Johnson Noise And The Measurement Of Absolute Temperature

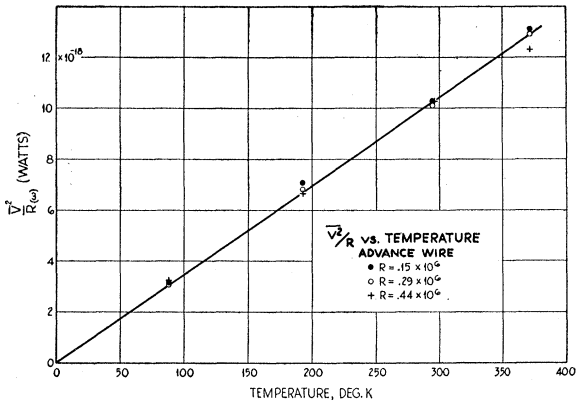
© *Frank van Kann, Department of Physics, The University Of Western Australia, Perth, 6009, Australia*

frank.van.kann@uwa.edu.au

1 Theoretical Background

This project is based on similar work at the University of Oregon (see "Undergraduate Experiment in Noise Thermometry", P. Kittel, W.R. Hackleman and R.J. Donnelly, Am. J. Phys., 46(1), pp. 94-100, Jan 1978). The fundamental concepts are the same but the procedure is substantially streamlined by the use of modern electronic instrumentation.

The microscopic random motion of the electrons in a conductor gives rise to macroscopic but tiny voltage fluctuations, which can be observed using sensitive instruments. These voltage fluctuations were first discovered by J.B. Johnson at Bell Laboratories in 1927 when he observed excess noise in his electronic amplifiers (see "Thermal Agitation of Electricity In Conductors", J.B. Johnson, Phys. Rev., (32) pp. 97-109, July 1928). It is worth reading his paper (it's available in LMS) as it conveys the excitement of the discovery. Figure 1 is an extract from the paper and summarises the observations.



. Apparent power vs. temperature, for Advance wire resistances.

Figure 1 *Summary of Johnson's original measurements*

This plots the average power of the voltage fluctuations for a number of different resistors as a function of the temperature of the resistor. It is essential to plot the average power, averaged over a suitable time interval because the instantaneous power is a fluctuating random variable but the average converges to a particular value as shown in the figure.

Strictly speaking, in the language of Many Particle Physics, the average should be taken over an infinite ensemble of identical systems but this is impractical and a time average approximates this provided that the system is stationary.

It is noteworthy that H. Nyquist (also at Bell) published the theory of Johnson noise in the same edition of Phys. Rev. (pp 103-107). This work underpins what is now known as the Fluctuation-Dissipation Theorem (see "Irreversibility and Generalized Noise", H.B. Callen, T.A. Welton, *Phys. Rev.* **83** (1) pp 34–40, 1951).

This applies to any physical system, which contains a mechanism of converting energy (for example electrical or mechanical) into heat. It is that same mechanism, which also gives rise to the observed random fluctuations, including both Johnson noise and Brownian motion.

The aim of this project is to measure this Johnson noise and make quantitative comparison with theoretical predictions based on statistical mechanics and thermodynamics.

The electrical circuit for the measurement of the noise in a resistor R can be represented (in accordance with Thevenin's theorem) as an ideal voltage source V_R in series with the resistance R as shown in Figure 2.

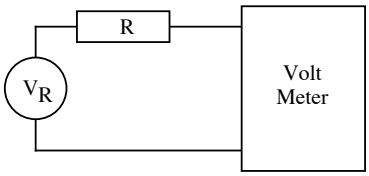


Figure 2 *Thevenin equivalent circuit for the measurement of Johnson noise.*

An equivalent representation (in accordance with Norton's theorem) consists of an ideal current source in parallel with a resistor as shown Figure 3. In this equivalent circuit the current source is given by

$i_R = \frac{V_R}{R}$ with parallel resistance R.

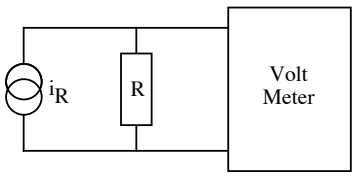


Figure 3 *Norton equivalent circuit for the measurement of Johnson noise.*

In either representation the electrical signal (voltage or current) is a random variable, which cannot be predicted in detail but which is described in terms of its statistical properties. These include for example the form of the probability distribution function and the parameters, which define it. There are two ways to characterise the probability.

The first is in terms of a parent population, which, consists of an infinite set (ensemble) of identical systems. The random variable (in this case the current or voltage) in each of these is sampled at the same instant of time to produce the probability distribution. If the system is "stationary", this distribution is independent of the time at which the sample is taken (for a discussion of this, see Reif, "Fundamentals of Statistical and Thermal Physics", McGraw-Hill, 1965).

A second way to obtain the probability distribution is to take an infinite set of consecutive samples from a single system. If the system is both "stationary" and "ergodic" (see Reif) then these two averages are the same.

For a resistor, which is in thermal equilibrium with its environment, the voltage noise observed by Johnson is described in terms of its average power

$$\langle P \rangle = \alpha T \qquad 1$$

or equivalently

$$\langle V^2 \rangle = \alpha R T \,, \qquad 2$$

where the term on the left is the mean square of the observed voltage and α is a constant. This depends on both a fundamental constant (Boltzmann's constant) and also on the properties of the measuring apparatus, namely its amplification factor and its bandwidth.

Then Nyquist's theorem for an electrical system can be stated as

$$\langle V^2 \rangle = 4 k_B R T \Delta f \qquad 3$$

or

$$V_{rms} \equiv \sqrt{\langle V^2 \rangle} = \sqrt{4 k_B R T \Delta f} \,, \qquad 4$$

where k_B is Boltzmann's constant, T is the thermodynamic temperature of the resistor, R its resistance and Δf is the *equivalent noise-bandwidth* (ENBW) of the measuring apparatus and is described below. Implicit in these equations is that the mean power is independent of frequency, but depends only on the frequency interval of the measurement. The term "white noise" is used to describe this frequency-independent noise;

$$\left\langle \frac{V^2}{\Delta f} \right\rangle = 4 k_B R T \,. \qquad 5$$

Equation 5 describes a fundamental characteristic of the signal source and the term on the left is the mean-square voltage per unit bandwidth, which is valid for even infinitesimally small bandwidths. This mean-square voltage per unit bandwidth is often denoted by the symbol $S_V(f)$ but in the case of Johnson noise it is independent of frequency and can be written as $S_V(0)$ or simply S_V, giving

$$S_V = 4 k_B R T \,. \qquad 6$$

If the bandwidth of the measuring instrument is known, the mean-square voltage measured is simply

$$\langle V^2 \rangle = S_V \Delta f = 4 k_B R T \Delta f \,. \qquad 7$$

No physical apparatus can measure to infinitely high frequency. Even the most perfectly constructed resistor inevitably has "stray" capacitance – particularly once leads are attached – and this forms a low pass filter, which attenuates the higher frequencies. This defines an effective bandwidth or frequency interval within which it can function and this is the Δf in equation 7. The actual attenuation of such a filter is plotted as the solid curve in figure 4. This shows a very gradual droop and we must somehow decide what is the effective "cut-off" frequency Δf.

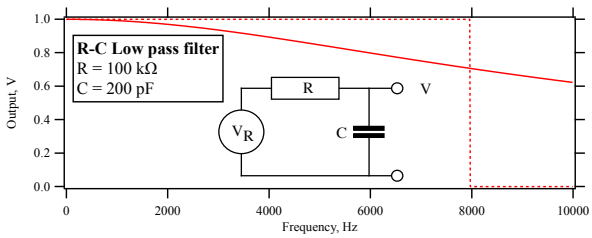


Figure 4 The R-C low-pass filter.

The answer is simply to construct a hypothetical filter with a sharp cut-off (depicted by the dashed line in the figure), such the output power is identical to that of the actual filter. The derivation is outside the scope of the present work but the result for the case of white noise is simply

$$\Delta f = \frac{\pi^2}{RC} \,. \qquad 8$$

There are many other electronic filters, and each has it's own ENBW.

The volt meter depicted in figures 2 and 3 is more than a simple volt meter. Firstly, it must contain an amplifier, because the voltage fluctuations are quite tiny and secondly it must have a means to determine average of the square of the voltage. Johnson had to use a specially built bolometer to measure the mean-square but we can do it simply by digitally sampling the voltage (at a sufficiently fast sample rate) and computing the mean-square numerically.

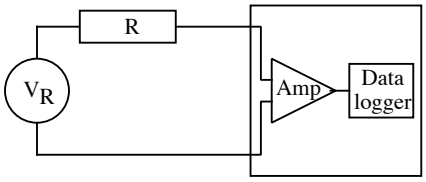


Figure 5 The actual measuring apparatus

The amplifier also contributes to the overall ENBW of the apparatus, and this must be determined experimentally as outlined in the next section.

2 Measurement of Noise

The Johnson noise fluctuations are very small; for example if R = 1 MΩ then at room temperature the voltage signal is only 127 nV rms if the bandwidth is 1 Hz and still only 18 μV rms over the entire audio frequency spectrum. Nevertheless, the noise can be observed and measured as a function of both R and T using a sufficiently sensitive amplifier together with suitable recording instruments as shown in the block diagram of Figure 6.

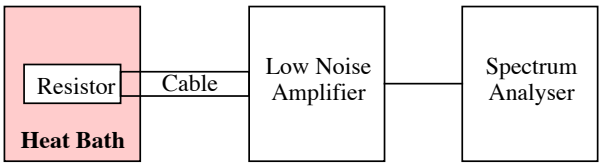


Figure 6 Block diagram of the apparatus for measuring Johnson noise

The resistor, which is the source of the noise, is connected to the input of the low noise amplifier (LNA).

The resistor is in equilibrium with a heat bath at temperature T, which may be different from the ambient temperature of the LNA and the other instruments. Thus, the resistor is shown connected by means of a short length of thin coaxial cable, which can withstand the required temperature difference. Two forms of heat bath will be used in the experiment, one the earth's atmosphere at 300 K and the other of a thermos flask of boiling liquid nitrogen at 77 K.

The amplified noise appears at the output of the LNA and its frequency spectrum is measured directly by means of an audio frequency spectrum analyser. This is in the form of a software package, which uses the sound inputs of a personal computer to digitise the signal and then computes the spectrum of the sampled wave form.

The block diagram of figure 6 is shown schematically in figure 7.

Here V_R is the Johnson noise "signal" and R is the resistance, which produces it. The capacitance C represents the stray capacitance of the coaxial cable and the amplifier input, estimated to be about 30 pF. The LNA is represented by its frequency response function A(f).

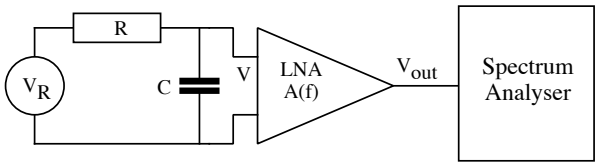


Figure 7 Schematic diagram of the apparatus for measuring Johnson noise

3 Frequency Dependence

The unwanted but unavoidable stray capacitance at the input forms a low pass filter, which attenuates the signal at high frequency. The magnitude of the frequency response of this low pass filter is given by

$$|G(f)| = \frac{1}{\sqrt{1-(f/f_o)^2}},$$

where $f_o = \frac{1}{2\pi RC}$. The ENBW of such a filter is

$$\Delta f = \frac{\pi^2}{RC} = \frac{\pi f_o}{2}.$$

The amplifier itself has a frequency response, which must also be considered. It incorporates a band pass filter between about 150Hz and 5kHz as shown in figure 5, where the large dots represent the measured transfer function.

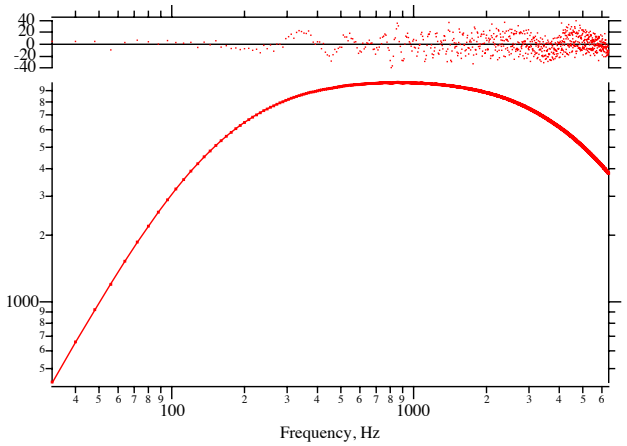


Figure 8 Frequency response of the LNA

This frequency dependence is stated explicitly by writing the magnitude of the gain A(f) as a function of frequency,

$$|A(f)| = \frac{A_o\left(\frac{f}{f_H}\right)}{\sqrt{\left(1+\left(\frac{f}{f_H}\right)^2\right)\left(1+\left(\frac{f}{f_L}\right)^2\right)}}.$$

The best fit function of this form, with A_o = 1.04 x 10^4, f_H = 154 Hz and f_L = 4.95kHz is plotted as the solid curve and is evidently a good fit to the observed data. The small dots show the residuals, or differences between the data and the model function.

When measuring Johnson noise, the spectrum analyser measures the Fourier transform of the output of the amplifier given by

$$V_{out}(f) = G(f)\, A(f)\, V_R.$$

The frequency response of both G(f) and A(f) must be known in order to obtain V_R (which is independent of frequency) from the measured $V_{out}(f)$

The required transfer function has been obtained experimentally for each amplifier, an example of which is shown in figure 8. This frequency response is effectively an array of band-pass filters, each one having its own equivalent noise bandwidth (ENBW) and each one being centred on one of the discrete frequency "bins" plotted in the graph. The actual output signal is the sum of these and therefore the ENBW of the amplifier can be conveniently obtained by summing the ENBWs of all of the individual bins.

$$\Delta f_{amp} = \sum_{0}^{N} \left|A(f_i)\right|^2 \delta f\,,$$

where δf is the bin width and the index i runs from 0 to the N, the total number of bins in the graph.

Moreover, it is convenient to define an ENBW for a particular subset of contiguous bins, spanning the frequency interval between f_1 and f_2. This is denoted by Δf_e and is given by

$$\Delta f_e = \sum_{(f_1)}^{(f_2)} \left|A(f_i)\right|^2 \delta f\,,$$

where (f_1) denotes the index of f_1 and similarly for f_2. This is useful in the case that you wish to ignore some parts of the spectrum because they are contaminated by unwanted interference, for example the first few harmonics of the power mains. As described in section 5 the ENBW (including the effect of the shunt capacitance) has been measured for the amplifiers used in this experiment. The chosen frequency interval is from 400 Hz to 5 kHz. The lower limit is chosen to avoid mains harmonics and the upper limit is chosen because the effect of shunt capacitance is small below that frequency.

Defining the symbol $\left\langle V_{out}^2 \right\rangle_{f_1}^{f_2}$ to represent the measured mean square voltage in the frequency interval from f_1 to f_2, this gives

$$\left\langle V_{out}^2 \right\rangle_{f_1}^{f_2} = 4 k_B T R \Delta f_e\,.$$

Thus the measurement of this quantity allows you to obtain Boltzmann's constant by rearranging this equation to obtain

$$k_B = \frac{\left\langle V_{out}^2 \right\rangle_{f_1}^{f_2}}{4 R T \Delta f_e}\,.$$

Alternately, if you know Boltzmann's constant, you could rearrange it to obtain T, effectively making the apparatus into an absolute thermometer.

The measurement of $\left\langle V_{out}^2 \right\rangle_{f_1}^{f_2}$ can be obtained from the Fourier transform of the amplifier output, which is recorded by the spectrum analyser. The calculation is implemented within the app and makes use of Parseval's theorem to obtain the result. You can perform this calculation at the press of a button (namely the ***Area*** button).

4 Interpretation Of Measured Power Spectra

Some analysis is required in order to make a quantitative comparison between the observed spectral density and that predicted by Nyquist's theorem as the effects of both the amplifier noise the frequency dependence of the circuit must be taken into account.

To simplify analysis of the circuit, it is convenient to use the Norton equivalent representation of the circuit as shown in figure 10.

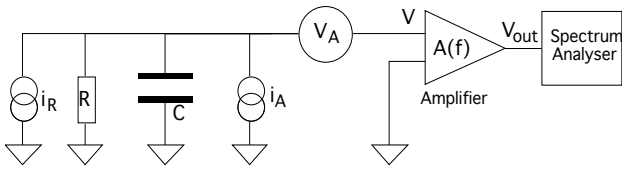


Figure 10 Equivalent circuit with the amplifier noise sources included.

This shows that for any particular set of values of R and T it is impossible to distinguish between the amplifier current noise and the Johnson noise. They are connected in parallel and both have the same white spectral characteristic.

An elementary analysis of the circuit in figure 7 gives the voltage V at the input of the amplifier,

$$V = V_A + (i_A + i_R)\, Z,$$

where Z is the impedance of the parallel combination of R and C. In the approximation that the effect of C can be neglected, this gives Z = R, and

$$V = V_A + (i_A + i_R)\, R = V_A + i_A\, R + V_R\,.$$

For the amplifier used in the experiment (LT1793), the current noise is negligible and indeed, it is difficult to devise an experiment to measure it.

The dependence of the power spectral density on both resistance and temperature should be experimentally tested. It would be desirable to cover the entire two-dimensional grid of R and T, but this would take too much time.

For the first part of this project it is adequate to make two measurements with a single resistor at room temperature (300K) and the boiling point of liquid nitrogen (77 K).

Ideally, the experiment would be performed over a large range of R and T. However, for the larger values of R the effect of C becomes significant and must be allowed for in the analysis.

The approximation made in the analysis above, $Z \approx R$ is no longer valid and must be replaced by

$$Z = \frac{Z_C R}{Z_C + R}\,, \text{ where } Z_C = \frac{1}{j2\pi C}\,.$$

This gives Z = R G(f), where G(f) is the transfer function of the R-C low pass filter defined in section 3.

The effect of this is illustrated in figure 11. Here, the ENBW is calculated using f_1 = 400Hz and f_2 = 5kHz.

The ENBW is a function of R and a calibration must be performed for each value of R being used.

The example of figure 11 shows the ENBW for amplifier number 10. The table in the graph lists the ENBW between 400 Hz and 5.00 kHz for resistor values between 0 and 10 MΩ. There variation for resistors below 1 MΩ is less than a few %, so the R = 0 value could be used for these.

Also, it is found that the variation of the ENBW between different amplifiers is small compared with the reproducibility of the measurement. Thus, the value of 2.8×10^{11} could be used with 3% uncertainty for all the amplifiers with any resistor values up to 1 MΩ.

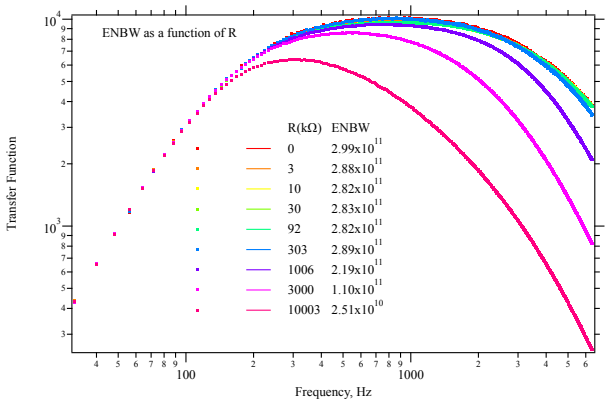


Figure 11 The effect of the R-C low pass filter at the input of the LNA number 10.

5 Precision of the Measurements

Two types of error affect the measurement, systematic and random. The random nature of the power spectra is evident from the measurements and since the statistical parameters of the sources of noise are known, the effects of random errors can be calculated using statistical analysis. However, systematic errors are less obvious and any possible sources of these must be considered. These include

- Accuracy of electrical measurements

The measurements require a knowledge of both voltage and resistance. These can be measured with high accuracy, typically 1 part in 10^5 using a suitable commercial digital volt meter, such as the HP34401A. The accuracy of the sound inputs is not so great but this can be calibrated using a signal source and comparing the measured amplitude with an AC volt meter.

- Temperature measurement

Temperature measurement represents the biggest source of systematic error. Several thermometers are available but their calibration is uncertain.

The mercury in glass thermometer can be read to a resolution of 0.5°C and is likely to be accurate at this level.

The K-type thermocouple is connected to a hand held multimeter in temperature mode and has a resolution of 1°C but its accuracy is likely to be somewhat worse than this; it gives readings inconsistent with the mercury in glass device.

The platinum resistance thermometer (PRT) is an encapsulated platinum resistor with a quoted temperature coefficient of 0.385%/°C and resistance of 100.0Ω at 0°C, with an accuracy of 0.12%. This corresponds to about 0.3C at room temperature and its readings are consistent with the mercury in glass device at this level.

Using the HP34401A volt meter it is possible to read the PRT to a part in 10^5, corresponding to 0.003°C; a much better resolution than its accuracy. Thus, an accurate measurement of the Johnson noise of a resistor in thermal equilibrium with the PRT could be used to calibrate the thermometer; the noise measurement becomes a temperature measurement, on the assumption that Boltzmann's constant is known.

- Random errors

Each measurement of the power spectrum represents a finite sample from an infinite population of such measurements and the parameters derived from the measurement are themselves subject to random errors. Johnson noise follows the normal distribution and since the Fourier transform represents a linear operation, both the real and imaginary part of the transform are also normally distributed. Therefore, the power spectrum, which is the sum of the squares of these, follows the chi-square distribution with two degrees of freedom.

A power spectrum averaged over N samples follows the chi-square distribution with 2N degrees of freedom, which approaches the normal distribution for large N. In this case, the standard deviation of the estimate of the power spectrum in each frequency interval is proportional to the power spectrum of the parent distribution, with $\sigma_S = \frac{S}{\sqrt{N}}$. A large number of averages and therefore a large measurement time is required to obtain high precision.

To quantify this, it is significant that this standard deviation applies independently to each point in the spectrum. Thus, if the spectrum is "white" and consists of M frequency bins, the uncertainty is reduced to $\sigma_S = \frac{S}{\sqrt{MN}}$. This would suggest making M as large as possible to reduce the required number of averages.

However, the number of frequency bins determines the frequency resolution $\delta f = \frac{f_N}{M} = \frac{f_S}{2M} = \frac{1}{\tau}$, where f_S is the sampling frequency, F_N the corresponding Nyquist frequency and τ the duration of each sample.

Thus, the relative uncertainty is $\frac{\sigma_S}{S} = \frac{1}{\sqrt{N\tau\, f_N}}$,

where $N\tau$ represents the total duration of the measurement. It is unimportant how this is apportioned between the number of samples and the length of each sample. To achieve an accuracy of 3 digits (0.1%) would require $N\tau\, f_N = 10^6$, that is a duration of about a minute if f_N is 20 kHz, the upper limit of the audio spectrum.

Several practical factors determine the maximum useable frequency. The most fundamental, as discussed in section 4, arises from the frequency dependence, which is due to the unavoidable stray capacitance at the input of the LNA. For the LNA used in this experiment, this limit is at about 5kHz (the amplifier also has a low pass filter at this frequency). Thus, it would take about 16 minutes to achieve an accuracy of 0.1%.

The situation is further exacerbated by the fact that it takes the computer a finite time to perform the Fourier transform, during which time no data are recorded, but with modern computers, this is not an issue.

Appendix A Data Acquisition

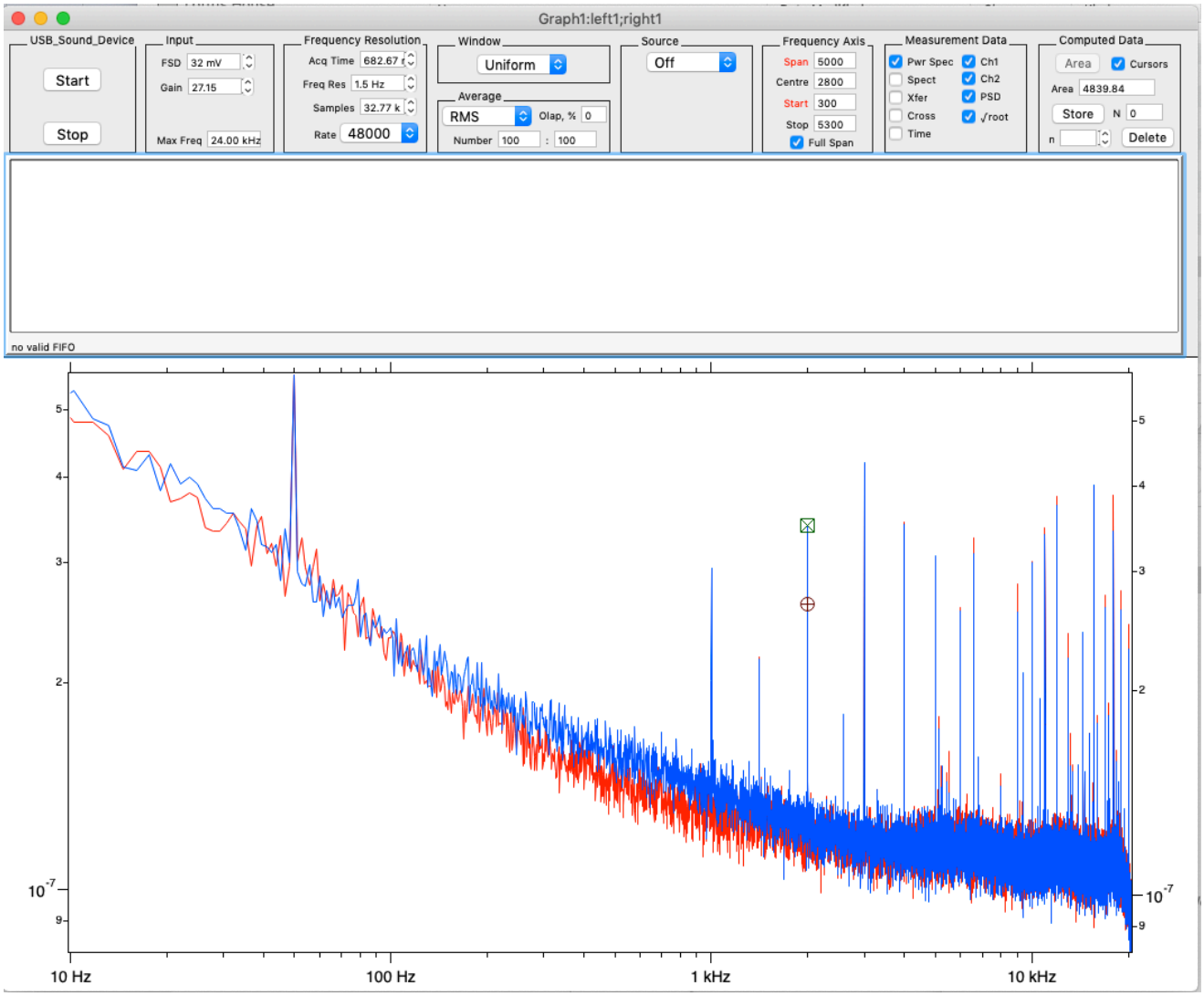


Figure 1 The spectrum analyser window.

Figure 1 shows an example of the spectrum analyser window. The grey region at the top is the control panel and the white area at the bottom shows a graph of the data, which you have chosen to display, and this is constantly updated whenever acquisition is active.

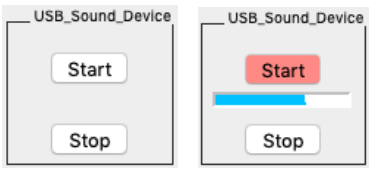
The control panel consists of two horizontal bands. At the bottom is a strip-chart, which shows live data as it is being acquired from the stereophonic sound input. The red trace shows the left channel, while the blue trace shows the right channel – and this "flat lines" in the case of a sound adapter with monophonic input.

At the very top is a row of boxes, which contain the controls. These are grouped into the boxes according to the nature of their function, and label at the top of each box indicates this function. The controls are dynamic and may change their colour or shape (or even disappear completely) depending on their settings.

For figure 1, the sound adapter is a Startech ICUSBAUDIO2D, with both inputs shorted and the plotted spectral density corresponds to the noise of the device, measured as an average of 100 samples. With the selected sensitivity, the noise above 2kHz is about 4nV/√Hz. Since the gain of the amplifier is 10000, this is many orders of magnitude smaller than the noise at its at its input.

The function of the controls is described group by group below.

USB_Sound_Device



This group controls the acquisition. The title corresponds to the name of the Audio Interface selected in the Sound Pane of System Settings and this must be selected before Igor is launched. To change it, you must

first Quit, select the device in the System Preferences, as described in appendix B and then re-launch Igor. The group contains two buttons and when acquisition is active, a progress bar.

Start

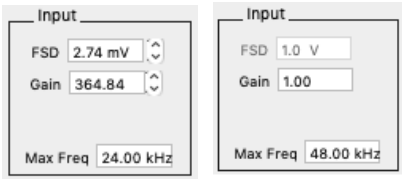
This button starts the acquisition of the selected number of averages. In the case of that averaging is set to either ***None*** or ***exp*** (exponential) the acquisition continues indefinitely. Otherwise it stops automatically when the desired number of averages is completed.

While acquisition is active, several changes can be seen. The ***Start*** button is coloured red and a progress bar is made visible, which indicates the time remaining until the next average is appended. Also, the strip chart is made active, with live data scrolling across the screen.

Stop

This button is used to stop the acquisition at any time. The progress bar disappears, the colour of the ***Start*** button returns to normal, and the strip chart becomes blank.

Input



This group determines the sensitivity of the input, but only for those audio devices, which support this function.

Max Freq

This displays the Nyquist frequency corresponding to the currently selected sampling rate.

Gain

This displays or sets the sensitivity of the input and exists in two states, depending on the sound adapter. The state, where the computer can set sensitivity of the adapter, is shown on the left. The sensitivity can be specified in two different but equivalent ways, in the form of a gain or FSD (full scale deflection). The gain is the ratio of the amplitude of the digitised signal relative to that of the input signal, whereas the FSD is the maximum amplitude of the signal before it becomes saturated. The relationship is simply FSD = 1/gain. Both can be varied in discrete steps, whose number and size are dictated by the adapter. The spectrum analyser app includes a list of these for a number of different adapters, and if the selected adapter is not recognised, the controls are changed as shown on the right. Here, there in no table corresponding to the selected adapter and the user must perform a manual calibration, calculate the corresponding gain and enter this value into the panel. This then sets the corresponding FSD value. The same applies to those devices, whose gain physically cannot be controlled by the computer. This includes the “high end” audio adapters, which tend to have manual gain control knobs.

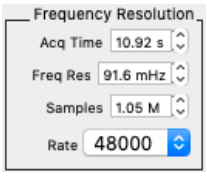
In all cases, the plotted data are scaled by the gain to reflect the actual measured voltages.

The small up-down arrows in the box on the left allow you to step through the available gains, possibly in quite small steps (for example 1 dB). Alternatively, typing a number into the gain box for these devices will search through the list for the nearest available discrete gain and display that value in the box. It also sets the corresponding FSD.

FSD

This displays the current value of the FSD setting, but only in the case that this can be computer controlled. Otherwise, it is blank.

Frequency Resolution



This group of controls sets the sampling frequency and the duration of the consecutive samples. The sampling rate is selected from those available with the audio adapter. Once this is selected, the duration of the sample (and therefore also the number of samples and the frequency resolution) is constrained by the fact that the number of samples must be an integer power of 2. There are natural limits to this. The maximum number of samples depends on the available memory and the time required to compute the transform. An arbitrary limit of about 1 million samples (actually 2^{20}) has been set at the high end, whereas at the low end, where the duration is short, the limit is set by the overheads involved in transferring the data and updating the graph. The lower limit for the number of samples corresponds to a minimum duration of about 0.5 seconds, and is adjusted so that the number of samples corresponds to the nearest power of 2. There are a discrete number of values of the sample size between the minimum and maximum number, because each must be a power of 2. Setting any one of the duration, the resolution or the number of samples automatically determines the others.

Window

This determines the window function, which is applied to the time series before computing the Fourier transform. The choices are ***Uniform***, ***Hanning*** and

Flat Top, which are identical to those used in the Hewlett Packard/Agilent/Keysight family of Dynamic Signal Analysers, as described in Appendix C.

Average

This group determines the nature and the size of the averaging. The options are ***None***, ***RMS***, ***RMS exp***, ***Vector*** and ***Vector exp***. RMS averaging is the most useful and is described in Appendix A. It uses a fixed number of samples, each with equal weight to compute the average. The ***exp*** (exponential) variant is similar, except the number of samples in indefinite and the older samples are given less weight than the more recent ones. Vector averaging requires that the sampling is synchronised with the signal, but this is not really possible using digitisers designed for audio applications.

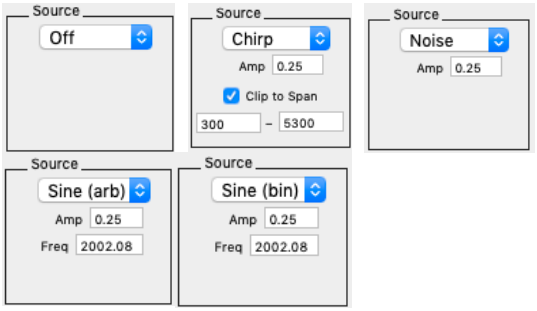
Number

The numerical value in the first box sets the desired number of averages for the non-exponential variants. For exponential averaging, it defined the time scale over which older samples are given less weight. The second numerical value, separated by the colon counts the number of averages, which have taken place.

Olap, %

This stands for ***Overlap***. When using a window function, the data at the beginning and end of the sample are given a lower weight and are in a sense "wasted" and setting an overlap reuses these. An overlap of 50% is commonly used with the Hanning window. See appendix C for details.

Source



This group determines the nature of the audio output signal. This is commonly used for characterising the frequency response of a system, and a number of different signals are available for this, namely ***Off***, ***Chirp***, ***Noise***, ***Sine (arb)*** and ***Sine (bin)***.

Chirp

This produces a periodic chirp, with an option to limit its frequency band.

Noise

This produces a Gaussian noise over the full frequency band of the sampling.

Sine (arb)

This produces a sinusoidal signal of any arbitrary frequency.

Sine (bin)

This produces a sinusoidal signal whose frequency is rounded to the nearest bin in the Fourier transform.

Amp

This sets the amplitude of the output signal. For Gaussian noise, it is three times the standard deviation of the noise.

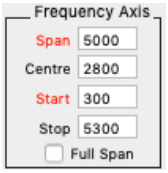
Frequency

This sets the frequency of the output signal for the two sinusoidal options. In the case of ***Sine (bin)*** the entered frequency is rounded to the nearest bin of the Fourier transform

Clip to Span

This applies only to the periodic chirp, and limits the frequency band as specified by the two numerical values immediately below the checkbox. By default, there are the stop and start frequencies of the graph but they can be independently changed. The default values are restored whenever ***Chirp*** is selected.

Frequency Axis



This group determines the horizontal scale of the graph. It does not alter any acquisition settings. The four parameters are not independent and are fully specified by any two. For convenience, three pairs are defined, which work logically together. The active pair is determined automatically and current pair is coloured in red. The pairs are Span & Start, Span & Centre and Start & Stop. A different pair is selected whenever a parameter is edited, which is not a member of the current pair.

Span

This sets the length of the frequency axis.

Centre

This sets the centre of the frequency span.

Start

This sets the left hand end of the frequency axis.

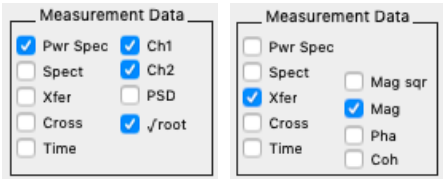
Stop

This sets the right hand end of the frequency axis.

Full Span

When this checkbox is ticked, the frequency specifications are ignored and the entire axis is shown.

Display



This group determines what is plotted on the vertical axes of the graph. There are two independent columns of checkboxes, which together determine the type and the units of the displayed traces. The left hand column is fixed, but the right hand column is altered according the item chosen in the left hand column.

Pwr Spec

This plots the rms averaged spectrum of channel 1 on the left axis and channel 2 on the right axis, as determined by the corresponding ***Ch1*** and ***Ch2*** checkboxes. The spectrum can be displayed as an amplitude or spectral density as determined by the PSD checkbox.

PSD

When ticked, the Power Spectral Density (PSD) is shown and otherwise the amplitude spectrum is shown.

√root

When ticked, the square root of the either the PSD (in units of Vrms/√Hz) or the amplitude spectrum (in units of Vpk) is shown. Otherwise, the square of these is shown ($\left\langle V^2 / Hz \right\rangle$ or V_{pk}^2).

Spect

This plots the linear spectrum of channel 1 on the left axis and channel 2 on the right axis, as determined by the corresponding ***Ch1*** and ***Ch2*** checkboxes. Linear spectra are not averaged in RMS mode and only the most recent sample is plotted.

Xfer

This plots the transfer function, with magnitude or magnitude squared on the left axis and phase or coherence on the right axis, as determined by the ***Mag Sqr***, ***Mag***, ***Pha*** and ***Coh*** checkboxes.

Cross

This plots the Cross Power Spectrum, with magnitude on the left axis and phase on the right axis, as determined by the ***Mag*** and ***Pha*** checkboxes.

Time

This plots the time series of channel 1 on the left axis and channel 2 on the right axis, as determined by the corresponding ***Ch1*** and ***Ch2*** checkboxes. Time series are not averaged in RMS mode and only the most recent sample is plotted.

Computed data

This group of controls performs additional processing of the acquired data.

Area

This button computes the area of certain traces displayed in the graph, optionally either over the entire range of the displayed frequency axis or that specified by the positions of the cursors. The button is enabled for only certain combinations of the options selected in Measurement Data, either ***PSD*** (and only if √ square root is ***not*** selected) or ***Xfer*** (and only if ***Mag sqr*** - magnitude squared - ***is*** selected). For all other combinations it is disabled.

Cursors

When ticked, the calculation of the area uses the cursors to determine the limits of the integration. Otherwise, the entire displayed frequency axis is used.

Store

This button stores the current data for subsequent use. All of the data from the current measurement are stored and given a unique identifying number, and the counter (***N***) is incremented to correspond to the number of stored data sets.

delete

This button deletes the data corresponding to the index number (***n***), and then decrements the count (***N***). If necessary, ***n*** is adjusted to be between 1 and ***N***.

N

This numerical indicator displays the count of the stored data sets. It cannot be edited.

n

This control sets the index, which points to one of the stored data sets. Whenever new data are acquired, this indicator is made blank, but if a change is made (either by using the up-down arrows or by typing a valid value) then the corresponding data are displayed in the graph, overwriting the currently displayed data.

It is easy to tell whether the graph shows live date or stored data, because the index is blank whenever live data are displayed. If the index shows a number, then that represents the index of the displayed data.

Appendix B Setting the System preferences for the sound.

Figure A1 The sound adapter

Figure A1 is a photograph of the Startech ICUSBAUDIO2D sound adapter, with the input and output connectors clearly labelled. Both are stereo, with two channels in a single connector. You can use a cable with a mating stereo connector at one end and a pair of single-channel connectors at the other. These are usually coloured white for the left and red for the right channel as shown in figure A2.

Figure A2 The sound cable with the 3.5mm stereo plug at one end and the two RCA plugs for right and left at the other end

The ***System Preferences*** must be used to select this device for use. On the Mac, system preferences can be opened either in the Dock or by choosing the second item in the Apple menu.

Figure A3 The System Preferences Pane

Figure A3 shows the ***System Preferences***, with the ***Sound*** item highlighted in the red oval. Selecting this opens the control panel for selecting the sound device for both input and output.

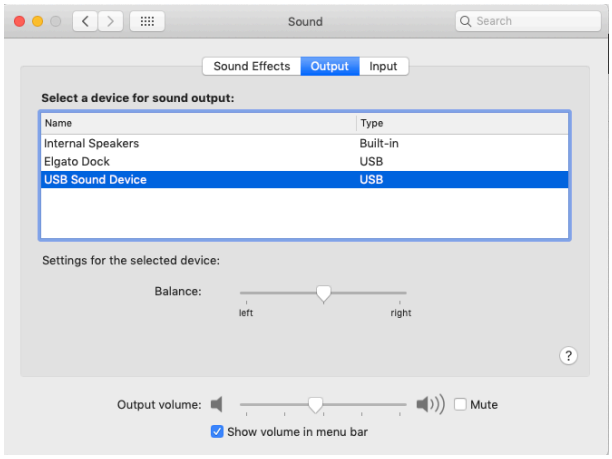


Figure A4 The output device selector

Figure A4 shows the panel when the output tab is selected. Here, you should choose the USB Sound Device as your output device.

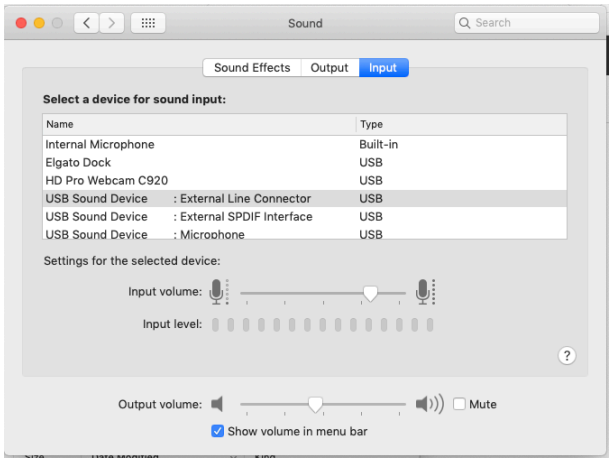


Figure A5 The input device selector

Figure A5 shows the panel when the input tab is selected. Here, you should choose the USB Sound Device: External Line Connector as your input device.

This tab also displays the gain sliders for both the input and output gains, labelled as “volume” with the icon of a microphone and loudspeaker respectively.

The input gain slider is set by the Audio DSA app and you can see it move as you step either the FSD or Gain controls.

The app cannot control the output volume slider and you should set it near its mid point. The volume control in this panel works together with the amplitude setting of source in the app, so that the actual output signal is proportional to the product.

Appendix C: Scaling and Windows in the Discrete Fourier Transform (DFT)

Frank van Kann, Department of Physics, The University of Western Australia, Perth, 6009 Australia

frank.van.kann@uwa.edu.au

1 Introduction

The following paragraph is quoted from the Igor Pro reference manual.

"No universal standard exists for the scaling of the FFT or for the sign of the imaginary component of the FFT. There are three commonly used methods for scaling. The method used by Igor is to not scale the FFT results but to scale the IFFT by 1/N. Another method is to scale the FFT by 1/N but not scale the IFFT. The third is to scale both by 1/√N. Both the scaling and the sign convention used by Igor are those used in "Numerical Recipes" due to the fact that this book has become a bible for most scientists and engineers."

Although the scaling method is arbitrary, it is essential to preserve the property of the transform that the reverse transform is the inverse of the forward transform; a factor of N occurs when going "round the loop" to recover the initial data by applying both the forward and reverse transforms in succession.

The forward transform used by Igor is

$$\mathbf{X}_n = \sum_{k=0}^{N-1} x_k \, e^{\frac{j2\pi kn}{N}} \quad ,n = 0, \ldots N-1 . \qquad 1.1$$

This mapping can be expressed in the form

$$\mathbf{X}_n \Leftrightarrow x_k .$$

The input and output data are in general both complex arrays with N points labelled 0, 1, ... N-1. The input data are assumed to be sampled at regular time intervals δt over a total time duration T

$$T = N\,\delta t = \frac{1}{\delta f} = \frac{N}{f_S} ,$$

where f_S is the sampling frequency. This is

$$f_S = \frac{1}{\delta t} = \frac{N}{T} = N\delta f ,$$

where δf is the "bin width", or interval between samples in the spectrum.

The spectrum contains both positive and negative frequencies and the maximum frequency represented is the Nyquist frequency,

$$f_N = \frac{f_S}{2}$$

A number of different scaling equations are discussed in sections 3 and 4 and are summarised in the section 2.

The definitions of the linear and power spectra and also of the transfer and coherence functions are the same as those used in the Keysight/Agilent (originally Hewlett Packard) spectrum analysers and dynamic signal analysers, including the 3562A, 3582A, 35665A 35670A and 89410A. The HP3582A represented ground-breaking technology on several fronts and was described in the HP Journal in September, 1978.

2 Summary

Given two time series x_k and u_k, the following can be defined.

$\mathbf{X}_n$ "raw" FFT of x_k as defined in equation 1.1

$\mathbf{U}_n$ "raw" FFT of u_k as defined in equation 1.1

$\mathbf{Y}_n$ scaled double sided linear spectrum of x_k

$\mathbf{V}_n$ scaled double sided linear spectrum of u_k

$\mathbf{S}_n$ double sided power spectral density of x_k

X_n single sided "raw" FFT of x_k

U_n single sided "raw" FFT of u_k

Y_n scaled single sided linear spectrum of x_k

V_n scaled single sided linear spectrum of u_k

S_n single sided power spectral density of x_k

H_n single sided transfer function between u_k and x_k

C_n single sided coherence function between u_k and x_k

Table 2.1 summarises the scaling of the respective spectra, where the angle brackets $\langle\ \rangle$ represent a frequency-by-frequency average taken over an infinite ensemble of identical systems. This has practical implications, which are discussed in section 4.

Table 2.1 *Summary of the scaling relationships*

Double Sided Spectra

Linear Spectrum

$$\mathbf{Y}_n = \frac{\mathbf{X}_n}{N}, 0 \le n \le N-1$$

Power Spectrum

$$\mathbf{S}_n = \left\langle \frac{|\mathbf{X}_n|^2}{N^2 \delta f} \right\rangle = \left\langle \frac{|\mathbf{Y}_n|^2}{\delta f} \right\rangle, 0 \le n \le N-1$$

Single Sided Spectra

Linear Spectrum

$$Y_n = \frac{2\mathbf{X}_n}{N}, 1 \le n \le \frac{N}{2} - 1$$

$$Y_n = \frac{\mathbf{X}_n}{N}, n = 0, \frac{N}{2}$$

Power Spectrum

$$S_n = \left\langle \frac{2|\mathbf{X}_n|^2}{N^2 \delta f} \right\rangle, 1 \le n \le \frac{N}{2} - 1$$

$$S_n = \left\langle \frac{|\mathbf{X}_n|^2}{N^2 \delta f} \right\rangle, n = 0, \frac{N}{2}$$

Windowed Single Sided Spectra

Linear Spectrum

$$Y_n = \frac{2 w_A \mathbf{X}_n}{N}, 1 \le n \le \frac{N}{2} - 1$$

$$Y_n = \frac{w_A \mathbf{X}_n}{N}, n = 0, \frac{N}{2}$$

Power Spectrum

$$S_n = \left\langle \frac{2 w_A^2 |\mathbf{X}_n|^2}{w_S N^2 \delta f} \right\rangle, 1 \le n \le \frac{N}{2} - 1$$

$$S_n = \left\langle \frac{w_A^2 |\mathbf{X}_n|^2}{w_S N^2 \delta f} \right\rangle, n = 0, \frac{N}{2}$$

For the Hanning window, $w_A = 2$ and $w_S = \frac{3}{2}$.

Causal Relationships between Two Signals

Transfer Function

$$H_n = \frac{\langle U_n X_n^* \rangle}{\langle |X_n|^2 \rangle} \quad \textbf{not} \quad \sqrt{\frac{\langle |U_n|^2 \rangle}{\langle |X_n|^2 \rangle}}$$

Coherence Function

$$C_n = \frac{\left| \langle U_n X_n^* \rangle \right|^2}{\langle |U_n|^2 \rangle \langle |X_n|^2 \rangle}$$

where

$$\langle U_n X_n^* \rangle = \lim_{M \to \infty} \frac{1}{M} \sum_{m=0}^{M-1} (U_n)_m \left(X_n^* \right)_m$$

$$\langle |U_n|^2 \rangle = \lim_{M \to \infty} \frac{1}{M} \sum_{m=0}^{M-1} \left| (U_n)_m \right|^2$$

$$\langle |X_n|^2 \rangle = \lim_{M \to \infty} \frac{1}{M} \sum_{m=0}^{M-1} \left| (X_n)_m \right|^2$$

3 Scaling

The appropriate scaling depends on the desired interpretation of the spectrum. If the requirement is to obtain the apparent amplitudes Y_n for sinusoidal signals, the spectrum should be scaled so that

$$\mathbf{Y}_n = \frac{\mathbf{X}_n}{N} \quad . \qquad 3.1$$

This gives the exact amplitude for sinusoidal signals whose frequencies are centred on the discrete frequency bins; the situation with other sinusoidal signals is discussed in section 7, Window Functions.

Such a requirement to obtain the amplitude of monochromatic signals is somewhat arbitrary and not directly related to the scaling of the continuous Fourier transform. In that case, the transform of a monochromatic signal is a Dirac delta function, with infinite amplitude. However, The average over any finite frequency interval, which contains the frequency of the signal, is equal to the amplitude. Thus, the scaling of the discrete transform is equivalent to the average value of the continuous transform within the frequency bin.

4 Power Spectral Density

The appropriate scaling for the power spectral density function can be obtained from Parseval's theorem in the discrete domain

$$\sum_{k=0}^{N-1} |x_k|^2 = \frac{1}{N} \sum_{n=0}^{N-1} |X_n|^2 \quad .$$

The mean power in the original continuous time signal is

$$\left\langle x^2 \right\rangle = \frac{1}{T} \int_0^T x^2 (dt) dt \approx \frac{1}{N} \sum_{n=0}^{N-1} |X_n|^2 \quad .$$

This suggests the definition for the discrete spectral density function S_n such that

$$\left\langle x^2 \right\rangle = \int_{-f_N}^{f_N} x^2 (t) dt \approx \sum_{n=0}^{N-1} S_n \delta f$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} |x_k|^2 = \frac{1}{N^2} \sum_{n=0}^{N-1} |X_n|^2 \quad .$$

This gives

$$S_n = \frac{|X_n|^2}{N^2 \delta f} = \frac{|Y_n|^2}{\delta f} \quad .$$

This is an estimate of the power spectral density from a single sample of the signal and has a large statistical uncertainty, since the standard deviation of this estimate is equal to the magnitude of the estimate itself.

This uncertainty can be reduced by averaging, using a method known as Welch's modified periodogram, which takes consecutive samples of the signal to compute the average. It is only an approximation of the spectral density, which is defined as the average over an infinite ensemble of identical systems. Since an infinite ensemble (or even a large but finite ensemble) is not generally available, it is necessary to compute the average using a succession of consecutive samples taken from a single system, which is possible provided that the system is stationary.

$$\mathbf{S}_n = \left\langle \frac{|\mathbf{X}_n|^2}{N^2 \delta f} \right\rangle = \left\langle \frac{|\mathbf{Y}_n|^2}{\delta f} \right\rangle$$

The angle brackets denote a pseudo ensemble average taken over the consecutive subsamples, so that

$$S_n = \frac{1}{M \delta f} \sum_{m=0}^{M-1} \left| (Y_n)_m \right|^2 \quad .$$

For the average, the time series of total length L sample points has been divided into M subsamples, as depicted in figure 4.1. These samples may overlap, which is useful to avoid loss of precision in the case that window functions are used (see section 7).

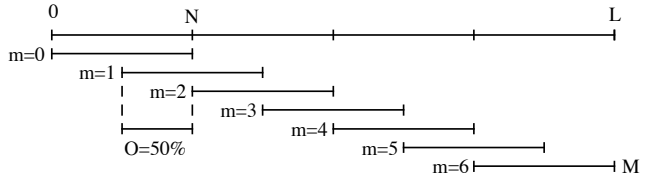


Figure 4.1 Illustration of subdivision of the time record into overlapping segments

The number of segments is given by

$$M = \frac{L - O N}{(1 - O) N} \quad ,$$

where O is the overlap, expressed as a fraction of N.

The index m in figure 4.1 represents the number of the sub-sample.

5 Coherence and Transfer Functions

The causal relationship between two signals is conveniently displayed in the form of a transfer function, exemplified by the block diagram in figure 5.1, which shows the relationship between x_k and u_k in the time domain and X_n and U_n in the Fourier domain.

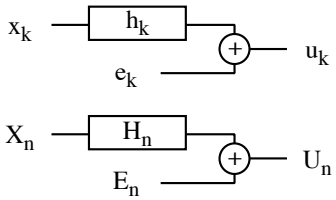


Figure 5.1 Block diagram depicting a transfer function in both the time and Fourier domains

The functional relationship between the signals is

$$u_k = h_k \otimes x_k + e_k$$

in the time domain and

$$U_n = H_n \cdot X_n + E_n ,$$

in the Fourier domain, where x_k is the input signal, u_k the output signal, $\otimes$ represents the convolution operator and e_k is random noise, which is added to the output signal. The transfer function is represented by h_k in the time domain and H_n in the Fourier domain.

Because of the added noise, the simple ratio of the measured rms-averaged transforms

$$\sqrt{\frac{\left\langle |U_n|^2 \right\rangle}{\left\langle |X_n|^2 \right\rangle}}$$

is likely to be a poor estimate of the magnitude of H, since averaging does not reduce the noise but only reduces the variance of the noise. Moreover, this ratio contains no information about the phase of the transfer function.

A better estimate of H is obtained from

$$H_n = \frac{\left\langle U_n X_n^* \right\rangle}{\left\langle |X_n|^2 \right\rangle}$$

and here averaging reduces the noise provided that the noise e is not correlated with the signal x. In reality H is an estimate of the transfer function, based on a finite number of averages. Since H is complex, the phase of the transfer function is also obtained.

No scaling is required in this calculation, since the result is a dimensionless ratio. Neither is the distinction needed between single and double sided spectra (see section 6); the equations are the same for both.

The added noise degrades the causal relationship between the signals x and u. This is quantified by the coherence function, given by

$$C_n = \frac{\left|\left\langle U_n X_n^* \right\rangle\right|^2}{\left\langle |U_n|^2 \right\rangle \left\langle |X_n|^2 \right\rangle} \approx \frac{\left\langle |U_n|^2 \right\rangle}{\left\langle |U_n|^2 \right\rangle + \left\langle |E_n|^2 \right\rangle},$$

where C_n is the estimate of the coherence, based on the finite number of averages M. Again, this is a dimensionless ratio and no scaling is needed. The approximation holds for large M, where C_n is essentially the signal to noise ratio of the output signal U_n.

6 Real Functions and Single Sided Spectra

If the input array is real then the negative frequency part of the output spectrum can be obtained from the positive part

$$X_{-n} = X_{N-n} = X_n^* .$$

In this case, the negative frequency part of the spectrum contains no additional information and is simply the positive part "folded back". This is illustrated in the following example using the function

$$x_k = x(t_k) = A \sin 2\pi f_1 t_k + B \cos 2\pi f_N t_k + C\, g(t_k) + D ,$$

where $t_k = k\,\delta t$, with k = 0, 1, ... N-1 and g(t) represents Gaussian noise with unit standard deviation. The parameter values are N = 256, A = 1, B = 1, C = 0.5, D = 0.75 and $f_1 = f_N/4$ (chosen to be centred on a frequency bin).

The corresponding linear amplitude spectrum Y_n is plotted in figure 6.1.

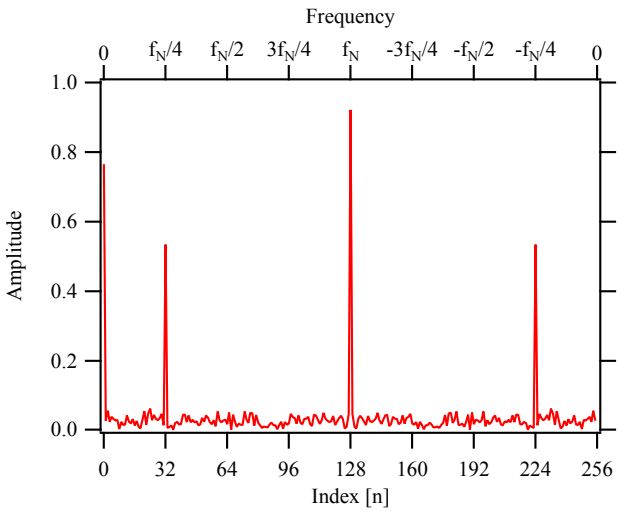


Figure 6.1 The linear amplitude spectrum of x(t).

The signal at frequency f_1 appears twice, firstly at n = N/8 = 32 and again at the folded position n = N - N/8 = 224, which corresponds to the negative frequency component $-f_1$ in the continuous time domain. The DC component occurs only once at n = 0; the folded

back position at n = N corresponds to DC in the next period of the spectrum. The n = N/2 position is also unique as it is folded back onto itself. This bin is occupied by the signal at the Nyquist frequency, which must be purely real, since $X = X^*$ there; only cosine waves at the Nyquist frequency can be represented in the transform.

Using the double sided scaling of equation 3.1, the magnitude of each monochromatic signal is A/2, where A is the amplitude at that frequency, except for the signals at DC and the Nyquist frequency, which have magnitude D and B respectively, equal to their respective amplitudes.

A more intuitive way to represent the spectrum is shown in figure 6.2. Here the data are rotated so that f = 0 is in the centre and the negative frequencies are to the left.

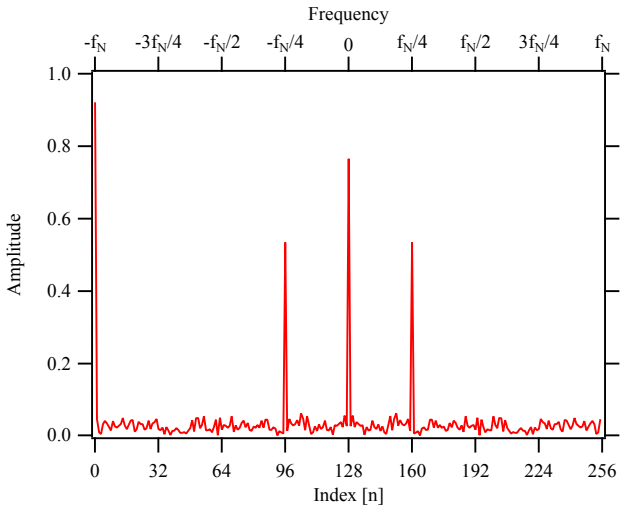


Figure 6.2 The rotated linear amplitude spectrum of x(t).

7 Single Sided Spectra

The transform of a real function is conveniently represented by the single sided spectrum X_n, which is obtained by taking the first N/2 + 1 points of $\mathbf{X}_n$ and discarding the remainder (the + 1 is to include the signal at the Nyquist frequency). Igor uses this approach when computing Fourier transforms of real functions. If the input array is real then N must be even (and ideally an integer power of 2 for computational efficiency) and only the first N/2 + 1 points of the complex transform are given.

In this case, the appropriate scaling is

$$Y_n = \frac{2\mathbf{X}_n}{N}, 1 \le n \le \frac{N}{2} - 1 \qquad 7.1$$

and

$$Y_n = \frac{\mathbf{X}_n}{N}, n = 0, \frac{N}{2} \qquad 7.2$$

where a special rule is required for the n = 0 or DC component, which is represented only once in the full spectrum and the n = N/2+1 or the component at the Nyquist frequency.

In most practical situations, the fact that only the cosine component at the Nyquist frequency is included is not a serious limitation, since an anti-alias filter would ensure that there is no significant power at that frequency.

The single sided linear spectrum is shown in figure 7.1.

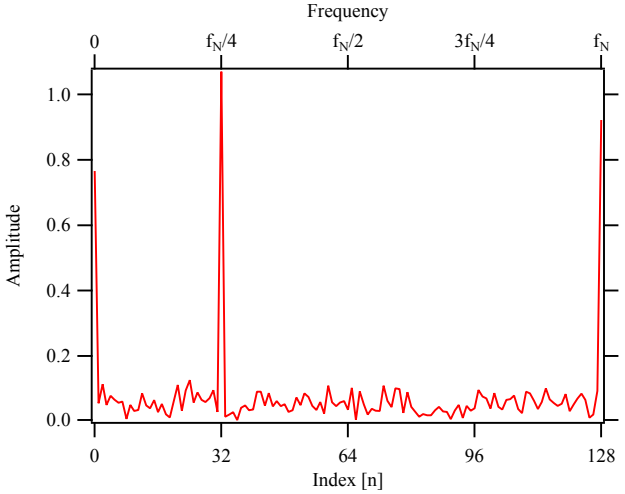


Figure 7.1 The single sided linear amplitude spectrum

The corresponding scaling for the spectral density is

$$S_n = \left\langle \frac{2|\mathbf{X}_n|^2}{N^2 \delta f} \right\rangle = \left\langle \frac{|Y_n|^2}{2\delta f} \right\rangle, 1 \le n \le \frac{N}{2} - 1$$

and

$$S_n = \left\langle \frac{|\mathbf{X}_n|^2}{N^2 \delta f} \right\rangle = \left\langle \frac{|Y_n|^2}{\delta f} \right\rangle, n = 0, \frac{N}{2}.$$

At the end points, both the linear spectrum and the power spectral density are scaled differently in the single sided case.

8 Windows

In the example shown in figures 6.1 to 7.1, care was taken to ensure that the sinusoidal signal was centred on a frequency bin in the transform. In such a case, the spectrum for a pure sinusoid is a close approximation to a Kronecker delta function, as shown in figures 8.1a and 8.1b.

The function chosen for this example is a pure sinusoid

$$x_k = x(t_k) = A \sin 2\pi f_1 t_k$$

with A=1 and $f_1 = f_N/4$ as before but with B = C = D = 0. The spectrum is plotted on both a linear scale and also a log scale to magnify the baseline of the curve, which is zero within the numerical precision of the computation.

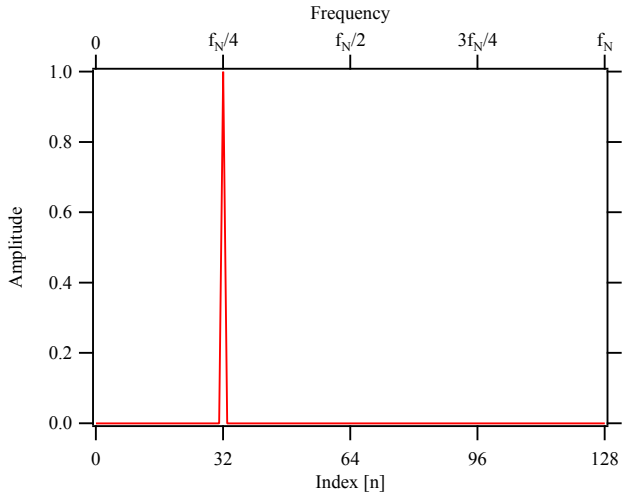


Figure 8.1a Spectrum of a pure sinusoid centred on a frequency bin plotted on a linear scale.

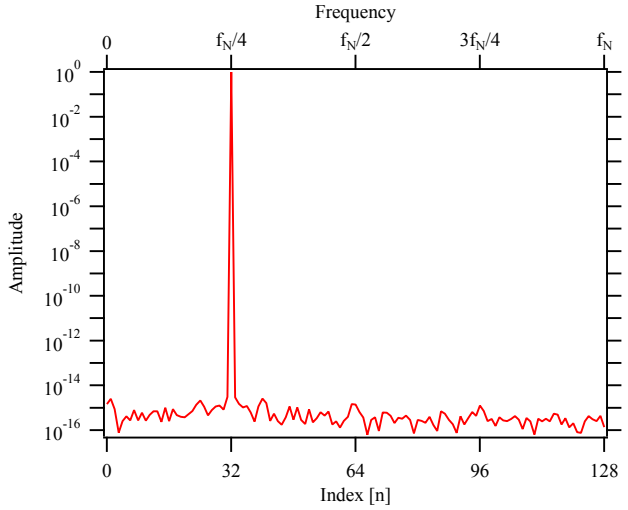


Figure 8.1b Spectrum of a pure sinusoid centred on a frequency bin plotted on a log scale.

If the signal is not centred in a frequency bin, this produces an effect known as sideband leakage and is illustrated by the following example of a sinusoidal signal located mid way between two bins at $n = \frac{N}{8} + \frac{1}{2}$, equivalent to $f_1 = f_N\left(\frac{1}{4} + \frac{1}{N}\right)$.

The corresponding DFT is plotted in the upper (red) curve of figure 8.2, which shows how the energy of the monochromatic signal has been spread across the spectrum, limiting the dynamic range to about two decades - worse than the equivalent of 8-bit resolution.

This effect is reduced by use of a weighting or window function, which attenuates the time record near the end points to reduce the effect of the discontinuity. The Hanning window,

$$W(t) = \sin^2 \frac{\pi t}{T} \quad ,$$

is in common use and produces the improvement, which is also illustrated in figure 8.2, where the lower (blue) curve shows the amplitude spectrum of the windowed signal. Although the window function has broadened the peak relative to the ideal case of figure 8.1, the dynamic range in increased to well over 6 decades at frequencies away from the peak.

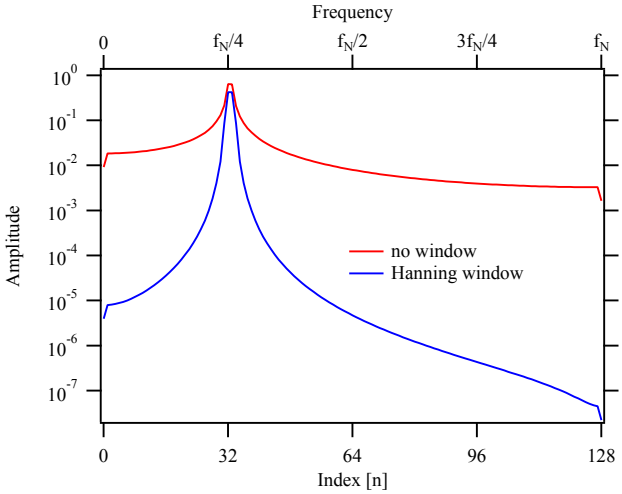


Figure 8.2 The effect of a window function for signals not on a frequency bin.

The two curves in figure 7.3 show the signal in the time domain before and after the Hanning window is applied.

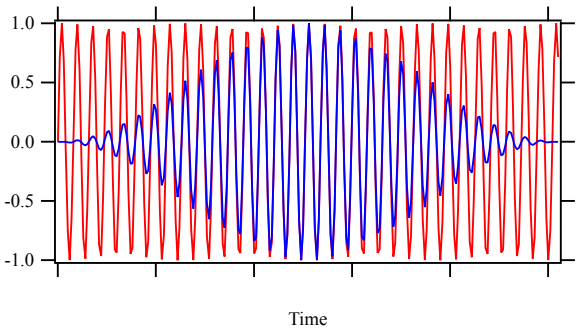


Figure 8.3 Time domain signal before and after the window is applied

Windowing is required whenever the signal is not periodic in the sample duration T. The signal is periodic if a copy of the sampled signal can be appended after the sampling interval without a discontinuity in either the value or the derivative. The sine function located between frequency bins as in figure 7.2 is an extreme example of a discontinuity in the derivative. A cosine function at the same frequency would have a discontinuity in the value. As illustrated in figure 7.3, the application of the window nulls the signal at both ends of the sample and ensures periodicity.

The implied periodicity in the time domain is a consequence of the discrete sampling in the frequency domain. It is analogous to the implied periodicity in the frequency domain produced by the finite sampling rate of the signal, which is characterised by the Nyquist frequency.

9 Special Cases Of Periodic Functions

Any linear combination of sine and cosine functions, whose frequencies are centred on a bin in the Fourier transform is periodic. There are three other important periodic functions, which in addition have the property that the magnitude of their Fourier transform is independent of frequency ("white"). These are the delta function ("impulse"), Gaussian noise and a periodic chirp. All of them are useful for experimentally probing the frequency response of a system, as outlined in section 5.

The impulse is mainly of theoretical importance, since the fact that the power of the signal is concentrated in a single sample places great demands on the linearity of the system under test. White noise is more favourable in this respect, since 99.7% of the samples lie within 3 standard deviations of the mean (usually zero). For the chirp, 100% of the points are within the specified amplitude. Window functions are not required for either the impulse or white noise, except in the case where the output signal is contaminated by incoherent noise or interference (which may be non-periodic). As described in section 5, averaging reduces the effect of these.

Windows should not be used with the periodic chirp, since the frequency content of the signal is a function of time and windowing would produce a non-white spectrum and therefore parts of the frequency domain are not probed. Because of this, a chirp is not suitable if the output signal is significantly contaminated by non-periodic interference and white noise would be a better choice.

10 Scaling of Windowed Spectra

It can be seen in figure 8.2 that the Hanning window reduces the amplitude of the spectrum. As illustrated in figure 10.1, signals centred on a frequency bin are reduced by exactly a factor of 2 compared to the un-windowed spectrum and this should be allowed for when scaling the spectrum.

The scaling appropriate for windowed amplitude spectra is then

$$Y_n = \frac{2w_A \mathbf{X}_n}{N}, 1 \le n \le \frac{N}{2} - 1$$

and

$$Y_n = \frac{w_A \mathbf{X}_n}{N}, n = 0, \frac{N}{2},$$

where w_A is the amplitude scaling factor, which depends on the window function used. For the Hanning window, $w_A = 2$ gives the proper amplitude spectrum for monochromatic or "narrow band" signals.

However, this overestimates the total power in the spectrum, because the window broadens the peak, increasing its area. This broadening also raises the amplitude of broadband signals such as random noise. This too is illustrated in figure 9.1, which shows the amplitude spectrum computed both with and without window of a signal consisting of Gaussian noise added to a pure sinusoid,

$$x_k = x(t_k) = A \sin 2\pi f_1 t_k + C\, g(t_k),$$

with C = 2 and A = 1 and with $f_1 = \frac{f_N}{4}$ as before to ensure the peak is centred in the frequency bin. The choice of f_1 centred on a bin is important as it permits the spectrum to be represented accurately without the need for a window and permits the effect of the window to be unambiguously demonstrated.

To reduce the "noise of the noise", the rms average of 64 consecutive spectra is shown.

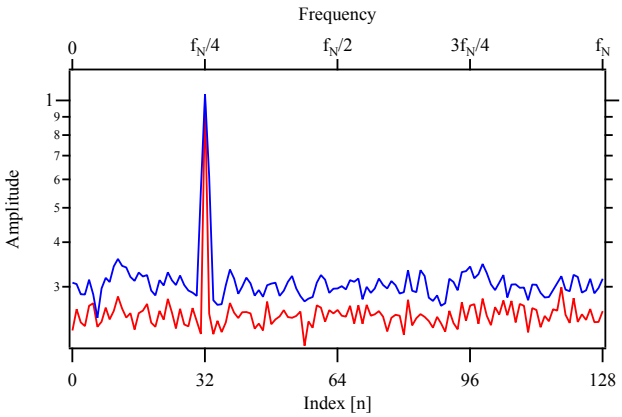


Figure 10.1 Amplitude spectrum of with and without window.

The amplitude scaling factor indeed gives the proper amplitude for the narrow band signal but its width is increased. Also, the amplitude of the noise is clearly seen to be increased by such overlap and the total power is over estimated. Parseval's theorem properly applies to the windowed signal and its transform but in practice it is desired to relate the spectrum back to the original un-windowed signal. The amplitude scaling factor satisfies this requirement to give the proper amplitude for narrow band signals but at the expense of incorrectly representing the power. The amplitude at each discrete point in the spectrum represents the total power in the corresponding frequency band. Without a window, this bandwidth is equal to the width of the bin but the window function increases the effective bandwidth and therefore the total power contributed by that frequency.

A different scaling is required to represent the spectral density of the noise when a window function is used. This depends on the window function and the appropriately scaled density for the Hanning window used in the above example is shown in figure 10.2.

$$S_n = \left\langle \frac{2w_A^2 |\mathbf{X}_n|^2}{w_S N^2 \delta f} \right\rangle, 1 \le n \le \frac{N}{2} - 1,$$

and

$$S_n = \left\langle \frac{w_A^2 |\mathbf{X}_n|^2}{w_S N^2 \delta f} \right\rangle, n = 0, \frac{N}{2}.$$

The required scaling is obtained with $w_A = 2$ and $w_S = \frac{3}{2}$,

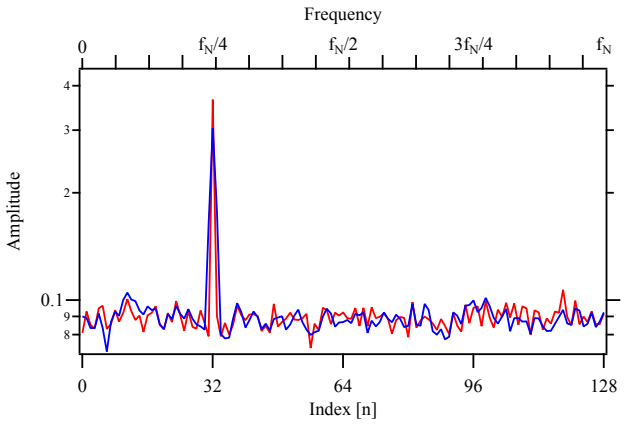


Figure 10.2 Scaled spectral density with and without window.

This gives the proper spectral density for the noise but underestimates the power of narrow band signals (because some of their energy has leaked into adjacent bins). This makes clear the distinction between narrow and broadband signals; a narrow band signal is one whose bandwidth is much less than the width of a bin and the amplitude spectrum properly represents its amplitude. Conversely, a broadband signal is much wider than the bin width and the spectral density properly represents its power.

The scaling factors are related to the identities

$$\frac{1}{\pi}\int_0^{\pi} \cos^2 x\, dx = \frac{1}{2} = \frac{1}{w_A}$$

and

$$\frac{1}{\pi}\int_0^{\pi} \cos^4 x\, dx = \frac{3}{8} = \frac{w_S}{w_A^2}.$$

There are many window functions described in the literature but the most useful examples are embodied in the firmware of the Keysight/Agilent/Hewlett Packard family of FFT spectrum analysers and dynamic signal analysers.

The window functions are

Uniform - that is, no window function; $w_A = 1$, $w_S = 1$.

Hanning - optimised for frequency discrimination (described above); $w_A = 2$, $w_S = 1.5$.

Flat top - optimised for amplitude accuracy; $w_A = 1$, $w_S = 3.811$. This window is proprietary to Keysite/Agilent/Hewlett Packard but has been reverse engineered to produce these examples.

Examples of these windows are shown in figures 10.3 and 10.4.

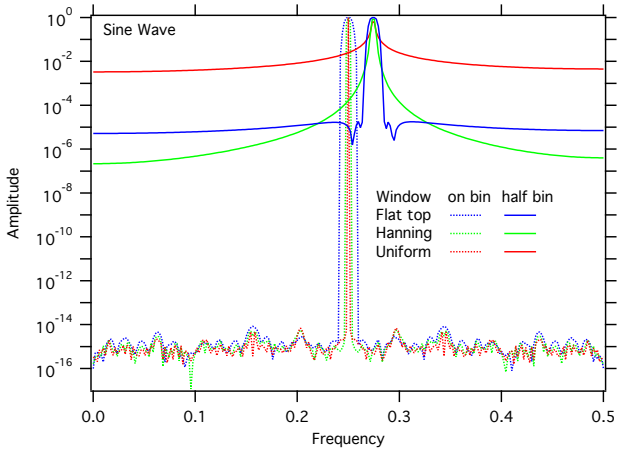


Figure 10.3 Examples of the three window functions with sine waves centred on a bin and situated in the mid point between two bins

Figure 10.3 shows the effect when each of the three windows is applied to two sinusoidal signals, one situated exactly on a frequency bin of the spectrum and the other located at the mid-point between two bins. Clearly, if the signal is centred, no window is required. Conversely, if the signal is between bins, either the Hanning or flat top produces superior results and the choice is dictated by whether frequency resolution or amplitude accuracy is desired. This is made clearer by figure 10.4, which shows a zoomed-in view of the peaks in figure 10.3. At worst, the flat-top has an amplitude error of 0.05% but has 3 times poorer frequency resolution than the Hanning. However, The Hanning has an amplitude error of about 15%. The uniform, the Hanning and the (reverse engineered) flat top windows are incorporated into the software package described in the next section.

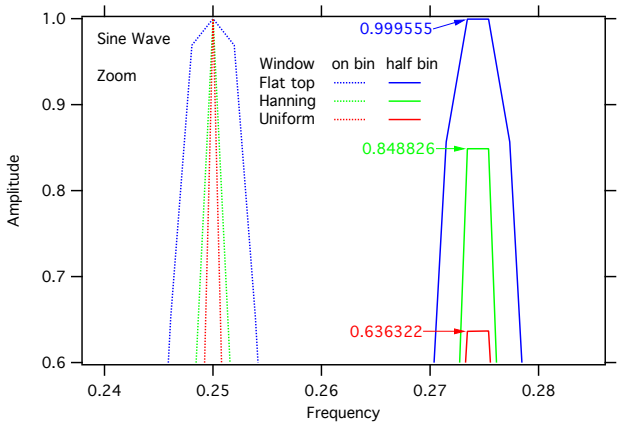


Figure 10.4 Zoomed-in view of the peaks in figure 9.3

Appendix D – measuring $\left\langle V_{out}^2 \right\rangle_{f_1}^{f_2}$

The value of $\left\langle V_{out}^2 \right\rangle_{f_1}^{f_2}$ is calculated from the measured Fourier transform, as described below.

Parseval's theorem describes the relationship between the power of a signal x(t) and that of its Fourier transform X(f).

$$x^2\Big|_{f_1}^{f_2} = \int_{f_1}^{f_2} |X(f)|^2 \, df \ ,$$

where $x^2\Big|_{f_1}^{f_2}$ is the value of the square of the signal after being passed through an infinitely sharp band-pass filter between f_1 and f_2. While such a sharp filter acting on the actual signal x(t) would be impossible to realize, it is easy using the Fourier transform, simply by specifying the limits of the integral when applying Parseval's theorem.

In practice, we perform a discrete Fourier transform so the integral is replaced by a summation over the corresponding frequency bins of the discrete transform. Also, we are interested in the mean square voltage rather than its instantaneous value, giving the desired measurement as

$$\left\langle V_{out}^2 \right\rangle_{f_1}^{f_2} = \left\langle \frac{V^2}{\delta f} \right\rangle \sum_{(f_1)}^{(f_2)} |A(f_i)|^2 \, \delta f \ ,$$

where we have replaced the integral by the summation and have taken the average over a number of consecutive samples of the Fourier transform. Since the amplifier gain doesn't change between samples, it can be taken out of the average as a constant.

We can recognize the first term on the right hand side from Nyquist's theorem

$$\left\langle \frac{V^2}{\delta f} \right\rangle = 4 k_B R T$$

and the second term is simply Δf_e.

Appendix E – Amplifier Noise

This section is for interest only. It is not used in the experiment. Also, it refers to the concept of a power spectral density, which is not yet familiar.

The amplifier itself contributes some additional noise, which must be considered. A commonly used model of amplifier noise is shown in figure E1 and consists of a voltage noise source V_A in series with the input and a current noise source I_A in parallel with the input.

The noise sources have spectral densities S_V and S_I respectively, which are usually assumed to be "white", or independent of frequency as well as independent of each other and of the impedance of the circuit connected to the input.

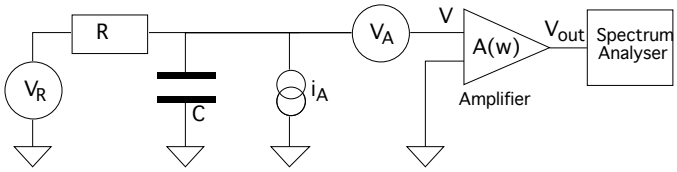


Figure E1 *The noise measurement circuit with the amplifier noise sources included.*

For any given amplifier, these assumptions about the characteristics of the noise should be verified experimentally. The output of the amplifier due to the voltage noise alone is simply $A(f)V_A$ and its spectral density S_V can easily be measured by "short circuiting" the input. The current noise is not easy to measure directly but can be determined as described in section 5.

In the ideal amplifier, where the input impedance is very large, the noise current flows mainly in the external circuit. Ignoring the effect of C, this produces a voltage given by $V = i_A R$ at the input of the LNA (in addition to the voltage noise V_A). There is a value of R for which $V_A = i_A R$, when the effect of current noise and voltage noise are equal. This is the so-called noise matched resistance, given by

$$R_A = \sqrt{\frac{S_V}{S_i}} \ .$$

For the measurement of Johnson noise, this gives the maximum possible signal to noise ratio.

Using such a resistor permits Johnson noise to be measured (with unity signal to noise ratio) for a particular temperature of the resistor, which is given by

$$4 k_B T R_A = S_V + S_i R_A^2 \ .$$

This is $T_A = \dfrac{\sqrt{S_V S_i}}{2 k_B}$, which is the so-called noise temperature of the amplifier.

The University of Western Australia

Department of Physics

Second Year Laboratory

Johnson Noise

Appendix F Recipe sheet

The aim of the first week's experimentation is to measure the Johnson noise for a single resistor (supplied) at two temperatures, room and liquid nitrogen.

The resistor (nominally 100kΩ) is permanently connected to a short length of coaxial cable and the assembly is terminated at one with a BNC connector. At the other end, the resistor is fully enclosed by the braid and is visible only as a slight bulge.

The BNC plug can be plugged directly into the input socket if the LNA.

There are also a few examples where the coaxial cable is terminated at **both** ends with a BNC connector. At the end furthest from the resistor is a BNC plug, which can be plugged directly into the input socket if the LNA. At the other end is a BNC socket. The ends of the resistor are connected between the centre pins of the plug and respectively and a "shorting" link (supplied) is needed to measure Johnson noise. This arrangement, although it's a distraction for the simple noise measurements, permits the amplifier to be calibrated. We have already done that, so the students don't have to worry about it - except to utilise the results we provide.

So, to measure Johnson noise:

1) Plug the resistor into the input of the LNA.
2) Make sure the shorting link is connected to the other end (if it's one of those).
3) Plug the output of the LNA into the sound input (either Left or Right - or both).
4) Use the sound control panel to ensure the source is set to "Line in" and if there is a level slide control for the input, make sure it's at maximum.
6) Launch the Igor spectrum analyser.
7) Set the input to the most sensitive range which doesn't saturate the input - you need to press Start to get a sample to check this. The Chart Graph window shows the raw input, which must be within ±1 without clipping.
8) Measure the spectral density - use a large number of averages! (See the notes on how long you need for 1% accuracy).
9) Move the round cursor to 400Hz and the square cursor to 5kHz and find the area of the spectral density. Record this number in the lab book.
10) Repeat 8 and 9 with the input of the amplifier shorted and measure the spectral density and area of its voltage noise.
11) Use the data to perform a spot calculation of k_B and compare with accepted value. (Remember to subtract the voltage noise from the Johnson noise data.)
12) Repeat at liquid nitrogen.
13) Plot a graph to find both k_B and S_i .
14) Probably, k_B won't agree with expected value. You will need to calibrate the sound input. Feed in a sine wave of known amplitude and measure the peak. You could get the amplitude of the sine wave with a DMM in AC volts mode.