

Module I, Lecture 03: Introduction to Digital Signal Processing

Fourier transform spectroscopy, particularly NMR and ESR, can draw upon a large family of digital signal processing techniques that can often transform a noisy bump into a collection of beautifully resolved signals or salvage a data set that standard spectral processing software would deem completely corrupted. Before we start with the formal theory of spin systems, it would be wise to get our datasets into the best shape possible.

Sampling rate and digital resolution

Sampling rate is the number of points per unit time used to represent a time domain signal in a discrete form. Modern spectrometers always digitize the time domain signals at the highest rate their hardware is capable of (usually between 1 MHz and 1 GHz). The signal is then *downsampled* to the desired sampling rate by averaging nearby points – meaning that “low” sampling rates yield a better signal-to-noise ratio. The minimum sampling rate is dictated by the highest frequency present in the signal.

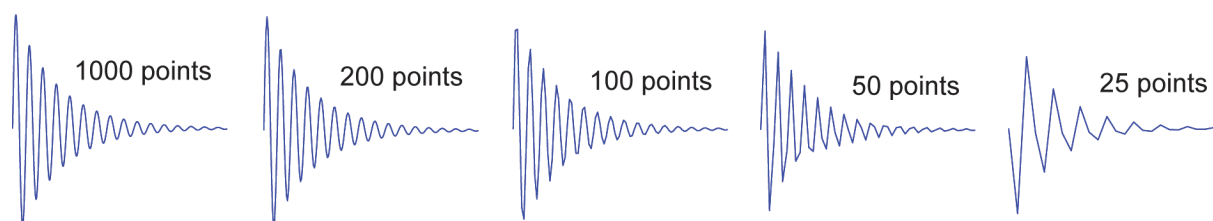


Figure 1. A decaying oscillation digitized at different sampling rates. Note the frequency distortion in the rightmost trace.

Nyquist condition: perfect reconstruction of a time domain signal from digital samples is possible if and only if the sampling rate exceeds two points per period of the fastest oscillation found in the signal. For example, correct digitization of an NMR spectrum located between -7 kHz and 15 kHz requires at least 30 kHz sampling rate.

In the NMR context, the consequences of *undersampling* a free induction decay are quite severe – high-frequency signals end up in the wrong place. This phenomenon is known as *aliasing*.

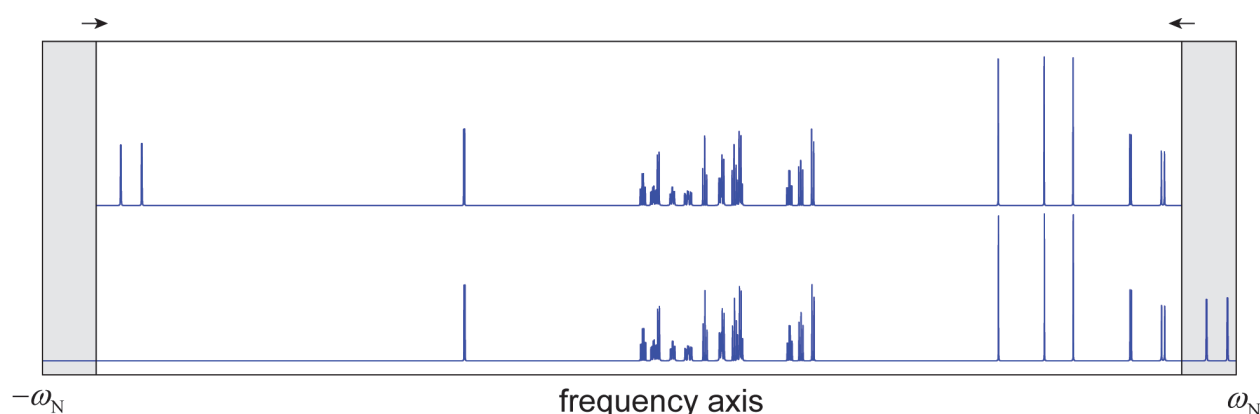


Figure 2. Frequency distortion caused by sub-Nyquist sampling. Note the two methyl peaks folding over the spectrum edge.

DSP software in modern spectrometers would often cause such signals to be suppressed rather than aliased, in which case they are lost completely. If a free induction decay is *oversampled*, the only practical consequence is increased noise level.

Conclusion: the best sampling possible in almost any experiment is the Nyquist sampling.

Digital resolution is the number of points per unit frequency used to represent a frequency domain signal in a discrete form.

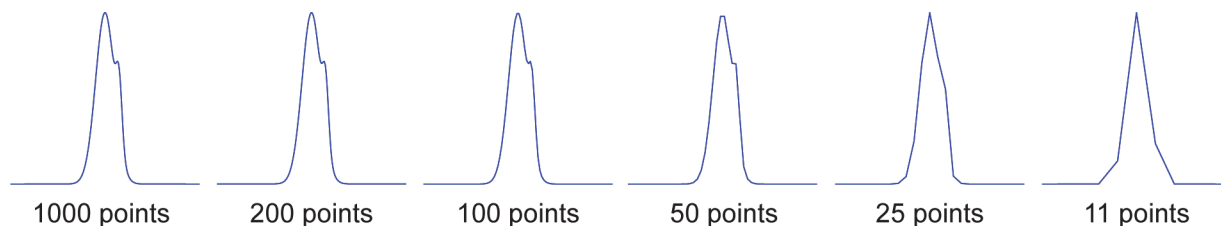


Figure 3. A pair of overlapping peaks digitized with different resolution. Note the gradual loss of detail from left to right.

The accuracy of spectral quantification (peak picking, integration, etc.) is directly determined by the digital resolution available. Integration in particular is very sensitive to the quality of the digital representation of the signal shape.

Harmonic interpolation and zerofilling

Padding the time-domain signal with zeros amounts to harmonic interpolation in the frequency domain:

$$\begin{aligned} \text{original spectrum: } F_n &= \frac{1}{N\sqrt{2\pi}} \sum_{k=0}^{N-1} f_k e^{-ikn/N} \\ \text{after zerofilling: } F_n &= \frac{1}{2N\sqrt{2\pi}} \sum_{k=0}^{2N-1} f_k e^{-ikn/2N} = \frac{1}{2N\sqrt{2\pi}} \sum_{k=0}^{N-1} f_k e^{-ikn/2N} \end{aligned} \quad (1)$$

After zerofilling the frequency sampling step (digital resolution) becomes finer, resulting in a more accurate quantification. Sharp signals often require zerofilling to be integrated accurately, particularly in multi-dimensional experiments.

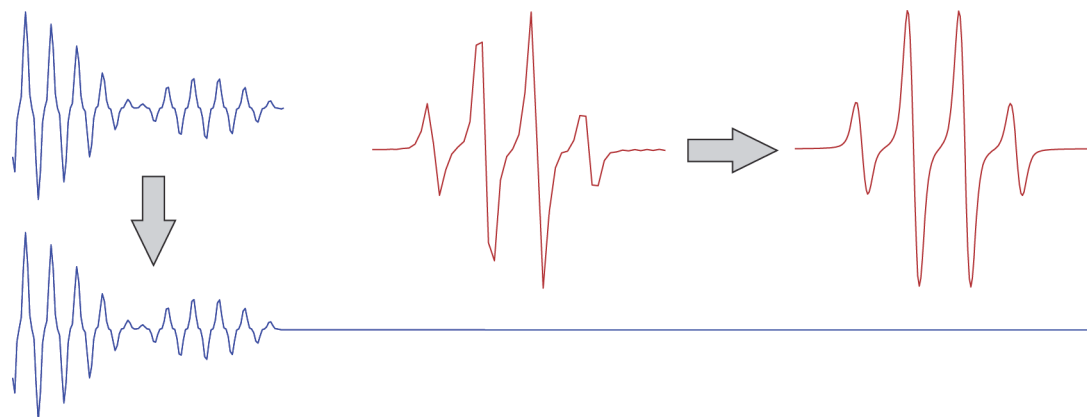


Figure 4. Improvement in digital resolution of an ESR signal resulting from time-domain zerofilling.

Band-pass filters

In many types of experiments (notably in chemical kinetics) the frequencies present in the signal are much lower than the frequencies of the noise. The following transformation performs a forward Fourier transform, multiplies the unwanted frequencies by zero and performs a backward Fourier transform:

$$f^*(t) = F_- \{ M(\omega) F_+ \{ f(t) \} \}, \quad M(\omega) = \begin{cases} 0 & \text{for unwanted } \omega \\ 1 & \text{for wanted } \omega \end{cases} \quad (2)$$

It is called a *band-pass filter* – it eliminates undesired frequencies from the time-domain signal. The selection function $M(\omega)$ is called a *magnitude transfer function*.

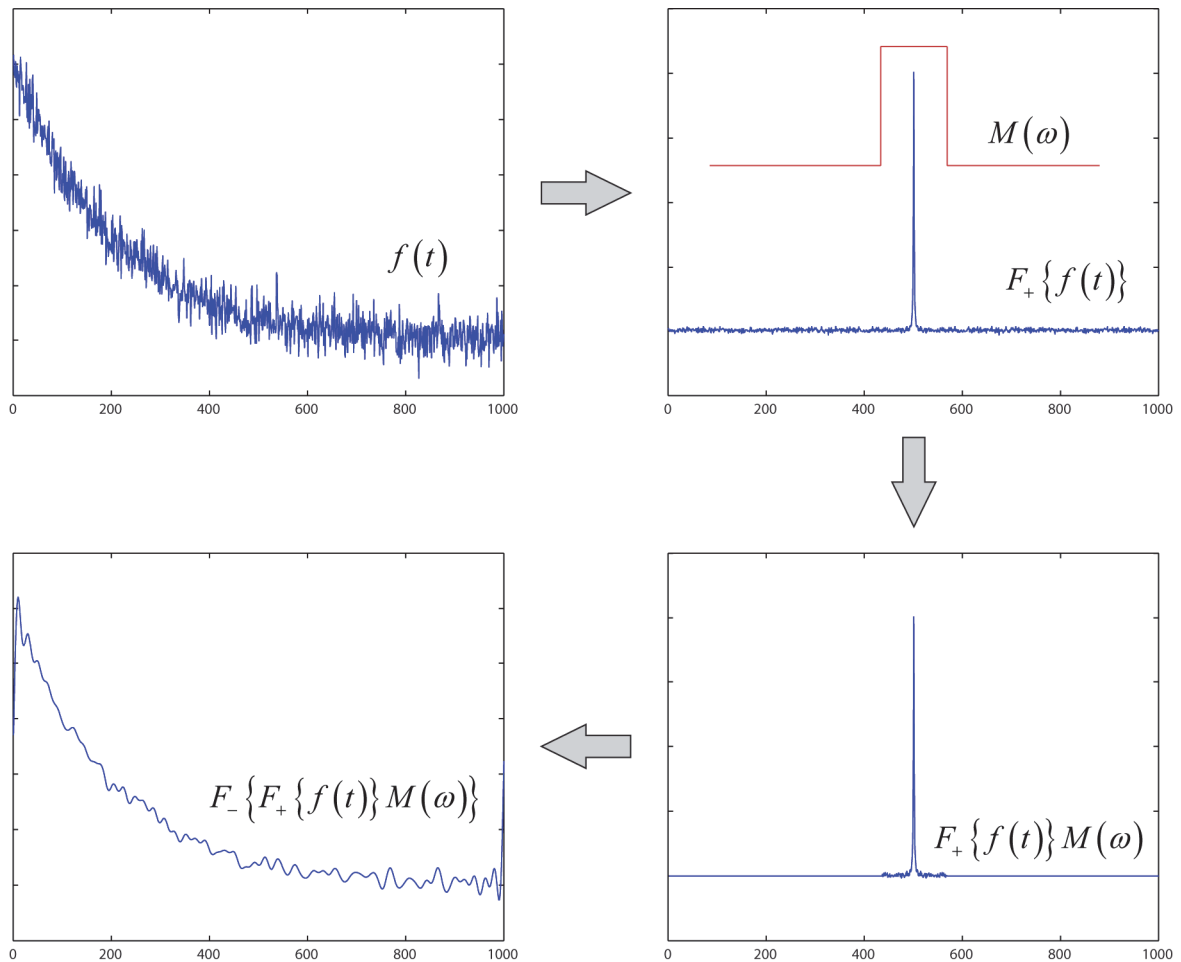


Figure 5. Band-pass filter applied to a noisy chemical kinetics decay trace.

Note the sharp jumps at the edges of the filtered signal – the Fourier transform implicitly assumes periodic boundary conditions and jumps to midpoints at discontinuities. This behaviour may be avoided by concatenating the signal with its mirror image prior to the forward Fourier transform and then taking the first half of the signal after the backward Fourier transform.

Despite their complicated appearance, band-pass filters are linear because the application of a discrete Fourier transform may be represented as an action by a matrix on a vector of digitized data values:

$$\left[F_+ \{ f(t) \} \right]_n = \frac{1}{N\sqrt{2\pi}} \sum_{k=0}^{N-1} f_k e^{-\frac{ikn}{N}} = \sum_k F_{nk} f_k, \quad F_{nk} = \frac{1}{N\sqrt{2\pi}} e^{-\frac{ikn}{N}} \quad (3)$$

Magnitude transfer function zeroing unwanted frequencies may be represented by a diagonal matrix. Therefore, $f^*(t) = \hat{F}^{-1} \hat{M} \hat{F} f(t)$, which is linear in $f(t)$.

Convolution filters (aka window functions)

Consider a smoothing filter (Figure 6) that collects data points from a certain interval in the original signal, applies Lorentzian weighting, calculates the average and places one point into the resulting “filtered” signal. Mathematically speaking, this is a convolution operation:

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(\nu) g(\omega - \nu) d\nu \quad \text{OR} \quad \tilde{f}(\omega) = f(\omega) * g(\omega) \quad (4)$$

It was demonstrated in Lecture 1 that this procedure may also be carried out by computing inverse Fourier transforms of $f(\omega)$ and $g(\omega)$, multiplying them together and performing a forward Fourier transform on the result.

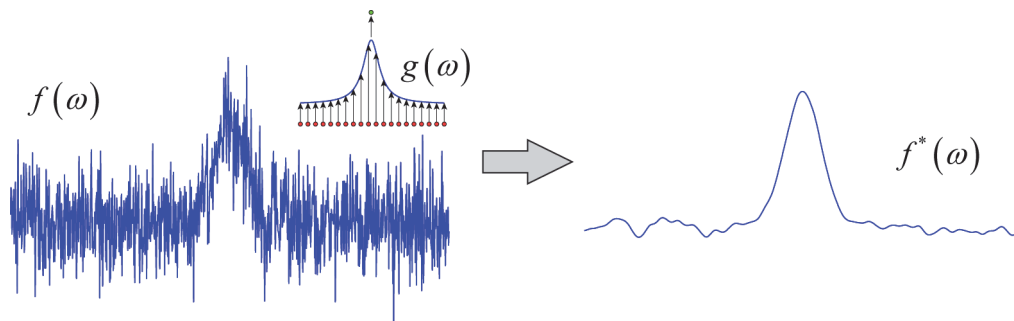


Figure 6. The result of the application of a Lorentzian local averaging filter to a noisy peak.

The inverse Fourier transform of $f(\omega)$ is the free induction decay, and the inverse FT of a Lorentzian function is an exponential function:

$$\int_{-\infty}^{\infty} \frac{1}{1 + \omega^2} e^{i\omega t} d\omega = e^{-t} \quad (5)$$

Therefore, *the application of a Lorentzian smoothing convolution filter in the frequency domain is equivalent to exponential apodization in the time domain.*

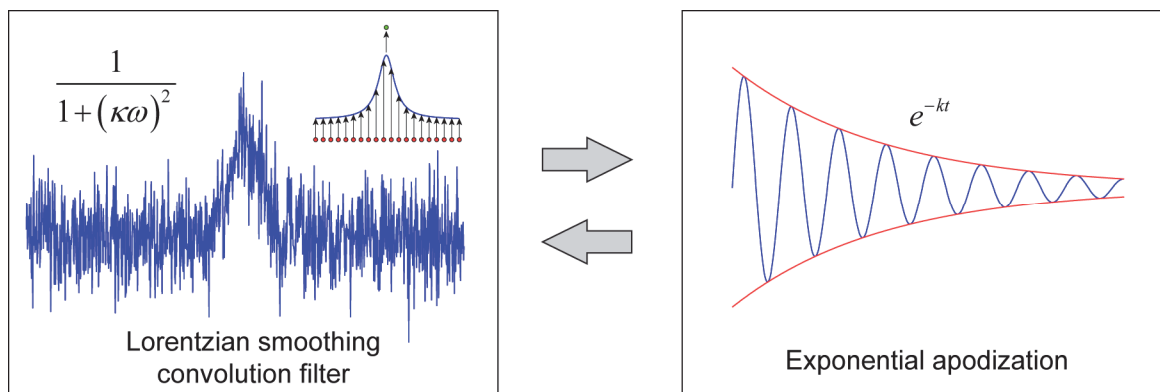


Figure 7. Illustration of the equivalence of Lorentzian convolution filtering in the frequency domain and exponential apodization in the time domain.

It is instructive to calculate the Fourier transform of a pure exponential decay and a pure sinusoidal modulation. For the exponential decay function we get:

$$f(t) = e^{-kt}, \quad F(\omega) = \int_0^{\infty} e^{-kt} e^{-i\omega t} dt = \dots = \frac{1}{k + i\omega} = \left(\frac{k}{k^2 + \omega^2} \right) - i \left(\frac{\omega}{k^2 + \omega^2} \right) \quad (6)$$

The real part of the result is the *Lorentzian curve*. The imaginary part is known as the *dispersion curve*. These are the real and the imaginary components of an NMR signal – so the Fourier transform of an exponential yields a standard NMR line shape at zero frequency. A Fourier transform of a pure non-

decaying oscillation $f(t) = e^{-i\omega_0 t}$ with a frequency ω_0 yields a delta-function (an infinitely sharp peak) at the frequency of the oscillation.

$$f(t) = e^{-i\omega_0 t}, \quad F(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-i\omega_0 t} e^{-i\omega t} dt = \dots = \sqrt{2\pi} \delta(\omega - \omega_0) \quad (7)$$

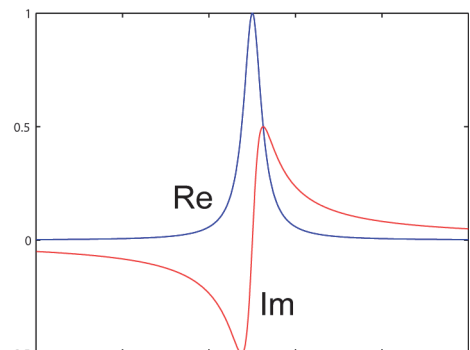


Figure 8. The real and the imaginary part of the Fourier transform of an exponential decay.

Therefore, in the absence of signal decay, all lines in the Fourier transform are infinitely sharp – it is the decay of the time domain signal that gives frequency domain lines their shape. The decay envelope is relatively easy to manipulate – one envelope may be divided away and another one multiplied in. The choice is often determined by the signal-to-noise budget. Some signal-to-noise ratio can be traded for some resolution and *vice versa*.

More generally, any form of time domain apodization is equivalent to a convolution filter in the frequency domain. The filter need not be a smoothing filter and may perform signal sharpening instead. There is a large number of case-specific apodization functions, for example:

- exponential weighting: aka soft low-pass filter – accelerates the decay and makes signals broader, but reduces noise.
- Lorentz-Gaussian transformation: replaces $\exp(-kt)$ envelope with $\exp(-kt^2)$. This makes signals narrower (Gaussian curve has shorter wings), but enhances the noise.

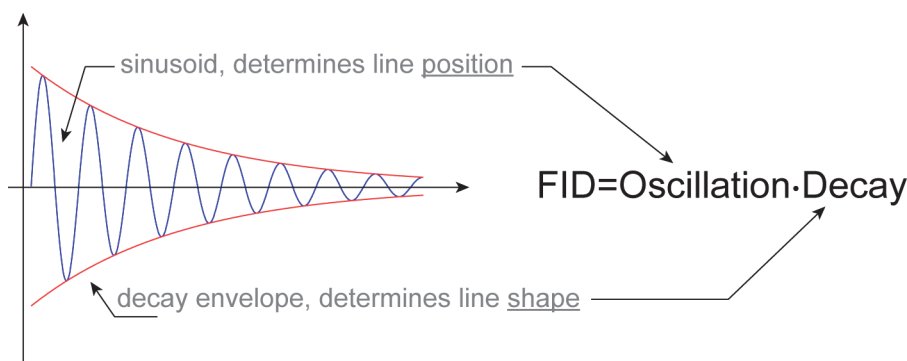


Figure 9. Schematic illustration of the factors affecting position and shape of a spectral line.

Notably, the Fourier transform of a Gaussian function is another Gaussian function – Gaussian apodization would therefore produce Gaussian line shapes. Another important observation is that the decay rate occurs in the denominator of Equation (6) – the longer the FID, the better the resolution.

Reference deconvolution

In a poorly shimmed NMR magnet, all spectral lines are distorted in *precisely* the same way:

$$S_{\text{exp}}(\omega) = \int_{-\infty}^{\infty} S_{\text{id}}(\omega') R(\omega - \omega') d\omega' \quad (8)$$

where $S_{\text{exp}}(\omega)$ is the distorted “experimental” spectrum, $S_{\text{id}}(\omega)$ is the “ideal” spectrum that we would like to recover and $R(\omega)$ is the “distortion” – the shape that a perfectly sharp signal would have acquired in the current magnet. We know from the convolution theorem given in Lecture 1 that the time domain version of Equation (8) is:

$$S_{\text{exp}}(t) = S_{\text{id}}(t)R(t) \quad (9)$$

from which we could easily recover $S_{\text{id}}(t)$, if we knew the envelope function of the distortion $R(t)$. This is where the “reference” comes in – if we have an *a priori* standalone signal (e.g. TMS) $s(\omega - \omega_0)$, then shifting it to zero and performing an inverse Fourier transform recovers the envelope:

$$R(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(\omega) e^{i\omega t} d\omega \quad (10)$$

If the desired FID envelope (e.g. an exponential decay) is $P(t)$, then the following transformation gets rid of the poor shimming:

$$S_{\text{id}}(\omega) = F_+ \left\{ F_- \left\{ S_{\text{exp}}(\omega) \right\} F_- \left\{ s(\omega) \right\}^{-1} P(t) \right\} \quad (11)$$

It should be noted that the signal-to-noise ratio cost of reference deconvolution is often huge.

SVD de-noising of data arrays

The singular value decomposition (SVD) of a complex matrix M is defined as follows:

$$M = U \Sigma V^\dagger \quad (12)$$

where $U = \{\vec{u}_1, \dots, \vec{u}_p\}$ is a complex matrix of *left singular vectors*, $V = \{\vec{v}_1, \dots, \vec{v}_q\}$ is a complex matrix of *right singular vectors* and Σ is a $p \times q$ diagonal matrix with positive real *singular values* $\{\sigma_1, \dots, \sigma_n\}$ along the diagonal. All arrays are sorted in such a way as to put the singular values in decreasing order. A rank- k approximation to the matrix M is then defined as:

$$M^{(k)} = \sum_{i=1}^k \sigma_i (\vec{u}_i \otimes \vec{v}_i^\dagger), \quad \|M - M^{(k)}\| = \sum_{i=k+1}^n \sigma_i^2 \quad (13)$$

where $\|\cdot\|$ is a Frobenius norm. It means that, to a good approximation, contributions to Equation (12) from near-zero singular values may be dropped.

In practice, the singular vectors associated with small singular values are often filled with random noise and the procedure improves the signal-to-noise ratio. Some components can correspond to undesired signals (such as the solvent) and may likewise be dropped. The SVD method, however, is only applicable to *arrays* of spectral data – e.g. 2D experiments or kinetics runs with multiple spectra recorded at regular intervals (as in Figure 10 below).

The SVD de-noising recipe is quite simple:

1. Run the SVD on the data matrix, drop the singular values and singular vectors associated with random noise (they correspond to the smallest singular values, the threshold is left to user discretion).

2. Re-construct the data using the remaining singular values and Equation (12).

The result can be quite dramatic. This is illustrated below with a ^{19}F NMR spectrum of an unfolding fluorotyrosine-labelled Green Fluorescent Protein.

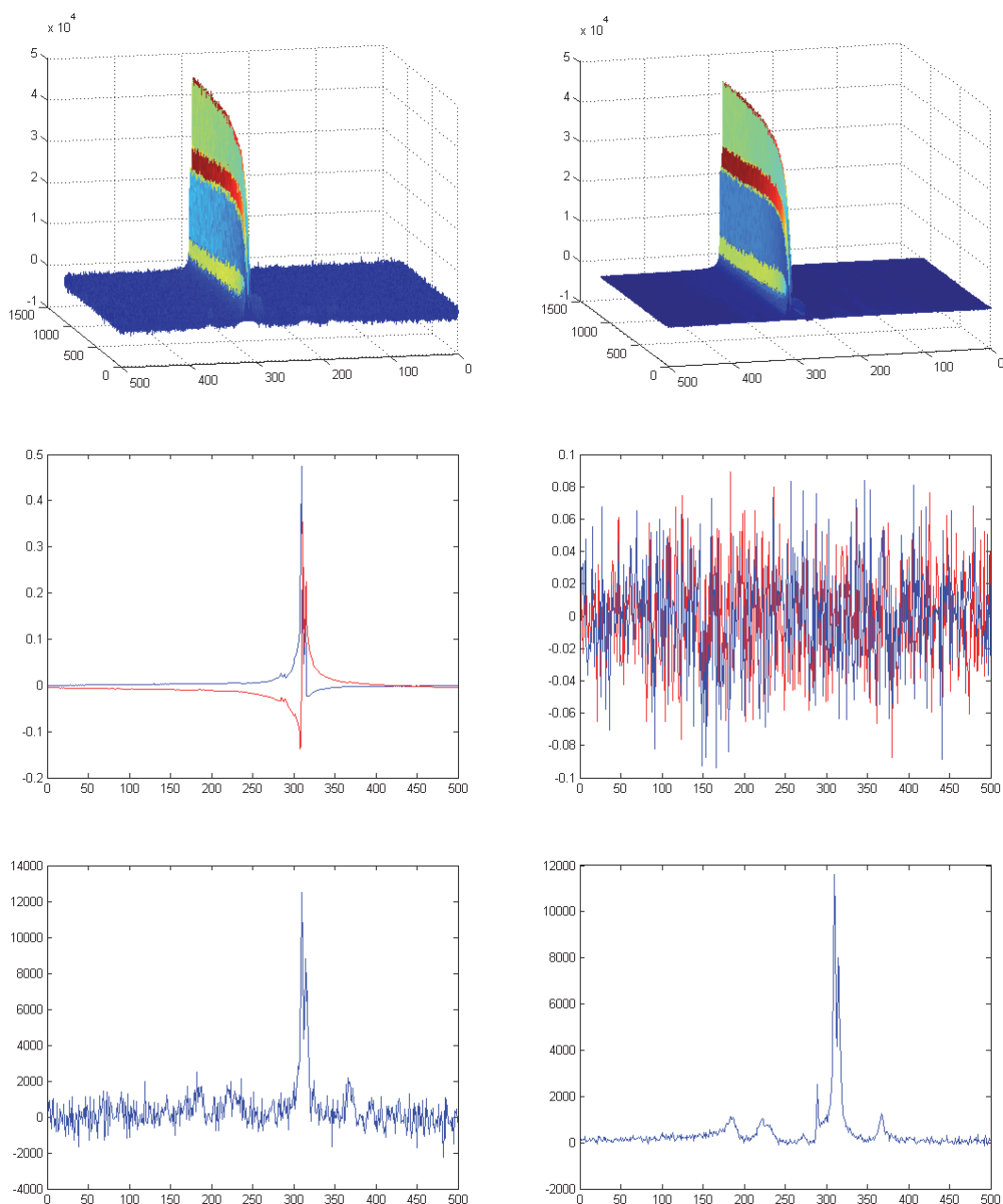


Figure 10. An illustration to SVD de-noising using ^{19}F NMR kinetics data (protein unfolding). **Upper left:** initial data array, **upper right:** de-noised data array, **middle left:** real and imaginary part of the largest left singular vector, **middle right:** real and imaginary part of the smallest singular vector, **lower left:** one of the spectra from the original data array, **lower right:** one of the spectra from the SVD de-noised data array.

The signal-to-noise ratio of the experiment is limited by low intrinsic sensitivity (^{19}F NMR lines are broad) and the requirement to capture the time dynamics of a rapid unfolding process. SVD de-noising improves the signal-to-noise ratio by at least a factor of ten in this case.

An interesting philosophical question is about the extent to which such filtering may be applied to experimental data – as of 2005, strict guidelines are in place for digital image manipulation. The communi-

ty consensus at the moment is that *linear* signal processing can be applied freely, whereas *non-linear* methods must be declared and detailed in full.

Reconstructing complex signals

A common situation, particularly in EPR spectroscopy, is that the imaginary part of the spectrum is missing and it can no longer be phased. Fortunately, *causality* (“no output before the input”) of linear time-invariant systems places sufficient constraints on the properties of their response functions for the imaginary part of the spectrum to be reconstructed from the real part.

If the excitation pulse arrives at time zero, the response signal $f(t)$ must vanish for $t < 0$, and therefore should not change if we multiply it by the *Heaviside step function* $h(t)$:

$$f(t)c(t) = f(t)h(t), \quad h(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t > 0 \end{cases} = \int_{-\infty}^t \delta(\tau) d\tau, \quad c(t) = 1 \quad (14)$$

In the frequency domain this means that the corresponding convolutions should be equal, that is:

$$f(\omega) * c(\omega) = f(\omega) * h(\omega) \quad (15)$$

The Fourier transform of a constant function we already know from the previous lecture:

$$c(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} c(t) e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} dt = \sqrt{2\pi} \delta(\omega) \quad (16)$$

The Fourier transform of the Heaviside function is a very subtle matter – we will not dwell on it here and simply state the textbook result:

$$h(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \left(\pi \delta(\omega) - \frac{i}{\omega} \right) \quad (17)$$

Substitution of these results into Equation (15) yields:

$$\sqrt{2\pi} f(\omega) * \delta(\omega) = \frac{1}{\sqrt{2\pi}} f(\omega) * \left(\pi \delta(\omega) - \frac{i}{\omega} \right) \quad (18)$$

According to the definition of the delta function, $f(\omega) * \delta(\omega) = f(\omega)$. After applying this simplification, we get the following convolution identity relating $f(\omega)$ to itself:

$$f(\omega) = \frac{1}{i\pi} f(\omega) * \frac{1}{\omega} \quad \Leftrightarrow \quad f(\omega) = \frac{1}{i\pi} \int_{-\infty}^{\infty} \frac{f(\nu)}{\nu - \omega} d\nu \quad (19)$$

This identity is known as *Hilbert transform*. After separating the real and the imaginary parts of $f(\omega)$, we arrive at *Kramers-Kronig relation*:

$$f_{\text{re}}(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f_{\text{im}}(\nu)}{\nu - \omega} d\nu \quad f_{\text{im}}(\omega) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f_{\text{re}}(\nu)}{\nu - \omega} d\nu \quad (20)$$

With the practically useful result that the real and the imaginary part of an NMR / ESR spectrum are not independent and may be calculated from one another using the Hilbert transform. In practice this means that, if we digitize a spectrum from a 1960 printed paper, we should still be able to phase it.

Correcting phase distortions

In a spectrometer with a perfectly linear (*i.e.* frequency-independent) detector and perfectly sharp pulses, the complex phase of the detected time-domain signal would be frequency-independent:

$$f(t) = e^{i\omega_0 t + i\varphi - kt} = e^{i\varphi} e^{i\omega_0 t - kt} \quad (21)$$

Because $e^{i\varphi}$ is a constant, it would also appear in front of the Fourier transform of $f(t)$:

$$\begin{aligned} f(\omega) &= \frac{e^{i\varphi}}{\sqrt{2\pi}} \int_0^{\infty} e^{i\omega_0 t - kt} e^{-i\omega t} dt = \dots = \frac{e^{i\varphi}}{\sqrt{2\pi}} \left[\frac{1}{k + i(\omega - \omega_0)} \right] = \\ &= \frac{e^{i\varphi}}{\sqrt{2\pi}} \left[\frac{k - i(\omega - \omega_0)}{k^2 + (\omega - \omega_0)^2} \right] = \frac{e^{i\varphi}}{\sqrt{2\pi}} \left[\frac{k}{k^2 + (\omega - \omega_0)^2} + \frac{-i(\omega - \omega_0)}{k^2 + (\omega - \omega_0)^2} \right] \end{aligned} \quad (22)$$

with the result that the real and the imaginary part of the spectrum

$$f_{\text{re}}(\omega) = \frac{k}{k^2 + (\omega - \omega_0)^2} \quad f_{\text{im}}(\omega) = \frac{-i(\omega - \omega_0)}{k^2 + (\omega - \omega_0)^2} \quad (23)$$

become mixed or, equivalently, rotated in the complex plane relative to one another:

$$\begin{aligned} R &= f_{\text{re}}(\omega) \cos \varphi + f_{\text{im}}(\omega) \sin \varphi \\ I &= -f_{\text{re}}(\omega) \sin \varphi + f_{\text{im}}(\omega) \cos \varphi \end{aligned} \quad (24)$$

The *phase correction* procedure (manual or automatic) consists in finding such φ as would minimize the amount of dispersive signal $f_{\text{im}}(\omega)$ in R .

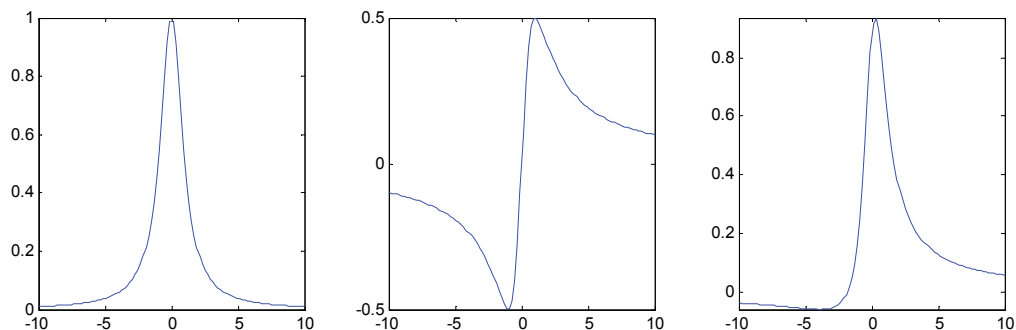


Figure 11. Left: real part of an NMR signal. Middle: imaginary part. Right: a phase-twisted line resulting from mixing the imaginary part into the real one.

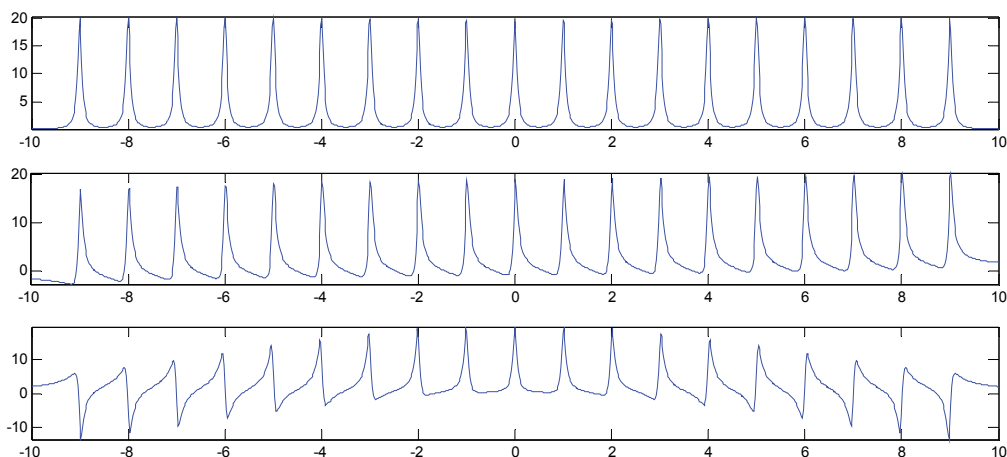


Figure 12. **Top:** correctly phased spectrum. **Middle:** spectrum with a zero-order phase distortion. **Bottom:** spectrum with a linear frequency-dependent phase distortion.

In real-life situations the phase correction required is rarely frequency-independent, and so a more sophisticated polynomial phase correction is employed:

$$\varphi(\omega) = \varphi_0 + \varphi_1\omega + \varphi_2\omega^2 + \dots \quad (25)$$

where the index k in φ_k is known as the *correction order*. With frequency-dependent phase correction, Equations (24) must be applied to each spectrum point separately. The most common source of the linear phase distortion is the modulation theorem:

$$F_+ \{f(t)\} = f(\omega) \quad F_+ \{f(t-t_0)\} = e^{i\omega t_0} f(\omega), \quad (26)$$

In practical spectroscopy a time shift in the free induction decay is inevitable because spectrometer electronics needs time to switch from pulse transmission to signal detection.

Correcting point defects in time domain signals

The conventional way of correcting a rolling baseline in a spectrum (fitting a polynomial through the signal-free areas and subtracting it out) is as simple as it is indefensible – essentially a phenomenological correction with no physical meaning to it. A particular problem that it often suffers from is deciding (particularly with noisy data) which regions are “signal-free”. Such methods should be avoided whenever possible and replaced with physically meaningful ones.

Most types of baseline distortions are very broad features, meaning that they are determined by just two or three initial points in the time domain.

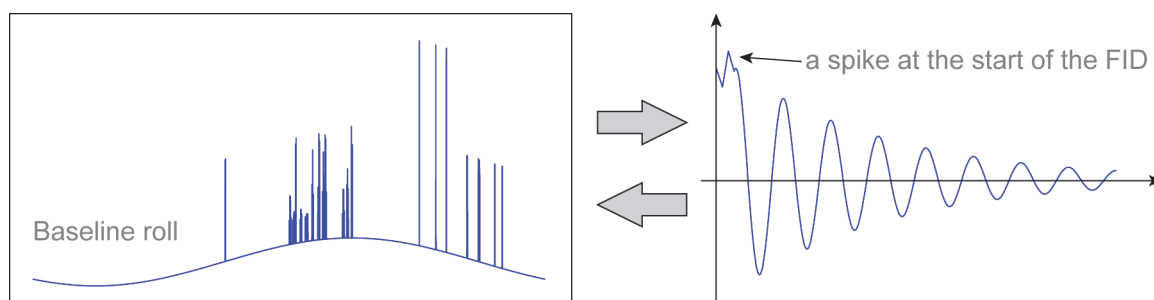


Figure 13. Baseline roll phenomenon caused by point defects in the time domain data.

The causes may vary – most commonly the spike at the start of the FID comes from imperfect electronics switching (some of the pulse ‘breaks’ into the detection circuit) or probe parts containing the nucleus that is being detected (often the case with ^{19}F). Another popular cause is a colleague accidentally hitting the ‘eject’ button in the middle of a 2D experiment.

A general way of correcting all types of point defects in the time domain is called *linear prediction*. It makes a (reasonable) assumption that the future values of a discrete time-domain signal may be obtained approximately as a linear combination of the past values:

$$f_n = \sum_{k=1}^m a_k f_{n-k} \quad (27)$$

The number of coefficients m determines prediction accuracy (the more, the better, but generally m should be greater than the number of distinct peaks expected in the spectrum). The coefficients a_k are found from the so-called “training set”.



Figure 14. Schematic illustration of training and prediction steps in the linear prediction procedure.

Let us assume that we have $2m$ reliable undistorted points in the data. Then:

$$\begin{cases} f_{m+l} = \sum_{k=1}^m a_k f_{m+l-k} \end{cases} \quad (28)$$

so the known points give a system of linear equations for the prediction coefficients a_k . If more points than $2m$ is available for training, the system would be usually overdetermined and should be solved in a least squares sense. The resulting coefficients are used to reconstruct the missing points by sequentially applying Equation (27) to the data.

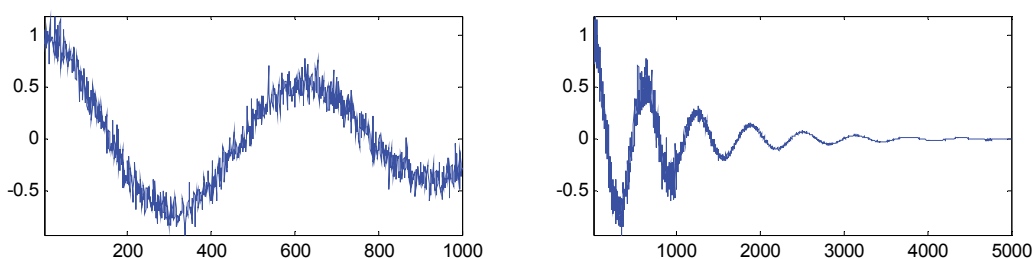


Figure 15. A truncated oscillatory signal and its linear prediction to five times the duration with 500 predictor coefficients in Equation (27).

Linear prediction is useful for:

- dealing with most types of baseline distortion (“*backward LP*”).
- completing truncated FIDs as an alternative to zerofilling (“*forward LP*”).
- eliminating point defects within the experimentally acquired time-domain data.

Reconstructing badly corrupted data

In situations where large groups of time domain points are missing, a more radical approach, called *retrofitting*, may be taken, where the “Fourier transform” (which would no longer be uniquely defined, hence the quotes) is reconstructed using a least squares fit to the available time domain data:

$$F_+ \{f(t)\} = \arg \min_{f(\omega)} \left\{ \left\| f(t) - F_- \{f(\omega)\} \right\|^2 + R \{f(\omega)\} \right\} \quad (29)$$

where the norm is calculated using the available data points and $R \{ \}$ is a *regularization functional*, intended to constrain the solution to a particular class of functions when a solution is not unique. Popular choices for $R \{ \}$ include *Tikhonov regularization*, which has the following general form:

$$R \{f(\omega)\} = \left\| \hat{R}f(\omega) \right\|^2 \quad (30)$$

where \hat{R} is a linear operator. In particular, Tikhonov regularization is often set to favour smooth solutions, *i.e.* those for which the norm of the first derivative is small:

$$R \{f(\omega)\} = k \left\| \frac{d}{d\omega} f(\omega) \right\|^2, \quad k > 0 \quad (31)$$

The “Fourier transform” in Equation (29) is also applicable to non-uniformly sampled data, which is encountered more and more often as researchers try to save spectrometer time by only detecting time domain points that they absolutely have to detect.

Maximum entropy method

The real part of a magnetic resonance spectrum is typically non-negative and may be viewed, after normalization, as a probability distribution for resonance frequencies of the spin system. For a given statistical distribution, *information entropy* is a measure of the total amount of uncertainty that the distribution contains. For N equally spaced frequencies with a set of probabilities $\{p_k\}$, entropy is given by:

$$S = -k \sum_{n=1}^N p_n \ln p_n \quad (32)$$

In situations where the time-domain data is noisy or limited, and therefore the minimum prescribed by the first term in Equation (29) is not unique, selecting the solution with maximum uncertainty guarantees that the information content of the resulting spectrum has not been artificially inflated. The total error functional for the least-squares retrofitting using maximum entropy regularization is therefore:

$$\Omega = \left\| f(t) - F_- \{f(\omega)\} \right\|^2 - k \sum_{n=1}^N p_n \ln p_n \quad (33)$$

where the weighting coefficient k is left to user discretion – a considerable volume of literature exists on its proper choice in various circumstances.

Literature

1. J.C. Hoch, *NMR Data Processing*, Wiley, 1996.
2. A.E. Derome, *Modern NMR Techniques for Chemistry Research*, Pergamon, 1987.
3. J.T. Moore, *NMR Spectroscopy: Processing Strategies*, Wiley, 2000.

Spin Dynamics, Module I, Lecture 03 – Dr Ilya Kuprov, University of Southampton, 2013
(for all lecture notes and video records see <http://spindynamics.org>)

4. M. Bydder, J. Du, <http://dx.doi.org/10.1016/j.mri.2006.03.006>
5. <http://www.nature.com/nature/journal/v311/n5985/abs/311446a0.html>
6. [http://dx.doi.org/10.1016/0022-2364\(91\)90277-Z](http://dx.doi.org/10.1016/0022-2364(91)90277-Z)