated, it simply corresponds to the repetition period of the sequence.

## 6. Application to Interferometry

Of course in the interferometry case we do not have repeating sequences of irregular samples, indeed we only have a finite number of samples. Another difference is that we know that the spectrum of the visibility (image) is completely real. A simplified version of this case can be handled by matrix algebra considering the case where there is one unknown in the image for each beam area and each unknown is linked to the visibility measurements by a matrix (e.g Schwartz 1978). Again the condition for unique reconstruction is that the number of visibility measurements (which if we use the condition that the image is real, means including uv measurements and their conjugates as seperate measurements) is greater than the number of unknowns. This condition corresponds to having a mean separation between uv points of less than 1/source size (Conway in prep 1). Despite this however to be well conditioned with respect to noise it is required that the samples be fairly evenly distributed. In the limit of more uv points than image beam areas Scharwz(1978) shows that in the abscence of noise there is a unique solution and provided certain constraints hold CLEAN will find this unique solution. In this limit CLEAN can be considered as algorithm that performs the required interpolation between discrete uv samples to estimate the continuous uv distribution.

A more accurate model of the imaging process, which is presently being investigated, is one where there is more than one unknown per resolution element (i.e. pixel spacing less than the resolution) and where we we wish to make the best estimate not of the original image but of the image convolved with a restoring beam via purely linear (i.e. matrix) methods (Conway in prep 2). In this case to give no increase in noise compared to the dirty map it is expected that the uv point density is required to follow a suitably tapered form which equals the FT of the gaussian restoring beam and in which the largest gap between samples is $< 1/B$ (Boone 2001, in prep). What is interesting from the practical point of view is to determine how the noise in a purely linear reconstriction is increased when this ideal is not quite reached (when say at the uv edge the mean seperation between uv points is 1.5/B, or uv density half of critical). Formally speaking a linear solution is still likely to be possible if the total number of uv points is greater than the number of image beam areas, but noise must increase. This increase in noise at high spatial frequencies will however be limited by the fact that we are attempting to estimate not the true image but the image convolved by the restoring beam. If the increase in noise in the final target image were small (few percent say) then such uv coverages could still be considered as allowing well-condtioned linear reconstructions even if at the edge the mean seperation between uv points were $> 1/B$.

## References

Bracewell, R., 1986, 'The Fourier Transform and it Applications', Chapter 10.
Schwarz U.J. 1978, AA 65, 345.